Université du Québec
Institut National de la Recherche Scientifique
Centre Eau Terre Environnement

Adaptation d'outils de modélisation pour l'analyse d'impact sur le régime thermique des rivières dans le cadre de la gestion des barrages et du changement climatique

Par

Mostafa Khorsandi

Thèse présentée pour l'obtention du grade de

Philosophiae Doctor (Ph.D.)

en sciences de l'eau

Jury d'évaluation

Président du jury et examinateur interne	Alain Rousseau INRS-ETE
Examinateur externe	Anne-Marie Laroche Université de Moncton
Examinateur externe	Kevin Shook University of Saskatchewan
Directeur de recherche	André St-Hilaire INRS-ETE
Codirecteur de recherche	Richard Arsenault École de technologie supérieure

© Droits réservés de Mostafa Khorsandi, mai 2024

REMERCIEMENTS

As a water science/engineering student, I started my journey in 2006. As a learner, much earlier. After finishing my bachelor's in water engineering, I became a scientific nomad, traveling to several cities, countries, and research institutes to gain more knowledge in this beautiful field of study and research. After this long scientific transhumance worldwide, this dissertation is like my child, whom I wanted with all my heart and nurtured for four years. Now that this beautiful baby is well nurtured, I am ready to present it.

I wrote this dissertation, but it is the result of the efforts, cooperation, comments, and love of countless people to whom I am supremely grateful. Alas, in these limited lines, I only can thank a limited number.

First, I thank my supervisors, André St-Hilaire and Richard Arsenault, for their kindness, support, motivation, and cooperation, especially in difficult moments. Andre, I am incredibly grateful to you that, along with being my supervisor, you did for me what only a close friend would do for his dear friend and a caring human for beloved ones. This dissertation is written by me, but it is actually a reflection of your knowledge and ability to guide students. Richard, I am happy and honored to have you as my supervisor. Your help and support shaped a big part of this thesis.

I'm supremely grateful to my jury members and the reviewer of published papers for assessing my work and our publications based on the highest criteria level. This thesis and the published papers would not be possible without their insightful and constructive comments. I'm grateful to my internal and external jury members, who showed me the expectations of excellent work and encouraged me to fulfill them. I thank Stephen Dery for his positive feedback on my research proposal.

This honor and happiness have been given to me because I was born into an infinitely loving family. I am grateful to my dear mother "Zahra Khorsandi." The reason for my existence and my long-term study/research is because of her motivation to learn, going forward, cracking complex problems, and her curiosity, which was instilled in me. I am grateful to my mother for being a strong woman and a giving mom; and for teaching us the value of this world as a beautiful place to live. I am thankful to this extraordinary woman, the first hero of my life, for giving me a lifetime of unconditional love.

I want to mention my father, who, although he has not been among us for a long time, left us with values that are still alive and guiding me. Peace be upon him. I am grateful to my dear sisters Maryam and Monir for being more than my sisters, my close friends. Monir, what better gift than having you as a sister who is intelligent, knowledgeable, full of love, and goes beyond

herself to create beauty. I am incredibly grateful to my dear brothers Mohammad, Mahdi, Mortaza, and Shahram, who have been my first friends and teammates before anyone else. Together, we grew up loving each other, and now, our hearts beat for each other, no matter how far we are. Thank you to my nieces and nephews, Nika, Ahura, Termeh, Baran, Hesam, and Farnam, for allowing me to grow and watching them flourish. I wish I could be beside you more.

I am fortunate and grateful to have lovable, warm-hearted, and life-loving friends. In these limited pages, it's impossible to mention all the dear ones, so I'll settle for a limited number.

Ramin, for the long and deep friendship and for introducing Andre to me. Mostafa, my best friend, has been by my side since childhood. Pieter and Farshad, thank you for your unwavering support in starting this Ph.D. in Canada. Vahid, Mohammad-Mehdi, Farzin, Rasoul, Seyed-Mohammad, Seyed-Hossein, Mohsen, Jaber, Mohammad-Hossein, Hojjat, Hamid, Mohammad, Leah, Amir-Hossein, Sepideh, Ehsan, Mitra, Mehrdad, Elahe, Milad, Negin, Reza, Mohadeseh, Benjamin, Nazanin, Ali, Mahsa, Niloufar, Ehsan, Sepideh, Ali, Aram, Saman, Tayebeh, Jalal, Ursula, Zigfried, Eisinhower, Ilias, Veronique, Laureline, Anik, Muhammed, Suraj, Jean-Luc, Phillippe, Sally, Helen, Benoit, Eric, Amelie, and all my dear friends inside and outside Iran/Canada, who have shaped a beautiful part of my life, merci.

I am grateful to my dear roommates Anselm, Camilo, Pascal, and Julien for their support and friendship. To all my mountaineering friends in Iran and abroad who form my tribe, I express my gratitude and privilege to belong to this tribe.

The number of books I have read owes a debt to countless individuals who generously shared a lifetime of experience with others. I feel humility and deep respect in the face of the greatness of these writers. Throughout my life, I have been and remain indebted to dear librarians who have been my companions and friends in my quests. I am incredibly grateful to Mrs. Sichani, Mr. Samimi, dear Babak, and dear Jean-Daniel, librarians from various academic levels.

While completing this thesis, I lost some dear ones who were stars in the sky of my life. The great literary figure Mahmoud Andalib, the contemporary Iranian poet and my fatherly figure. Ali Reza-Gholi, my dear friend, whose heart beat for Iran and set a stream of thoughts and ideas for Iran's development. Peace be upon them.

I am thankful to my dear uncle Mahmoud, who, like a mountain of experience, taught me a lot. In my life, I have had many esteemed professors for whom I am incredibly grateful. Especially, I express my gratitude to (Professors and Doctors) Saied, Babak, Marzieh, Navid, Samaneh, Elaheh, Jörg, Alex, Klaus, Sadiqeh, Ahmad, and Lutfollah. Above all, I thank Andre and Pieter, who taught me how to live.

I am indebted to Dr. Saied Shah-Hoesseini, a dedicated ophthalmologist, for saving my vision. Thank you.

Mother Nature, the most important teacher, source of inspiration, and the reason for my curiosity, deserve the highest praise. Wherever I am on this planet, my heart beats for Iran. Iran is my reason for studying water science. Since 2006, I have decided to enter this field. My goal has been to do something to solve the water-related problems of this ancient, beloved, arid land and its people. Although far away, my thoughts, part of my research, and my heart are with you.

تن من پاره ای از آن تن توست *** چو ایران نباشد تن من مباد

I'm grateful to the First Nation land in Nechako (the traditional unceded territory of Wet'suwet'en Yintikh, Nadleh Whut'en, Stellat'en, Dakeł Keyoh, Dënéndeh), which is my case study, also Wendake-Nionwentsïo and Nitassinan (Innu) lands in which Quebec is located as my place to live, my shelter and the new home that I love. I appreciate their belonging since I am a global nomad with their values. As a scientific nomad, moving a lot and pursuing knowledge resources, I hope the readers will read and enjoy my work as I did doing that. Hard work with great hope in hard times pays off, and it was impossible without having a very unique person.

More than anything else, thank you, Andrée-Anne, the love of my life. How lucky I am to have you. This work was impossible without your love, support, and patience. Merci. Je t'aime.

RÉSUMÉ

La température de l'eau est l'une des variables critiques pour les évaluations de l'habitat du poisson dans les rivières. La modélisation hydrologique et thermique est utile pour fournir des estimations de la température de l'eau à l'échelle du bassin versant. Dans ce but, cette thèse propose des développements et des mises en œuvre de modèles sur le bassin versant de la rivière Nechako (Colombie-Britannique, Canada) pour obtenir de meilleures simulations de la température de l'eau. La thèse prend la forme de quatre articles de recherche. Les débits de la rivière Nechako sont partiellement contrôlés par le barrage Kenney et le déversoir du lac Skins. Rio Tinto, le propriétaire du barrage Kenney, utilise le modèle CEQUEAU pour la prévision du débit. La rivière Nechako est un couloir de migration vers les frayères pour certains poissons migrateurs. La présence du barrage, les impacts à long terme du changement climatique sur la température de l'eau et la sensibilité des poissons migrateurs aux changements thermiques font de cette région une excellente étude de cas pour le présent projet de recherche.

L'article 1 compare les méthodes de calage mono-site et multisite du module thermique du modèle CEQUEAU. Le calage multisite et la mise à l'échelle des paramètres n'avaient pas été utilisés jusqu'à présent pour calibrer un modèle thermique. Dans un premier temps, un calage hydrologique est effectué. Les mêmes résultats du modèle hydrologique étalonné sont utilisés pour tous les calages thermiques. Le calage local sur un seul site a fourni les meilleures performances pour les sites individuels, mais ces paramétragesont donné de mauvais résultats lorsqu'ils ont été appliqués aux autres sites. Le calage multisite et les méthodes de mise à l'échelle ont des performances adéquates pour tous les sites avec une racine de l'erreur quadratique moyenne ($RMSE_{global}$) < 2°C comme seuil acceptable. Cependant, la méthode multisite, pour laquelle toutes les mesures de sites différents sont combinées pour un seul calage, a donné les meilleures mesures de performance et le temps de calcul le plus court.

L'article 2 présente les effets des changements climatiques sur le régime thermique de la rivière Nechako. Le modèle hydrologique-thermique CEQUEAU a été calibré à partir des observations de débit et de température de l'eau. Les intrants météorologiques pour le modèle proviennent des réanalyses ERA5 pour le passé et les projections météorologiques des modèles climatiques pour les scénarios futurs. Le calage hydrologique a été réalisé pour la période 1980-2019 à l'aide des données de deux stations hydrométriques, et le calage de la température de l'eau a été effectué à l'aide des observations de 2005 à 2019 provenant de huit stations de température de l'eau. Les changements de température de l'eau ont été

évalués pour deux périodes futures (2040-2069 et 2070-2099) à l'aide de huit modèles climatiques inclus dans le projet de comparaison de modèles CMIP6 et de deux scénarios de voie socioéconomique partagée (SSP) (4,5 et 8,5 W/m^2 d'ici 2100) pour chaque période. Les résultats montrent que les températures de l'eau supérieures à 20°C (un seuil supérieur pour un habitat thermique adéquat pour la migration du saumon rouge (*Oncorhynchus nerka*) dans cette rivière) à la station Vanderhoof augmenteront en fréquence selon ces scénarios. Alors que la fréquence d'occurrence de ce phénomène est de 1% sur la base des observations estivales de 2005-2019, cette fourchette de chiffres est de 3.8 à 36% selon l'ensemble des scénarios de changement climatique. Ces résultats montrent la situation contraignante pour le saumon rouge et l'importance de la gestion de l'eau dans le cadre des changements climatiques dans ce contexte. Ces résultats fournissent également aux parties prenantes impliquées un aperçu de la manière dont les changements climatiques peuvent affecter les services écosystémiques et la gestion de l'eau pour conserver les habitats thermiques des poissons.

L'article 3 montre que le rayonnement infrarouge (à longues longueurs d'ondes; LI) est l'une des composantes du bilan énergétique en partie responsable des fluctuations thermiques de l'eau pendant les étés chauds dans le bassin versant de la rivière Nechako. Le LI entrant, émis par l'atmosphère, et le LI sortant émis par la surface terrestre (ou par l'eau) ne sont pas largement ni fréquemment mesurés. L'influence des nuages sur le bilan thermique du LI rend encore plus difficile l'établissement de formulations fiables pour les conditions du ciel en entier.

L'article 3 utilise les données de température de l'air, de température du point de rosée, de couverture nuageuse et de l'indice de surface foliaire (LAI) issues de la base de données ERA5 pour comparer neuf modèles d'irradiance infrarouge à ondes longues à la surface de la Terre (DLI) et les comparer avec le produit DLI d'ERA5. Notre travail repose sur des données continues de DLI recueillies sur une longue période à deux stations au Canada, ainsi que sur la réanalyse ERA5, une source fiable pour les régions dépourvues de données, comme le centre de la Colombie-Britannique et le nord du Québec (Canada). Les résultats démontrent la faisabilité de l'utilisation de différents modèles DLI avec les données de réanalyse ERA5, avec des RMSE variant de 30 à 38 W/m^2, ce qui est comparable à des études antérieures et peut facilement être appliqué à des échelles plus larges en l'intégrant dans des modèles hydrologiques. De plus, les données maillées d'ERA5 pour le DLI présentent les meilleurs résultats avec un RMSE de 30 à 32 W/m^2 aux deux stations. Cette performance accrue suggère l'utilisation directe des données ERA5 comme intrants pour les modèles thermiques.

Enfin, l'article 4 quantifie explicitement le DLI à l'aide de neuf formulations différentes qui ont été comparées dans l'article 3 et mises en œuvre dans le modèle CEQUEAU. De plus, cet article évalue l'utilisation directe du DLI ERA5 pour la modélisation thermique. Les résultats

indiquent que le modèle qui prend explicitement en compte les effets de la canopée grâce à la LAI, produit les meilleurs résultats. Ces résultats fiables sont utilisés pour la modélisation de la température de l'eau par le biais d'un calage multisite à l'échelle du bassin versant. Diverses formulations de l'équation de Stephan-Boltzmann se révèlent efficaces pour quantifier l'impact du DLI sur la température de l'eau à cette échelle. De plus, les données maillées d'ERA5 pour le DLI présentent des performances acceptables. La performance améliorée du meilleur modèle met en évidence le potentiel de partitionner le DLI en zones ouvertes et en zones avec canopée, en utilisant des formulations mathématiques et les données ERA5 comme données d'entrée pour les modèles thermiques. La signification de ces résultats est discutée, ce qui permet d'avancer dans notre compréhension de la modélisation de la température de l'eau et des études hydrologiques avec le modèle CEQUEAU.

Mots-clés : CEQUEAU, Modélisation, Température de l'eau de rivière, Calibrage, Impacts des changements climatiques, Rayonnement à ondes longues, Données ERA5, Modèles empiriques, Bassin versant de la rivière Nechako.

ABSTRACT

The temperature of the water is one of the critical variables for fish habitat assessments in rivers. Hydrological and thermal modeling is useful for providing water temperature estimates at the watershed scale. To this end, this thesis presents modeling developments in the Nechako River Basin (British Columbia, Canada) for improved water temperature modeling in the form of four interrelated research articles. River flows in the Nechako River are partially controlled by the Kenney Dam and the Skins Lake spillway. Rio Tinto, the owner of the Kenney Dam, employs the CEQUEAU model for flow forecasting. The Nechako River serves as a migration corridor to spawning grounds for some migratory fish. The presence of the dam, long-term impacts of climate change on water temperature, and the sensitivity of migratory fish to thermal changes make this region an excellent case study for the current research project.

Article 1 compares single-site and multisite calibration methods for the thermal module of the CEQUEAU model. Initially, a hydrological calibration is conducted, and the same calibrated hydrological model results are used for all thermal calibrations. Local calibration at a single site provides the best performance for individual sites, but these parameterizations perform poorly when applied to other sites. Multisite calibration and scaling methods demonstrate adequate performance for all sites, with a root mean square error ($RMSE_{global}$) of < 2°C as an acceptable threshold. However, the multisite method, where measurements from several sites are combined for a single calibration, provides the best performance measures and the shortest computation time. Article 2 presents the effects of climate change on the thermal regime of the Nechako River. The hydrothermal model CEQUEAU is calibrated based on flow and water temperature observations. Meteorological inputs for the model come from ERA5 reanalyses for the past and meteorological projections from climate models for future scenarios. Hydrological calibration is performed for the period 1980-2019 using data from two hydrometric stations, and water temperature calibration is conducted using observations from 2005 to 2019 from eight water temperature stations. Water temperature changes are assessed for two future periods (2040-2069 and 2070-2099) using eight climate models included in the CMIP6 model comparison project and two shared socioeconomic pathway (SSP) scenarios (4.5 and 8.5 W/m^2 by 2100) for each period. The results show that water temperatures above 20°C (an upper threshold for suitable thermal habitat for sockeye salmon (*Oncorhynchus nerka*) migration in this river at the Vanderhoof station will increase in frequency. While the occurrence frequency of this phenomenon is 1% based on summer observations from 2005-2019, this range varies from 3.8% to 36% depending on all climate change scenarios. These results illustrate the constraints on sockeye salmon and the importance of water management

in the context of climate change. They also provide stakeholders with insights into how climate change can affect ecosystem services and water management for preserving fish thermal habitats. Article 3 demonstrates that longwave radiation (LW) is one of the energy balance components partly responsible for temperature fluctuations in the water during hot summers in the Nechako River Basin. Incoming LW radiation, emitted by the atmosphere, and outgoing LW radiation emitted by the land (or water) are not widely or frequently measured. The influence of clouds on the LW thermal balance further complicates the establishment of reliable formulations for all sky conditions. Article 3 uses air temperature, dew point temperature, cloud cover, and leaf area index (LAI) data from the ERA5 database to compare nine models of downward longwave irradiance (DLI) at the Earth's surface and compares them with ERA5's DLI product. Our work relies on continuous DLI data collected over an extended period at two locations in Canada, in addition to the ERA5 reanalysis, a reliable data source for data-scarce areas, such as central British Columbia and northern Quebec (Canada). The results demonstrate the feasibility of using various DLI models with ERA5 reanalysis data for all sky conditions, with RMSE performance metrics ranging from 30 to 38 W/m^2. This performance is comparable to previous studies and can be easily applied on a broader scale by integrating DLI models into hydrological models. Furthermore, it is shown that ERA5 gridded data for DLI yields the best results with RMSE values of 30 to 32 W/m^2 at both locations. This improved performance suggests the direct use of ERA5 data as input for hydrological and ecological models. Lastly, Article 4 explicitly quantifies DLI using nine different formulations compared in Article 3 and implemented in the CEQUEAU model. Furthermore, this article evaluates the direct use of DLI from the ERA5 product for hydrological and thermal studies. The results indicate that the only model that explicitly considers canopy effects through LAI produces the best outcomes. These reliable results are utilized for water temperature modeling through multisite calibration at the watershed scale. Various formulations of the Stephan-Boltzmann equation prove effective for quantifying the impact of DLI on water temperature at this scale. Additionally, ERA5 gridded data for DLI demonstrates acceptable performance. The results show the potential for partitioning DLI into open areas and canopies, using mathematical formulations and ERA5 data as input for thermal models. The significance of these findings is discussed, advancing our understanding of water temperature modeling and hydrological studies with the CEQUEAU model.

Keywords: CEQUEAU, Modelling, River water temperature, Calibration, Climate change Impacts, Longwave radiation, ERA5 data, Empirical models, Nechako River watershed.

TABLE DES MATIÈRES

REMERCIEMENTS ..... **II**
RÉSUMÉ ..... **V**
ABSTRACT ..... **VIII**
TABLE DES MATIÈRES ..... **X**
LISTE DES FIGURES ..... **XV**
LISTE DES TABLEAUX ..... **XIX**
LISTE DES ABRÉVIATIONS ..... **XX**
1 INTRODUCTION **22**
1.1 CONTEXTE 22
1.2 OBJECTIFS MÉTHODOLOGIQUES 23
1.3 STRUCTURE DE LA THÈSE 24
2 REVUE DE LITTERATURE **26**
2.1 MODELES HYDROLOGIQUES ET THERMIQUES 26
2.2 SCENARIOS FUTURS ASSOCIES AUX CHANGEMENTS CLIMATIQUES: IMPLICATIONS POUR LA GESTION DU DEBIT ET DE LA TEMPERATURE D'ETE EN AVAL DES BARRAGES 30
2.3 MODELISATION DE L'IRRADIANCE DESCENDANTE A ONDES LONGUES A LA SURFACE DE L'EAU 31
3 SYNTHESE DES TRAVAUX DE RECHERCHE **34**
3.1 SITES D'ETUDE 34
3.2 MODELE CEQUEAU 35
3.3 DONNÉES DE RÉANALYSE ERA5 38
3.4 ALGORITHME DE CALAGE: CMA-ES 38
3.4.1 Calage du modèle hydrologique *39*
3.4.2 Calage du modèle thermique *39*
3.4.3 Calage des coefficients dans les équations DLI *40*
3.5 CALAGE MULTISITE D'UN MODELE HYDROLOGIQUE ET THERMIQUE SEMI-DISTRIBUE DANS UN GRAND BASSIN VERSANT CANADIEN 41
3.6 SCENARIOS FUTURS DE DEBIT ET DE TEMPERATURE DE L'EAU DANS UN BASSIN VERSANT AMENAGE : IMPLICATIONS POUR LA GESTION DU DEBIT ET DE LA TEMPERATURE D'ETE EN AVAL DU BARRAGE 42
3.7 ÉVALUATION DES MODELES D'IRRADIANCE A ONDES LONGUES DESCENDANTES AVEC DES DONNEES D'ENTREE ERA5 DANS LES ENVIRONNEMENTS CANADIENS 45
3.8 MODELISATION DE L'IMPACT DE L'IRRADIANCE DESCENDANTE A ONDES LONGUES SUR LA TEMPERATURE DE L'EAU A L'ECHELLE DU BASSIN VERSANT A L'AIDE DES DONNEES D'ENTREE ERA5 DANS L'OUEST DU CANADA 49

4 CONCLUSION ET PERSPECTIVES 51
PERSPECTIVES FUTURES 54
5 RÉFÉRENCES BIBLIOGRAPHIQUES 55
6 ARTICLE I 61
ABSTRACT 63
KEY WORDS 63
6.1 INTRODUCTION 63
6.2 METHODS 66
6.2.1 Study area 66
6.2.2 CEQUEAU model 69
6.2.3 CMA-ES calibration algorithm 72
6.2.4 Single-site and multisite calibration methods 73
6.2.5 Upscaling operators to derive parameters at spatial scales 74
6.2.6 Model calibration setup and implementation 75
6.3 RESULTS 80
6.3.1 Hydrological module parameters and performance metrics 80
6.3.2 Thermal module parameters and performance metrics 81
6.3.3 Energy flux contribution to water temperature and the data availability impact on calibration 83
6.4 DISCUSSION 84
6.4.1 The logic of calibration strategies for global parameters 85
6.4.2 Hydrological module calibration impacts on thermal calibration 85
6.4.3 The calibrated parameter values and their impacts 86
6.4.4 Final selection for the calibration method 88
6.5 CONCLUSION 88
ACKNOWLEDGEMENT 89
ARTICLE I BIBLIOGRAPHIE 90
7 ARTICLE II 94
ABSTRACT 96
KEY WORDS 96
7.1 INTRODUCTION 96
7.2 METHODS 99
7.2.1 Study area 99
7.2.2 CEQUEAU model 100
7.2.3 Modeling upstream boundary conditions 102
7.2.4 Fifth generation of ECMWF atmospheric reanalysis data 104
7.2.5 Model calibration and implementation 105
7.2.6 Future model projections 107
7.2.7 Duration curve method 108

7.3 RESULTS 108
7.3.1 Hydrological module calibration and parameters 108
7.3.2 Thermal model calibration 110
7.3.3 Future projections 113
7.4 DISCUSSION 119
7.4.1 Hydrothermal modeling 119
7.4.2 Impacts of future upstream boundary conditions on downstream water temperatures 120
7.4.3 Future water temperature simulation under climate change impacts 120
7.4.4 Ecological implications of Nechako River water temperature warming 121
7.5 CONCLUSION 121
ACKNOWLEDGEMENT 122
ARTICLE II BIBLIOGRAPHIE 123

8 ARTICLE III 128
ABSTRACT 130
KEY WORDS 130
8.1 INTRODUCTION 130
8.2 METHODS 132
8.2.1 Study areas 132
8.2.2 ERA5 reanalysis data 134
8.2.3 Longwave heat budget modeling 134
8.2.4 ECMWF scheme for DLI data 141
8.2.5 Calibration algorithm 141
8.2.6 Calibration objective function and performance metrics 142
8.2.7 Calibration process 143
8.2.8 Experimental uncertainty analysis 145
8.3 RESULTS 146
8.3.1 Model calibration 146
8.3.2 Model validations 151
8.3.3 Data measurement error and propagation using the ERA5 dataset 154
8.4 DISCUSSION: UNVEILING THE COMPLEXITIES OF DLI MODELING 157
8.4.1 Physical underpinnings of DLI formulations 157
8.5 CONCLUSION 165
ACKNOWLEDGMENT 166
ARTICLE III BIBLIOGRAPHIE 167

9 ARTICLE IV 171
ABSTRACT 173
KEY WORDS 173
9.1 INTRODUCTION 173

9.2 METHODS 175
9.2.1 Study area 175
9.2.2 ERA5 reanalysis dataset 177
9.2.3 CEQUEAU model 178
9.2.4 Model calibration 182
9.3 RESULTS 183
9.3.1 Model calibration 183
9.3.2 Multi-model ensemble simulations 186
9.3.3 Uncertainty quantification 188
9.4 DISCUSSION 193
9.4.1 Model parameter uncertainty 193
9.4.2 CEQUEAU model inadequacies affecting H_{long} 194
9.4.3 DLI model choice impact on model calibration parameters 195
9.4.4 The acceptable performance of models 2-8 196
9.4.5 ERA5 performance at the watershed scale 196
9.4.6 Model 9 superiority 197
9.5 CONCLUSION 198
ACKNOWLEDGMENT 199
ARTICLE IV BIBLIOGRAPHIE 200

APPENDIX I 204
APPENDIX I-A: UNDERSTANDING PARAMETERS IN THE CEQUEAU HYDROLOGICAL MODULE 204
APPENDIX I-B: THE FOUR STEPS OF THE CMA-ES ALGORITHM 205
APPENDIX I-C: SIMULATED TIME SERIES WITH SINGLE SITE CALIBRATION PARAMETERS FOR EACH STATION 207
APPENDIX I-D: SIMULATED TIME SERIES WITH MULTISITE CALIBRATION PARAMETERS FOR EACH STATION 211
APPENDIX I-E: THE MATLAB CODES IMPLEMENTED IN ARTICLE I 215

APPENDIX II 218
APPENDIX II-A: SUPPLEMENTARY INFORMATION, INCLUDING DATA FOR SPECIES CHARACTERISTICS, CLIMATE, HYDROLOGY, AND CALIBRATION IN THE STUDY 218
APPENDIX II-B: CLIMATE SIMULATION FIGURES AND DURATION CURVE ANALYSIS FOR VANDERHOOF STATION 223
APPENDIX II-C: WATER TEMPERATURE VARIABILITY FIGURES UNDER DIFFERENT CLIMATE SCENARIOS AND TIME HORIZONS 225
APPENDIX II-D: THE MATLAB CODES IMPLEMENTED IN ARTICLE II 230

APPENDIX III 243
APPENDIX III-A: SENSOR LOCATIONS IN THE NECHAKO AND BVESM WATERSHEDS 243
APPENDIX III-B: PARAMETER CALIBRATION RESULTS FOR DLI MODELS AT THE SLS STATION 244
APPENDIX III-C: THE MATLAB CODES IMPLEMENTED IN ARTICLE III 248

APPENDIX IV 249

APPENDIX IV-A: COOLING AND WARMING IMPACTS OF DLI FORMULATION ON WATER'S LONGWAVE HEAT BUDGET (H_{long}) 249

APPENDIX IV-B: PROCESS-BASED RIVER TEMPERATURE MODELS AND REFERENCES 250

APPENDIX IV-C: MODULAR CODE DIAGRAM OF CEQUEAU MODEL..... 251

APPENDIX IV-D: ALL PARAMETERS USED IN THE CEQUEAU MODEL CALIBRATIONS 252

APPENDIX IV-E: SIMULATED WATER TEMPERATURE TIME SERIES FOR DLI MODELS AT ALL STATIONS USING MULTISITE CALIBRATION PARAMETERS 254

APPENDIX IV-F: THE HEAT BUDGET CONTRIBUTIONS FOR VANDERHOOF STATION (CP = 23)..... 264

APPENDIX IV-G: THE C++ CODES IMPLEMENTED IN ARTICLE IV 265

LISTE DES FIGURES

FIGURE 2.1: TROIS SCHÉMAS DIFFÉRENTS SONT UTILISÉS POUR CRÉER DES UNITÉS DE RÉPONSE HYDROLOGIQUE (HRU) SUR UN BASSIN VERSANT [ADAPTE DE CRAIG *ET AL.* (2020)] 27

FIGURE 2.2: REPRÉSENTATION DES TERMES DU BILAN THERMIQUE POUR UNE SECTION DE RIVIÈRE SITUÉE À L'INTÉRIEUR D'UNE UNITÉ DE RÉPONSE HYDROLOGIQUE (HRU) 29

FIGURE 3.1 BASSIN VERSANT DE LA RIVIÈRE NECHAKO ET ZONE D'ÉTUDE 34

FIGURE 3.2 REPRÉSENTATION SCHÉMATIQUE DE (A) UN BASSIN VERSANT, DE CARRÉS ENTIERS ET DE CARRÉS PARTIELS, (B) CHAQUE CARRÉ PARTIEL (CP) EST UNE UNITÉ DE RÉPONSE HYDROLOGIQUE POUR LAQUELLE LES BUDGETS HYDROLOGIQUES ET THERMIQUES SONT CALCULÉS. LE CŒUR DU MODULE HYDROLOGIQUE EST CONSTITUÉ DE LA FONCTION DE PRODUCTION (PF) ET DE LA FONCTION DE TRANSFERT (TF), QUI CALCULE L'EAU DISPONIBLE À L'INTÉRIEUR DE CHAQUE CP ET PROCÈDE AU ROUTAGE SUBSÉQUENT VERS L'AVAL. (C) LE MODULE THERMIQUE CALCULE LE BUDGET THERMIQUE SUR CHAQUE CP. L'ÉNERGIE THERMIQUE EST ÉGALEMENT TRANSFÉRÉE VERS L'AVAL; (D) MONTRE LA PF SCHÉMATIQUEMENT, QUI SE COMPOSE DE TROIS RÉSERVOIRS INTERCONNECTÉS POUR MODÉLISER LES PROCESSUS HYDROLOGIQUES À L'INTÉRIEUR DE CHAQUE CP. 37

FIGURE 6.1 MAP OF THE NECHAKO WATERSHED 67

FIGURE 6.2 SCHEMATIC REPRESENTATION OF (A) WATERSHED, WHOLE SQUARES AND PARTIAL SQUARED, (B) EACH PARTIAL SQUARE AS A HYDROLOGICAL RESPONSE UNIT OR HRU, WHICH SHOWS PRODUCTION FUNCTION VARIABLES, WATER RESERVOIRS, AND CEQUEAU HYDROLOGICAL PROCESSES FOR EACH PARTIAL SQUARE (ST-HILAIRE ET AL. 2015). THESE VARIABLES ARE DEFINED IN TABLE I-A1 IN APPENDIX I-A 70

FIGURE 6.3 ESTIMATION OF PARAMETER ρ_k^1 AT LEVEL 1 FOR k^{TH} PARAMETER OF THE MODEL USING UPSCALING OPERATOR $\mathcal{U}_k\langle\cdot\rangle$ ON $\rho_{k,i}^0$ AT LEVEL 0 ON ALL SUB WATERSHEDS i $(i = 1, \dots, n)$ 75

FIGURE 6.4 IDENTIFICATION OF LOW AND HIGH FLOW PERIODS FOR THREE HYDROMETRIC STATIONS 77

FIGURE 6.5 (A) HYDROLOGICAL AND THERMAL CALIBRATION PROCESS DIAGRAM; (B) SINGLE-SITE CALIBRATION PROCESS (METHOD 1); (C) MULTISITE CALIBRATION METHOD (METHOD 2) AND (D) SINGLE-SITE AVERAGING METHOD (METHOD 3) 79

FIGURE 6.6 SIMULATED FLOW USING HYDROLOGICAL CALIBRATED PARAMETERS FOR THE VANDERHOOF STATION. THIS SET OF PARAMETERS AND SIMULATION FOR THE WHOLE BASIN WAS USED FOR THERMAL CALIBRATION AND SIMULATIONS 80

FIGURE 6.7 HEAT BUDGET COMPONENT FOR VANDERHOOF STATION (STATION 2). HEAT TERMS ARE: QRADSO = SOLAR SHORTWAVE RADIATION, QRADIN = LONGWAVE RADIATION, QEVAP = EVAPOTRANSPIRATION, QCONV = SENSIBLE HEAT, QRUISS = LOCAL ADVECTION 84

FIGURE 7.1 THE NECHAKO RIVER WATERSHED IS LOCATED IN THE CENTRAL PART OF BRITISH COLUMBIA, CANADA, WITH A 45,000 KM^2 DRAINAGE AREA. DOWNSTREAM OF OOTSA LAKE (THE LARGE LAKE IMMEDIATELY UPSTREAM OF SLS), THE FLOW OF THE RIVER IS FULLY REGULATED BY THE SLS BETWEEN THE NECHAKO RESERVOIR AND ITS CONFLUENCE WITH THE NAUTLEY RIVER, AFTER WHICH THE FLOWS OF THE NECHAKO AND NAUTLEY RIVERS COMBINE AND CONTINUE PAST THE TOWN OF VANDERHOOF AND EVENTUALLY INTO THE FRASER RIVER. THREE HYDROMETRIC STATIONS MEASURE THE DISCHARGE (ENVIRONMENT AND CLIMATE CHANGE CANADA, HTTPS://WWW.CANADA.CA/EN/ENVIRONMENT-CLIMATE-CHANGE/, DATA ACCESS: MAY 2022). THOSE ARE AT THE SLS, THE NAUTLEY RIVER (UPSTREAM OF THE CONFLUENCE WITH THE NECHAKO RIVER), AND THE TOWN OF VANDERHOOF. THE SLS FLOWS AND TEMPERATURES PROVIDE THE BOUNDARY CONDITIONS TO THE HYDROLOGICAL AND THERMAL MODULES AND ARE NOT USED AS TARGET STATIONS FOR HYDROLOGICAL CALIBRATION. EIGHT STATIONS MEASURE WATER TEMPERATURE (RIO TINTO, DATA ACCESS: MAY 2022) IN THE WATERSHED. THE STUDY AREA AND MONITORING STATIONS ARE EXPLAINED IN DETAIL IN OUELLET-PROULX ET AL. (2017A) AND KHORSANDI ET AL. (2022). 100

FIGURE 7.2 SCHEMATIC REPRESENTATION OF STEPS, CONCEPTS, AND MODELS TO SET UP THE CEQUEAU MODEL IN THIS STUDY, INCLUDING (A) PHYSIOGRAPHIC DATA FOR THE WATERSHED, WHOLE SQUARES, AND PARTIAL SQUARES; (B) INPUT METEOROLOGICAL DATA (PRECIPITATION, MIN AND MAX AIR TEMPERATURE, VAPOR PRESSURE, CLOUD COVER, NET SOLAR SHORTWAVE RADIATION), (C) THE MODELS INCORPORATED TO PREPARE UPSTREAM BOUNDARY CONDITION FOR CEQUEAU; (D) STRUCTURE OF A SAMPLE PARTIAL SQUARE (CP) WHICH IS A HYDROLOGICAL RESPONSE UNIT FOR WHICH BOTH HYDROLOGICAL AND THERMAL BUDGETS ARE COMPUTED. THE CORE OF THE HYDROLOGICAL MODULE IS THE PRODUCTION FUNCTION (PF) AND TRANSFER FUNCTION (TF), WHICH CALCULATES AVAILABLE WATER INSIDE EACH CP AND SUBSEQUENT ROUTING DOWNSTREAM. THE THERMAL MODULE CALCULATES THE HEAT BUDGET USING AVAILABLE WATER INSIDE EACH CP. 102

FIGURE 7.3 SIMULATED AND OBSERVED FLOWS USING THE CALIBRATED MODEL FOR VANDERHOOF AND NAUTLEY STATIONS FOR THE JUNE-SEPTEMBER PERIOD 109

FIGURE 7.4 THE SIMULATED (BLUE) AND OBSERVED (RED) WATER TEMPERATURE TIME SERIES FOR EACH STATION USING THE MULTISITE CALIBRATION APPROACH FOR THE JUNE-SEPTEMBER PERIOD 110

FIGURE 7.5 (A) THE YEARLY RANGE [FOR JULY-AUGUST ONLY] OF SIMULATED WATER TEMPERATURE VARIABILITY IS SHOWN USING BOXPLOTS FOR THE VANDERHOOF STATION'S BASELINE PERIOD 1980-2019. THE RED LINE REPRESENTS THE MEDIAN VALUE, AND THE BLUE CIRCLES SHOW THE MEAN SIMULATED VALUE. THE UPPER AND LOWER LIMITS OF THE BOXES SHOW THE 25TH AND 75TH PERCENTILES, RESPECTIVELY, AND THE WHISKERS SHOW THE MOST EXTREME DATA. THE MEAN ANNUAL TEMPERATURE OF THE OBSERVED DATA IS SHOWN WITH THE RED CROSSES (X) (2005-2019 ONLY). (B) RED CROSSES AND BLUE CIRCLES SHOW THE ANNUAL NUMBER OF DAYS ON WHICH THE MEAN DAILY TEMPERATURE EXCEEDS THE CRITICAL LIMIT OF 20°C FOR SIMULATED AND OBSERVED TIME SERIES, RESPECTIVELY. THE HORIZONTAL DOTTED LINE SHOWS THE MAXIMUM NUMBER OF OBSERVED DAYS ABOVE 20°C IN A YEAR, EQUAL TO 9 DAYS 112

FIGURE 7.6 (A) DURATION CURVE OF BASELINE (1980-2019) SIMULATED DATA AND OBSERVED WATER TEMPERATURES AT THE VANDERHOOF STATION (B) WATER TEMPERATURE TIMING FOR THE WARM SEASON (JUNE-SEPTEMBER) FOR OBSERVED AND SIMULATED TIME SERIES TOGETHER WITH THE LOWER AND UPPER BOUNDARY OF SIMULATED VALUES 113

FIGURE 7.7 DURATION CURVE OF SIMULATED DATA USING EIGHT CLIMATE MODELS FOR SSPS 4.5 AND 8.5 W/M^2 AND FOR THE 2040-2069 AND 2070-2099 TIME HORIZONS AT THE VANDERHOOF STATION. THE MODELS WITH THE LOWEST AND HIGHEST EXCEEDANCE PROBABILITIES WERE USED TO CALCULATE THE MINIMUM AND MAXIMUM EXCEEDANCE PROBABILITIES DEPICTED ON THE GRAPHS. 115

FIGURE 7.8 HIGH WATER TEMPERATURES TIMING FOR THE JUNE-SEPTEMBER PERIOD FOR THE SIMULATED TIME SERIES IN THE FUTURE, TOGETHER WITH THE LOWER AND UPPER BOUNDARY OF SIMULATED VALUES 116

FIGURE 7.9 THE YEARLY RANGE OF WATER TEMPERATURE VARIABILITY IS SHOWN USING BOXPLOTS FOR TWO TIME HORIZONS AND TWO SSPS. THE RED LINE REPRESENTS THE MEDIAN MULTI-MODEL VALUE, AND THE BLACK CIRCLES SHOW THE MEAN MULTI-MODEL SIMULATED VALUE. THE UPPER AND LOWER LIMITS OF THE BOXES SHOW 25TH AND 75TH PERCENTILES, RESPECTIVELY, AND THE WHISKERS SHOW THE MOST EXTREME DATA. THE HORIZONTAL DOTTED LINE ON THE LEFT PANELS SHOWS THE MAXIMUM OBSERVED MEAN OF TEMPERATURES FOR JULY-AUGUST MONTHS IN THE PAST. BLUE ASTERISKS ON THE RIGHT-SIDE PLOTS SHOW THE ANNUAL NUMBER OF DAYS ON WHICH THE TEMPERATURE EXCEEDS THE CRITICAL LIMIT OF 20°C FOR SIMULATED AND OBSERVED TIME SERIES USING MULTI-MODEL MEDIAN VALUES. THE HORIZONTAL DOTTED LINE SHOWS THE MAXIMUM NUMBER OF OBSERVED DAYS ABOVE 20°C IN A YEAR, WHICH OCCURRED DURING THE 2005-2019 PERIOD AND IS EQUAL TO 9 DAYS. 118

FIGURE 8.1 NECHAKO RIVER AND BASSIN VERSANT EXPÉRIMENTAL DE SAINTE-MARTHE WATERSHEDS AND MEASURING STATIONS AT SKINS LAKE SPILLWAY AND BASSIN VERSANT EXPÉRIMENTAL DE SAINTE-MARTHE (BVESM) 133

FIGURE 8.2 SCATTER PLOTS AND RMSE FOR ERA5 DLI AND NINE DLI MODELS IN SLS STATION CALIBRATION 148

FIGURE 8.3 SCATTER PLOTS AND RMSE FOR ERA5 DLI AND NINE DLI MODELS IN BVESM STATION CALIBRATION 149

FIGURE 8.4 SCATTER PLOTS AND RMSE FOR ERA5 DLI AND NINE DLI MODELS IN SLS STATION VALIDATION 152

FIGURE 8.5 SCATTER PLOTS AND RMSE FOR ERA5 DLI AND NINE DLI MODELS IN BVESM STATION VALIDATION 153

FIGURE 8.6 SCATTER PLOTS COMPARING ERA5'S 2 M AIR TEMPERATURES WITH HOURLY OBSERVED MEASUREMENTS AT (A) SLS AND (B) BVESM STATIONS, ALONG WITH THEIR RESPECTIVE RMSE PERFORMANCE METRICS. 154

FIGURE 8.7 SCATTER PLOTS OF $\beta_{Effective}$ VS. CLIMATIC VARIABLES (T, e_a, AND B) FROM ERA5 DATASET FOR (A, B, AND C) CLEAR-SKY CONDITIONS AND (D, E, AND F) ALL SKY CONDITIONS FOR THE SLS STATION 158

FIGURE 8.8 SCATTER PLOTS OF $\beta_{Effective}$ VS. CLIMATIC VARIABLES (T, e_a, AND B) FROM ERA5 DATASET FOR (A, B, AND C) CLEAR-SKY CONDITIONS AND (D, E, AND F) ALL SKY CONDITIONS FOR THE BVESM STATION 158

FIGURE 8.9 SCATTER PLOTS BETWEEN THE MEASURED EFFECTIVE EMISSIVITY ($\beta_{Effective}$) AND (e_a/T_{air}) FROM THE ERA5 DATASET FOR (A) CLEAR-SKY CONDITIONS AND (B) ALL SKY CONDITIONS AT THE SLS STATION. 161

FIGURE 8.10 SCATTER PLOTS BETWEEN THE MEASURED EFFECTIVE EMISSIVITY ($\beta_{Effective}$) AND (e_a/T_{air}) FROM THE ERA5 DATASET FOR (A) CLEAR-SKY CONDITIONS AND (B) ALL SKY CONDITIONS AT THE BVESM STATION. 161

FIGURE 8.11 SCATTER PLOTS BETWEEN THE MEASURED EFFECTIVE EMISSIVITY ($\beta_{Effective}$) AND LEAF AREA INDEX (LAI) FROM THE ERA5 DATASET FOR (A) CLEAR-SKY CONDITION AND (B) ALL SKY CONDITIONS FOR THE SLS STATION 163

FIGURE 8.12 SCATTER PLOTS BETWEEN THE MEASURED EFFECTIVE EMISSIVITY ($\beta_{Effective}$) AND LEAF AREA INDEX (LAI) FROM THE ERA5 DATASET FOR (A) CLEAR-SKY CONDITION AND (B) ALL SKY CONDITIONS FOR THE BVESM STATION 163

FIGURE 8.13 (A) RELATIONSHIP BETWEEN THE SKY VIEW VARIABLE (V_f) AND THE LEAF AREA INDEX (LAI) USING EQUATION (8.17) FOR THREE α VALUES. (B) COMPUTED CLOUD CORRECTION FACTORS DETERMINED FROM EQUATION (8.13) AND CALIBRATED u AND v VALUES IN THIS STUDY FOR BOTH SLS (FIRST 8 ROWS NUMBERED FROM MODEL 2 TO 9) AND BVESM (SECOND 8 ROWS NUMBERED FROM MODEL 2 TO 9) STATIONS (TABLE III-B1 IN APPENDIX III-B). 164

FIGURE 9.1 NECHAKO RIVER WATERSHED STUDY SITE, PRESENTING IMPORTANT LOCATIONS TOGETHER WITH LABELED MONITORING STATIONS. THE MODEL AND ERA5 GRIDS ARE ALSO SHOWN. AMONG ALL 615 PARTIAL SQUARES (CPS), ONLY THOSE RELATED TO THE MAIN RIVER ARE COLORED. 177

FIGURE 9.2 PERFORMANCE METRIC (RMSE) RESULTED FROM CALIBRATIONS FOR (A) EACH SITE USING DERIVED MULTISITE CALIBRATION (NUMBERED 1, 2, 3, AND 4 FOR ACCORDING STATIONS), AND (B) $RMSE_{MULTISITE}$ USING ALL STATIONS (1-4). 185

FIGURE 9.3 SIMULATED WATER TEMPERATURE TIME SERIES FOR THE FOUR STATIONS USING CALIBRATED CEQUEAU EQUIPPED WITH TEN DLI MODELING ALTERNATIVES. THE SHADED AREA SHOWS LOWER AND UPPER ESTIMATES FOR EACH TIME STEP, AND THE BLACK LINE IS THE OBSERVED VALUE. 187

FIGURE 9.4 SCATTER PLOTS FOR THE COMPARISON OF ERA5'S METEOROLOGICAL DATA AT 2 M HEIGHT WITH OBSERVED MEASUREMENTS AT THE SLS STATION ON A DAILY TIME BASIS (THE AIR TEMPERATURE DATA ARE AVERAGED, AND RADIATIONS ARE SUMMED USING HOURLY DATA) 190

FIGURE 9.5 OBSERVED AND SIMULATED FLOW FOR THE VANDERHOOF STATION USING CALIBRATED HYDROLOGICAL PARAMETERS. THE SIMULATION CURVE REPRESENTS THE MEAN OF EIGHT SETS OF PARAMETERS RESULTING FROM ITERATIVE CALIBRATIONS CONDUCTED FOR THE ENTIRE BASIN. THE SHADED AREA DENOTES THE UPPER AND LOWER UNCERTAINTY BOUNDS. 191

FIGURE 9.6 SIMULATED H_{long} TIME SERIES FOR THE VANDERHOOF STATION USING CALIBRATED CEQUEAU EQUIPPED WITH TEN DLI MODEL ALTERNATIVES. THE SHADED AREA SHOWS LOWER AND UPPER BOUNDS FOR EACH TIME STEP, AND THE BLUE LINE IS THE ENSEMBLE MEAN OF SIMULATED VALUES. 195

LISTE DES TABLEAUX

TABLEAU 2.1 FORME GÉNÉRALE DES TERMES DE BILAN THERMIQUE POUR UN PLAN D'EAU 28

TABLEAU 3.1 FORMULATION DES NEUF MODÈLES D'ESTIMATION DLI POUR CETTE ÉTUDE 47

TABLE 6.1 WATER FLOW STATION LOCATIONS AND DATA AVAILABILITY PERIOD FOR EACH STATION 68

TABLE 6.2 PARAMETERS USED IN THE CEQUEAU THERMAL MODULE 72

TABLE 6.3 SEVERAL METHODS FOR UPSCALING PROPOSED BY SAMANIEGO *ET AL.* (2010) TO DERIVE PARAMETERS AT LEVEL 1 USING THE PARAMETERS FROM LEVEL 0 74

TABLE 6.4 WATER TEMPERATURE MEASUREMENT LOCATIONS AND DATA AVAILABILITY PERIOD FOR EACH STATION 76

TABLE 6.5 DIFFERENT SCENARIOS DEFINED FOR CALIBRATION 78

TABLE 6.6 THERMAL PARAMETERS VALUES FOR DIFFERENT CALIBRATION SCENARIOS 81

TABLE 6.7 PERFORMANCE METRICS BY CALIBRATING/VALIDATING THE THERMAL MODULE PARAMETERS OF CEQUEAU USING TEMPERATURE MEASUREMENTS AT DIFFERENT LOCATIONS FOR: [1] FOCUSED SITE FOR CALIBRATION (METHOD 1) AND [2] THE OTHER TWO SITES WHICH ARE NOT THE FOCUS OF CALIBRATION 82

TABLE 7.1 THE ONSET AND END DATES FOR THE 17°C THERMAL THRESHOLD (FIGURE 7.6B) BASED ON THE STMP PERIOD (JULY 20 TO AUGUST 20) 116

TABLE 8.1 NAMES AND LOCATIONS OF LONGWAVE RADIATION MEASURING STATIONS 133

TABLE 8.2 FORMULATION OF THE NINE DLI ESTIMATION MODELS FOR THIS STUDY 144

TABLE 8.3 CALIBRATION PERFORMANCE METRICS FOR ERA5 DATA AND NINE DLI MODELS COMPARED WITH OBSERVED DATA AT SLS AND BVESM STATIONS (10 CALIBRATION ITERATIONS). BEST CANDIDATES ARE IN BOLD, AND ALTERNATIVES ARE UNDERLINED AND NUMBERED. 147

TABLE 8.4 VALIDATION PERFORMANCE METRICS FOR ERA5 DATA AND NINE DLI MODELS COMPARED WITH OBSERVED DATA AT SLS AND BVESM STATIONS. BEST CANDIDATES ARE IN BOLD, AND ALTERNATIVES ARE UNDERLINED AND NUMBERED. 151

TABLE 9.1 FORMULATION OF THE NINE DLI ESTIMATION MODELS FOR THIS STUDY 182

TABLE 9.2 PERFORMANCE METRICS BY CALIBRATING THE THERMAL MODULE PARAMETERS OF CEQUEAU USING TEMPERATURE MEASUREMENTS VIA MULTISITE CALIBRATION USING STATIONS 1-4. THE BEST DLI FORMULATION METHOD WITH THE BEST RMSE IS BOLD. THE MODEL MARKED BY A STAR IS THE CEQUEAU'S DEFAULT MODEL. 185

TABLE 9.3 THERMAL PARAMETERS VALUES FOR DIFFERENT MULTISITE CALIBRATION SCENARIOS 186

TABLE 9.4 MINIMUM AND MAXIMUM DIURNAL VARIATION ($|Observed_{max} - Observed_{min}|$) FOR TEMPERATURE MEASUREMENTS AT SIX STATIONS. 188

TABLE 9.5 THE MEAN CONTRIBUTION OF EACH SURFACE HEAT FLUX TO THE TOTAL LOCAL HEAT BUDGET VARIATIONS DURING THE DAY USING THE SIMULATED TIME SERIES FOR THE 2016-2019 PERIOD (THE MEANS ARE CALCULATED USING APRIL- NOVEMBER SIMULATED VALUES). FOR THIS SIMULATION, THE ERA5 DLI (MODEL 10) IS USED. 192

LISTE DES ABRÉVIATIONS

AET	Actual EvapoTranspiration
AI	Artificial Intelligence
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
BVESM	Bassin Versant Expérimental de Sainte-Marthe
CE	Carreaux Entiers
CMA-ES	Covariance Matrix Adaptation Evolution Strategy
CMIP6	Coupled Model Intercomparison Project version 6
CP	Carreaux Partiels
DEM	Digital Elevation Model
DLI	Downward Longwave Irradiance
DOY	Day of the Year
ECMWF	European Center for Medium-Range Weather Forecasting
ERA5	European ReAnalysis 5th generation
ESA	European Space Agency
ESRI	Environmental Systems Research Institute
ÉTS	École de Technologie Supérieure
GCM	General Circulation Model
GIS	Geographic Information System
HRU	Hydrologic Response Unit
IFS	Integrated Forecasting System
KGE	Kling–Gupta Efficiency
LH-OAT	Latin Hypercube One factor At a Time
LR	Longwave Radiation
MASL	Meters Above Sea Level
MEX	Matlab EXecutable
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NSE	Nash–Sutcliffe Efficiency
NSERC	Canadian Natural Sciences and Engineering Research Council
PET	Potential EvapoTranspiration
PF	Production Function
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
SLS	Skins Lake Spillway
SMSC	Simultaneous Multi Site Calibration
SRTM	Shuttle Radar Topography Mission

SSC	Single Site Calibration
SSP	Shared Socioeconomic Pathway
STMP	Summer Temperature Management Program
SWAT	Soil and Water Assessment Tool
TDC	Temperature Duration Curve
TF	Transfer Function
TN	Total Nitrogen
TP	Total Phosphorus
TSS	Total Suspended Sediment
VIC	Variable Infiltration Capacity

1 INTRODUCTION

Le début de l'ère anthropocène a entraîné de nombreux stress sur le biote aquatique, y compris l'extinction de certaines espèces (Grooten & Almond, 2018). L'un des impacts les plus importants est lié au changement climatique. De petits changements dans la température globale ont des impacts significatifs sur différents organismes vivants. De plus, la plupart des écosystèmes touchés sont aquatiques (Grooten & Almond, 2018).

Les poissons migrateurs anadromes comme les saumons sont touchés par les changements dans les habitats d'eau salée et d'eau douce. Depuis le début de l'Anthropocène [1964 et après, basé sur Lewis and Maslin (2015)], des études ont montré une réduction de 76% de l'abondance des populations de poissons de rivière migrateurs (Grooten & Almond, 2018). Les ectothermes, y compris les espèces migratrices ou non migratrices, sont sensibles aux changements de régime thermique de l'eau. L'évolution du climat risque d'avoir un impact sur les habitats thermiques pour les principales espèces de poissons d'eau froide. Utilisant une étude de cas sur une rivière avec barrage dans l'ouest du Canada, ce projet vise à adapter des outils de modélisation afin de permettre aux gestionnaires de simuler adéquatement des scénarios futurs de débit et température de l'eau.

1.1 Contexte

Le bassin hydrographique de la rivière Nechako (Figure 3.1) est une vaste zone du centre de la Colombie-Britannique et ses rivières abritent des espèces de poissons migrateurs, en particulier des saumons. Dans ce bassin versant les rivières Nechako et Nautley se rejoignent à Fort Fraser, puis à environ 80 km en aval, les rivières Nechako et Stewart se rejoignent près de Finmoore. Le barrage Kenney a altéré le débit naturel de la rivière Nechako et il est fortement régulé dans le sous-bassin versant de la Nechako par le déversoir de Skins Lake. Ce barrage appartient au partenaire industriel de cette recherche, Rio Tinto. Cette entreprise doit déverser suffisamment d'eau par le déversoir de Skins Lake en été (mi-juillet à mi-août) pour maintenir la température moyenne quotidienne de l'eau dans la rivière Nechako en-dessous d'un seuil de 20°C à Finmoore. Un des outils utiliséspour cette modélisation et l'analyse hydrothermique est le modèle CEQUEAU, qui a été reprogrammé et amélioré au cours des 15 dernières années (St-Hilaire *et al.*, 2015).

Comme le changement climatique et la régulation des barrages peuvent menacer les habitats des poissons en modifiant le régime thermique, Rio Tinto a décidé de faciliter la recherche autour des opérations du déversoir de Skins Lake dans le contexte du changement climatique. Par conséquent, une équipe de centres de recherche composée de l'INRS, de

l'École de technologie supérieure (ETS), de l'Université de Victoria (UVIC) et de l'Université de la Colombie-Britannique (UBC) est impliquée dans un projet multidisciplinaire depuis 2017. Cette présente recherche fait partie de ce travail d'équipe. En plus des travaux de l'INRS sur le modèle CEQUEAU, différents modèles/méthodes ont été mis en œuvre: le modèle HEC-RAS (ETS) a été utilisé pour simuler les niveaux et la température dans le tronçon en aval du barrage. Les modèles VIC-GL, RBM et CE-QUAL-W2 (UVIC) ont été utilisés pour simuler la thermie dans le réservoir et des travaux de laboratoire sur la physiologie de certaines espèces de poissons ont été exécutés (UBC).

À ce jour, plusieurs études ont été réalisées à l'INRS et ailleurs sur le développement et la modernisation du modèle CEQUEAU (Dugdale *et al.*, 2017b; St-Hilaire *et al.*, 2015) sur différents bassins versants au Canada. Spécifiquement pour le bassin versant de la Nechako, les chercheurs ont étudié (1) les prévisions d'ensemble de la température de l'eau (Ouellet-Proulx *et al.*, 2017b), (2) l'assimilation des données de température de l'eau et de débit pour la prévision de la température de l'eau d'ensemble (Ouellet-Proulx *et al.*, 2017a), et (3) l'implication des méthodes d'estimation des pertes par évaporation dans la modélisation des débits et de la température de l'eau (Ouellet-Proulx *et al.*, 2019). Les connaissances antérieures sur la modélisation déterministe à l'aide du modèle CEQUEAU et les études antérieures sur le bassin versant de la Nechako ont fourni une base de connaissances solide pour la recherche actuelle.

1.2 Objectifs méthodologiques

Dans le cadre de l'exploitation des barrages et du changement climatique, de nombreuses questions se posent dont certaines seront abordées dans cette thèse. Elles sont liées à l'objectif principal de fournir aux gestionnaires des ressources en eau des outils de modélisation améliorés pouvant être utilisés dans le cadre d'une optimisation multi-objectifs de la gestion des réservoirs. Compte tenu de cet objectif principal, les questions de recherche et les objectifs spécifiques comprennent :

- Question 1) Le module thermique de CEQUEAU, qui, dans le passé, a été principalement calé à l'aide d'observations provenant d'une seule station pour un bassin versant entier, peut-il être calé adéquatement pour fournir des simulations de température à différents endroits sur un grand bassin versant comme la rivière Nechako? Le calage multisite peut-il améliorer la performance du modèle comparativement au calage sur site unique?
    - Objectif 1) Développer et comparer des méthodes de calage multisites pour le module thermique CEQUEAU.

- Question 2) Quels sont les impacts potentiels des changements climatiques sur le régime de température de l'eau de la rivière Nechako?
    - Objectif 2) Développer des scénarios hydrothermaux pour le système Nechako pour l'horizon 2040-2100 et étudier les impacts potentiels.

- Question 3) Il existe une variété d'équations pour estimer le rayonnement infrarouge (à ondes longues) incident. Quelles équations sont les meilleures candidates pour le module thermique de CEQUEAU?
    - Objectif 3) Étudier et valider différentes équations pour calculer le flux de chaleur du rayonnement à ondes longues et les comparer.

    - Objectif 4) Développer et comparer plusieurs modules incorporés dans le modèle CEQUEAU qui utilisent différents algorithmes pour calculer le flux de chaleur par rayonnement à ondes longues et les comparer.

1.3 Structure de la thèse

Le chapitre 1 introduit le contexte, les objectifs et la structure de cette thèse. Dans le chapitre 2, des concepts issus de la littérature sur la modélisation hydrothermique sont expliqués à partir des questions de recherche.

Les méthodes de recherche utilisées pour atteindre les quatre objectifs spécifiques de recherche ainsi que les résultats sont détaillés sous forme de synthèse au chapitre 3. Ce chapitre décrit également le bassin versant de la rivière Nechako, qui a servi d'étude de cas.

La conclusion générale et les perspectives de cette thèse sont présentées au chapitre 4. Dans ce chapitre, les réponses aux quatre questions de recherche sont décrites et présentées sous forme de conclusions. Le chapitre 5 est la bibliographie pour la synthèse (chapitres 1-4). Les cinq premiers chapitres sont rédigés en français.

Les chapitres 6 à 9 sont les articles publiés ou soumis de cette thèse. Par conséquent, chaque chapitre contient le manuscrit tel que soumis aux revues avec comité de lecture, rédigé en anglais.

Le chapitre 6 décrit les résultats d'une analyse du calage multisite de la température de l'eau sur un bassin versant. À cette fin, une zone d'étude de sous-bassin à l'intérieur du bassin versant de la Nechako a été sélectionnée. Les calages mono-site et multisite ainsi que la méthode de mise à l'échelle sont comparés. Ce chapitre aborde le premier objectif de recherche.

Les connaissances acquises en répondant au premier objectif de recherche sont utilisées pour les horizons passés et futurs afin d'étudier les impacts du changement climatique sur la température de l'eau. Cette approche, sa validation et la mise en œuvre de l'analyse de l'habitat pour le bassin versant de la Nechako sont présentées et discutées au chapitre 7. Ce chapitre aborde le deuxième objectif de recherche.

Dans le chapitre 8, le troisième objectif de recherche est abordé. Dans ce chapitre, utilisant les mesures basées sur les stations pour le rayonnement à ondes longues vers le bas, les meilleures formulations pour le bassin versant de la Nechako ont été étudiées. La possibilité

d'utiliser les données de réanalyse ERA5 au lieu de calculer le rayonnement à ondes longues est évaluée.

Les développements effectués au chapitre 8 sont ensuite évalués à l'échelle du bassin versant dans le chapitre 9. Ce chapitre est la réponse au quatrième objectif de recherche. À travers ce chapitre, de nouvelles méthodes de calcul du flux de chaleur associé aux longues longueurs d'onde sont incorporées dans le modèle CEQUEAU et comparées. Le niveau de ces améliorations a été analysé.

Chaque article (I-IV) comprend plusieurs supports qui sont présentés dans les annexes I à IV sur la base des quatre objectifs de recherche.

2 REVUE DE LITTERATURE

2.1 Modèles hydrologiques et thermiques

Les modèles hydrologiques peuvent être mis en œuvre à différentes échelles spatiales et temporelles, et ils quantifient les processus hydrologiques en fonction de ces échelles. Ils peuvent être globaux, semi-distribués ou pleinement distribués. Ces processus comprennent l'évaporation et la transpiration (souvent combinées en évapotranspiration), l'infiltration, la fonte des neiges, les flux souterrains et l'écoulement des cours d'eau. À chaque pas de temps, une équation de bilan hydrique est résolue:

$$P = R + ET + \Delta S + \Delta G \tag{2.1}$$

P représente les précipitations, R représente le ruissellement, ET représente l'évapotranspiration dans le bassin versant sur une période Δt, ΔS représente le changement du stockage de l'eau dans les eaux de surface et ΔG représente le changement du stockage des eaux souterraines (toutes les unités en m^3 ou en mm). En fonction de l'échelle spatiale ou temporelle, différents modèles utilisent d'autres formes de l'équation (2.1).

Les caractéristiques physiographiques du bassin versant (utilisation des terres, couverture végétale, pente, altitude, plans d'eau) doivent être numérisées. Le but de la discrétisation est de former des cellules de calcul, dans lesquelles des équations de conservation de la masse et de l'énergie peuvent être utilisées pour calculer des variables inconnues telles que le ruissellement. La cellule discrète est souvent appelée unité de réponse hydrologique (HRU). Il existe plusieurs méthodes pour former les HRU (Craig *et al.*, 2020; Fortin *et al.*, 2001; Noël *et al.*, 2014), dont trois sont présentées dans la Figure 2.1 basée sur Craig *et al.* (2020).

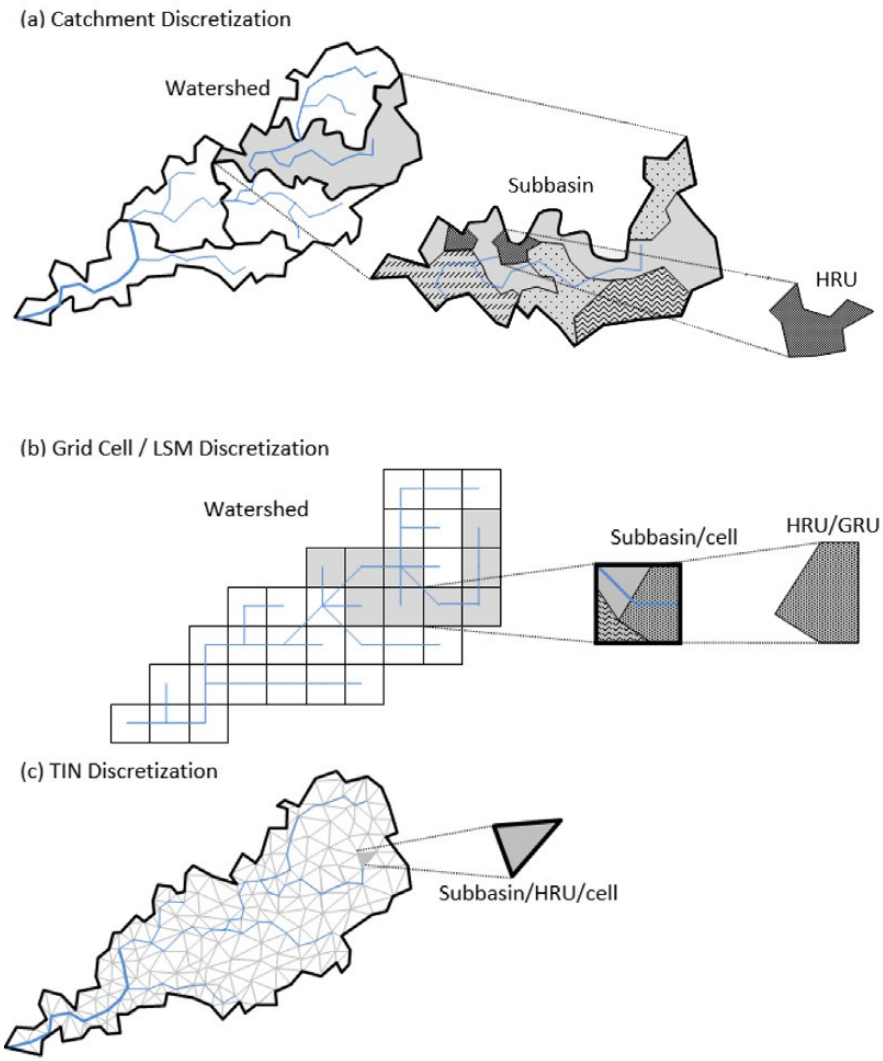

Figure 2.1: Trois schémas différents sont utilisés pour créer des unités de réponse hydrologique (HRU) sur un bassin versant [adapté de Craig *et al.* (2020)]

Il existe deux approches principales de modélisation de la température de l'eau, avec différents modèles dans chaque approche, variant en complexité et en exigences d'entrée de données. Ce sont les modèles statistiques et les modèles basés sur la physique (déterministes) de température de l'eau.

Benyahya *et al.* (2007) ont effectué une revue des modèles statistiques pour la prédiction de la température de l'eau. Ils ont classé les modèles statistiques de température de l'eau en deux catégories: paramétriques et non paramétriques. La catégorie paramétrique peut également être divisée en modèles de régression et modèles stochastiques. Les modèles de régression sont considérés comme plus adaptés pour des intervalles de temps supérieurs à une semaine (Benyahya *et al.*, 2007). Pour des intervalles de temps plus courts, d'autres modèles que ceux basés sur la régression sont mieux adaptés et plus faciles à appliquer. La

catégorie non paramétrique comprend des méthodes telles que les analogues, ou k-voisins les plus proches (kNN) et les réseaux de neurones artificiels (ANN), etc., que Zhu and Piotrowski (2020) ont inclus dans le groupe de modèles d'intelligence artificielle (AI).

Dugdale *et al.* (2017a) ont effectué une revue des modèles déterministes pour la prédiction de la température de l'eau. Cette catégorie de modèles est basée sur le calcul d'un bilan thermique. Le bilan thermique dans une portion de rivière peut être calculé en utilisant l'équation suivante:

$$\Delta H = H_{ups} + H_{local} + (H_s - H_{sr}) + (H_a - H_{br}) - H_e + H_c + H_p + H_{bed} \tag{2.2}$$

où les termes sont présentés dans le Tableau 2.1 et Figure 2.2 (toutes les unités dans MJ):

Tableau 2.1 Forme générale des termes de bilan thermique pour un plan d'eau

Terme (MJ)	Description
ΔH	changement du bilan thermique
H_{ups}	advection en amont
H_{local}	chaleur provenant des ruissellements locaux
H_s	rayonnement solaire entrant
H_{sr}	fraction de rayonnement solaire réfléchi
H_a	rayonnement infrarouge entrant
H_{br}	rayonnement infrarouge émis par le plan d'eau
H_e	chaleur latente
H_c	chaleur sensible
H_p	chaleur due aux précipitations liquides
H_{bed}	chaleur nette entrante ou sortant du lit de la rivière

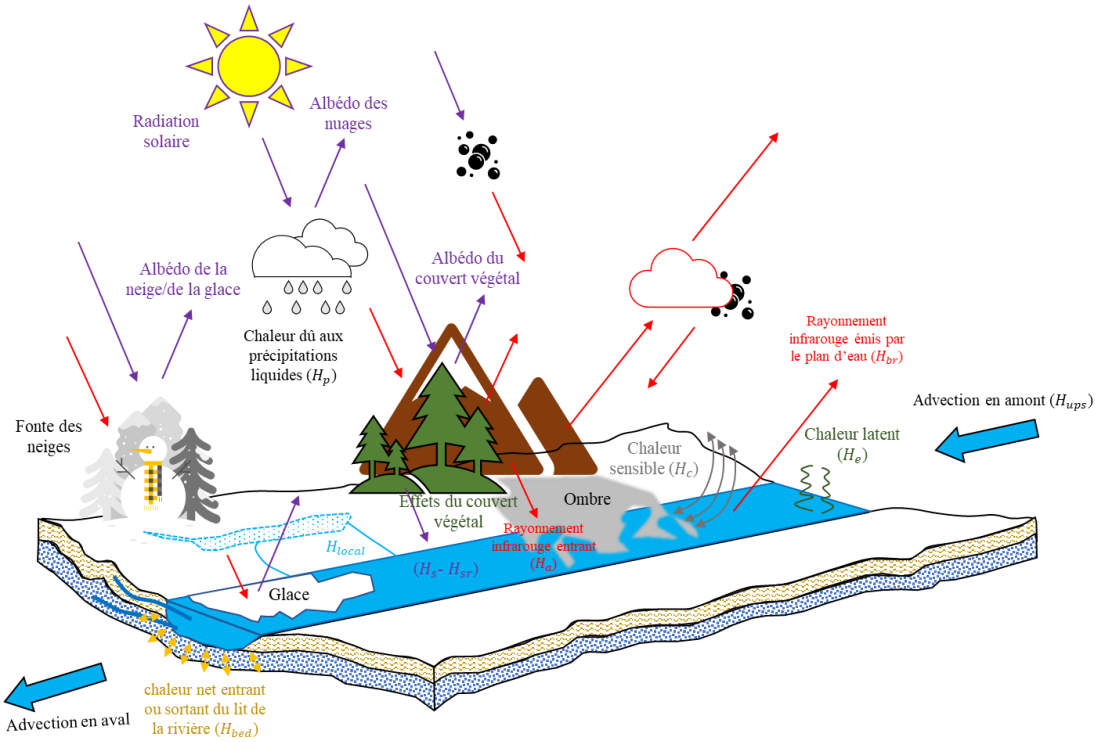

Figure 2.2: Représentation des termes du bilan thermique pour une section de rivière située à l'intérieur d'une unité de réponse hydrologique (HRU)

Différents modèles comprennent entièrement ou partiellement de chaleur de l'équation (2.2). Une fois la chaleur nette calculée, le changement de température peut être estimé en connaissant le volume d'eau à chauffer ou à refroidir et la capacité thermique de l'eau (4,186 MJ/°C/kg). Bien que de nombreux modèles hydrologiques bien connus puissent estimer ce volume d'eau à chaque période de temps, relativement peu d'entre eux disposent de modules thermiques pour la température de l'eau de ruissellement (Ficklin *et al.*, 2012). Pour ceux disposant d'un module thermique, la modélisation hydrologique ou hydraulique est une condition préalable à la modélisation de la température de l'eau.

Ouellet *et al.* (2020) ont effectué une revue critique de la littérature qui fournit un aperçu complet des progrès récents des outils de modélisation de pointe disponibles pour informer la recherche et la pratique écologiques. Ils fournissent des connaissances pour faire le lien entre les besoins scientifiques et les outils de modélisation disponibles dans la recherche sur la température des rivières. Ficklin *et al.* (2012) ont effectué une revue des différents modèles de température des rivières. Certains de ces modèles sont entièrement des modèles de température mécanistique qui sont généralement basés sur le calcul du bilan thermique de l'eau. Ils comprennent BasinTEMP, HSPF, SHADE-HSPF, HEAT SOURCE, QUAL2K, CE-QUAL-W2, CrUSTe, CEQUEAU, HEC-RAS, VIC-HMS, Streamline et SNTEMP. Ces modèles, ainsi que d'autres modèles disponibles, sont présentés au Tableau IV-B1, qui est initialement

basé sur Norton and Bradford (2009), Ficklin *et al.* (2012), Dugdale *et al.* (2017a) et Ouellet *et al.* (2020).

2.2 Scénarios futurs associés aux changements climatiques: implications pour la gestion du débit et de la température d'été en aval des barrages

Les écosystèmes aquatiques sont menacés par des pressions naturelles et anthropiques, notamment les changements d'utilisation des terres, les changements du régime d'écoulement et le changement climatique mondial (Algera *et al.*, 2022; Bosmans *et al.*, 2022). De toutes ces menaces, le changement climatique pourrait avoir les impacts les plus prolongés et à plus grande échelle (Nash *et al.*, 2017). La perte d'habitat a déjà de graves conséquences pour la vie aquatique, en particulier pour les salmonidés (Deinet *et al.*, 2020), avec une réduction de 76% de leur population mondiale et même plus dans certaines régions (93% en Europe) (Deinet *et al.*, 2020). On ne sait toujours pas quelle part de cette perte d'habitat est due au changement climatique. Cependant, il a été démontré que l'augmentation de la température de l'air a modifié le régime d'écoulement et les températures de l'eau en Amérique du Nord (Islam *et al.*, 2019). Les tronçons fluviaux situés en aval des ouvrages de barrage ont des débits gérés, ce qui peut exacerber (e.g., via des prescriptions de débits environnementaux inadéquates) ou atténuer (e.g., par des relâches d'eau) les impacts du changement climatique (Ahmad *et al.*, 2021; Algera *et al.*, 2022; Fullerton *et al.*, 2022; Sheedy, 2005; Sullivan & Rounds, 2021; Xiong *et al.*, 2020). Fullerton *et al.* (2022) ont suggéré la restauration de la ripisylve comme méthode possible pour contrer les impacts du changement climatique sur les saumons juvéniles. Sullivan and Rounds (2021) et Stratton Garvin *et al.* (2022) ont analysé les opérations de barrage comme méthode d'atténuation possible pour le même objectif.

Pour comprendre les impacts potentiels du changement climatique sur la température de l'eau, les chercheurs utilisent généralement des modèles thermiques (Dalcin *et al.*, 2023; Rajesh & Rehana, 2022; Simpson *et al.*, 2021). Ces modèles sont construits à partir de diverses sources de données, notamment des données météorologiques historiques, des données de réanalyse, des sorties de modèles climatiques et de modèles hydrologiques. Le processus de développement d'une étude sur le changement climatique pour la modélisation de la température de l'eau implique plusieurs étapes clés, notamment l'acquisition de données, la correction de biais, l'interpolation des données corrigées, le forçage sur le modèle hydrologique et la simulation du modèle pour les sorties de température de l'eau (Arsenault *et al.*, 2020; Bérubé *et al.*, 2022; Dallaire *et al.*, 2021; Tarek *et al.*, 2020a).

La première étape du processus consiste à acquérir les données. Cela peut impliquer la collecte de données de réanalyses (e.g., ERA5), de sorties de modèles climatiques globaux

et/ou locaux, et, dans le cas d'un module thermique, des débits/volumes simulés par des modèles hydrologiques et/ou hydrauliques. Une fois les données acquises, il est généralement nécessaire d'appliquer une correction de biais pour tenir compte des erreurs ou des inexactitudes dans les données des modèles climatiques (Cannon, 2018; Cucchi *et al.*, 2020). C'est une étape importante, car un biais dans les données peut entraîner des erreurs significatives dans les résultats du modèle final.

Une fois les données corrigées pour les biais, la prochaine étape consiste à interpoler les données corrigées pour créer une grille spatiale continue de données météorologiques et à forcer les données interpolées dans un modèle hydrologique. Il faut aussi coupler les données météorologiques avec d'autres données de surface terrestre, telles que la topographie, et la couverture terrestre (divers usages du territoire), pour simuler le cycle hydrologique et le mouvement de l'eau dans le système. Cette étape est cruciale, car elle permet aux chercheurs de modéliser les interactions complexes entre le climat, l'hydrologie et la température, et de comprendre comment les changements dans une variable peuvent avoir un impact sur l'ensemble du système. Enfin, le modèle est utilisé pour simuler des résultats de température de l'eau (Ouellet-Proulx, 2018; Ouellet-Proulx *et al.*, 2019).

2.3 Modélisation de l'irradiance descendante à ondes longues à la surface de l'eau

Le rayonnement à ondes longues (LR) est l'un des principaux termes du bilan thermique responsable de multiples processus environnementaux dans les milieux terrestres et aquatiques (Bernhardt *et al.*, 2022; McFarlane & Clark, 2021; Ouellet *et al.*, 2014). L'échange de LR est contrôlé par l'émissivité thermique de la surface, les objets atmosphériques, les nuages, la vapeur d'eau et le dioxyde de carbone (Herrero & Polo, 2012). Il existe deux variables liées à la LR (rayonnement incident et réémis vers le haut); l'irradiance descendante à ondes longues (DLI) agit comme une signature de l'état de l'atmosphère (Tang *et al.*, 2021). L'un des rôles de la LR est le possible refroidissement de l'eau pendant l'été, ce qui a un impact significatif sur la flore et la faune aquatiques. Les travaux décrits aux chapitres 2 et 3 ont démontré que cette composante du bilan thermique est en partie responsable du refroidissement de la température de l'eau dans le bassin versant de Nechako, dans le centre de la Colombie-Britannique, au Canada. Par conséquent, il est nécessaire d'utiliser un bilan radiatif à ondes longues plus précis dans cette région pour la modélisation du bilan énergétique.

Au cours des dernières décennies, de multiples études ont développé les fondements théoriques de la DLI (Zhou *et al.*, 2020). Les principaux contrôleurs de l'émissivité de la DLI sont la température de l'air, la couverture nuageuse et la pression de vapeur d'eau. Bien que

la physique de la DLI soit bien connue sur la base de la loi de Stephan-Boltzmann, l'étude de l'interaction des variables et les méthodes de calcul pour estimer l'émissivité atmosphérique sont encore en développement. En conséquence, plusieurs études ont paramétré et étudié la DLI à l'aide de différentes équations (Bárbaro *et al.*, 2009; Duarte *et al.*, 2006; Herrero & Polo, 2012; Huang *et al.*, 2007; Koll & Cronin, 2018; Sicart *et al.*, 2010).

Avec les progrès accomplis en ce qui a trait à l'acquisition de données, l'impact de la couverture nuageuse sur la DLI a été étudié de manière significative. Duarte *et al.* (2006) ont étudié la DLI en conditions de ciel clair et de ciel nuageux en utilisant des paramétrisations bien connues dans le sud du Brésil; Ces mêmes auteurs ont fait de même pour les glaciers andins. Herrero and Polo (2012) ont étudié la DLI pour un site montagneux, et Voortman *et al.* (2015) pour les écosystèmes de dunes.

Des études plus récentes ont accordé plus d'attention à tous les paramétrages des conditions du ciel (Li *et al.*, 2017; Long *et al.*, 2021). Par exemple, Li *et al.* (2017) ont comparé et recalibré quinze modèles paramétriques pour la DLI. Ils ont recommandé la famille de modèles de Brunt (Brunt, 1932) pour une utilisation future. Long *et al.* (2021) ont testé six modèles pour des conditions de ciel clair et huit pour des conditions nuageuses sur le plateau tibétain. Leurs résultats ont montré une RMSE moyenne de 26,4 W/m^2 pour des conditions de ciel clair. L'analyse de l'impact de la couverture nuageuse sur la DLI a convergé vers la comparaison avec les données de réanalyse. Tang *et al.* (2021) ont montré que les produits de réanalyse pour la DLI surpassent les données issues de satellites sur la surface terrestre. Ils ont souligné la grande possibilité d'utilisation de la réanalyse pour la modélisation DLI.

De plus, l'utilisation des données de réanalyse comme base de données pour les variables météorologiques et la DLI a été recommandée dans la littérature. Cependant, à notre connaissance, il n'existe pas de travaux de recherche antérieurs pour les régions où les données sont rares et qui peuvent bénéficier des données de réanalyse en tant qu'alternative aux mesures météorologiques aux stations pour calculer la DLI.

L'émissivité de l'air pour la DLI est essentiellement liée à la pression de vapeur et à la couverture nuageuse dans la colonne d'atmosphère (Niemelä *et al.*, 2001). Quel que soit le milieu d'étude, la DLI à la baisse affecte le système lotique sous deux formes: (a) comme un flux positif dans le bilan thermique de l'eau lorsque la température de l'air est supérieure à celle de l'eau ($T_{air} - \sqrt[4]{\frac{\beta_{water}}{\beta_{air}}} T_{water} > 0$) ou (b) en agissant comme un mécanisme de refroidissement dans le cas contraire ($T_{air} - \sqrt[4]{\frac{\beta_{water}}{\beta_{air}}} T_{water} < 0$).

La comparaison des modèles de DLI pour la modélisation hydrologique est possible en comparant les flux de DLI estimés avec les mesures sur le terrain (Ouellet *et al.*, 2014). En

raison d'obstacles économiques, logistiques et techniques, peu d'études ont effectué la validation de la DLI à l'aide de mesures sur le terrain (Brown *et al.*, 2021; Tang *et al.*, 2021). Par conséquent, la solution de substitution est la validation indirecte des modules de DLI en tant que composante de modèles thermiques (Ouellet-Proulx *et al.*, 2019). Bien qu'il existe des études comparant différents algorithmes pour d'autres termes de bilan thermique comme l'évaporation et l'évapotranspiration (e.g., Ouellet-Proulx et al., 2019), il n'existe aucune étude antérieure pour l'évaluation de l'impact de l'estimation de la DLI à l'aide de la modélisation hydrologique et thermique. À notre connaissance, Ouellet *et al.* (2014) ont réalisé la seule étude similaire, quoique dans un environnement expérimental contrôlé.

Tang *et al.* (2021) ont montré que le produit ERA5 pour la DLI surpasse en qualité les données issues des images satellites sur les zones terrestres. Plusieurs variables de la base de données ERA5 ont été évaluées à l'échelle de la station ou du bassin versant. Par exemple, les précipitations et la température (Taghizadeh *et al.*, 2021; Tarek *et al.*, 2020b), la couverture nuageuse (Yao *et al.*, 2020), la température de l'eau (Gatien *et al.*, 2022; Khorsandi *et al.*, 2022) , et les variables d'irradiance (Jiang *et al.*, 2020) sont quelques-unes des variables climatiques étudiées dans la littérature.

3 SYNTHESE DES TRAVAUX DE RECHERCHE

3.1 Sites d'étude

Le bassin versant de la rivière Nechako est situé dans la partie centrale de la Colombie-Britannique, au Canada, avec un bassin versant de 12 804.7 km^2 (Figure 3.1). En aval du déversoir du lac Skins (SLS), le débit de la rivière est entièrement régulé par l'évacuateur de crue entre le réservoir Nechako et sa confluence avec la rivière Nautley, après quoi les débits des rivières Nechako et Nautley se combinent et continuent au-delà de la ville de Vanderhoof et éventuellement dans le fleuve Fraser. Trois stations hydrométriques mesurent le débit (Environnement et Changement climatique Canada, https://www.canada.ca/fr/environnement-changement-climatique/, accès aux données: mai 2022). Elles sont localisées à SLS, à la rivière Nautley (en amont de la confluence avec la rivière Nechako) et à Vanderhoof (voir Figure 3.1). Les débits et les températures prises immédiatement en aval du déversoir SLS fournissent les conditions aux limites amont des modules hydrologiques et thermiques et ne sont pas utilisés pour le calage hydrologique. Huit stations mesurent la température de l'eau dans le bassin versant (Figure 3.1). La zone d'étude et les stations de suivi sont décrites en détail dans Ouellet-Proulx *et al.* (2017a) et au Chapitre 6.

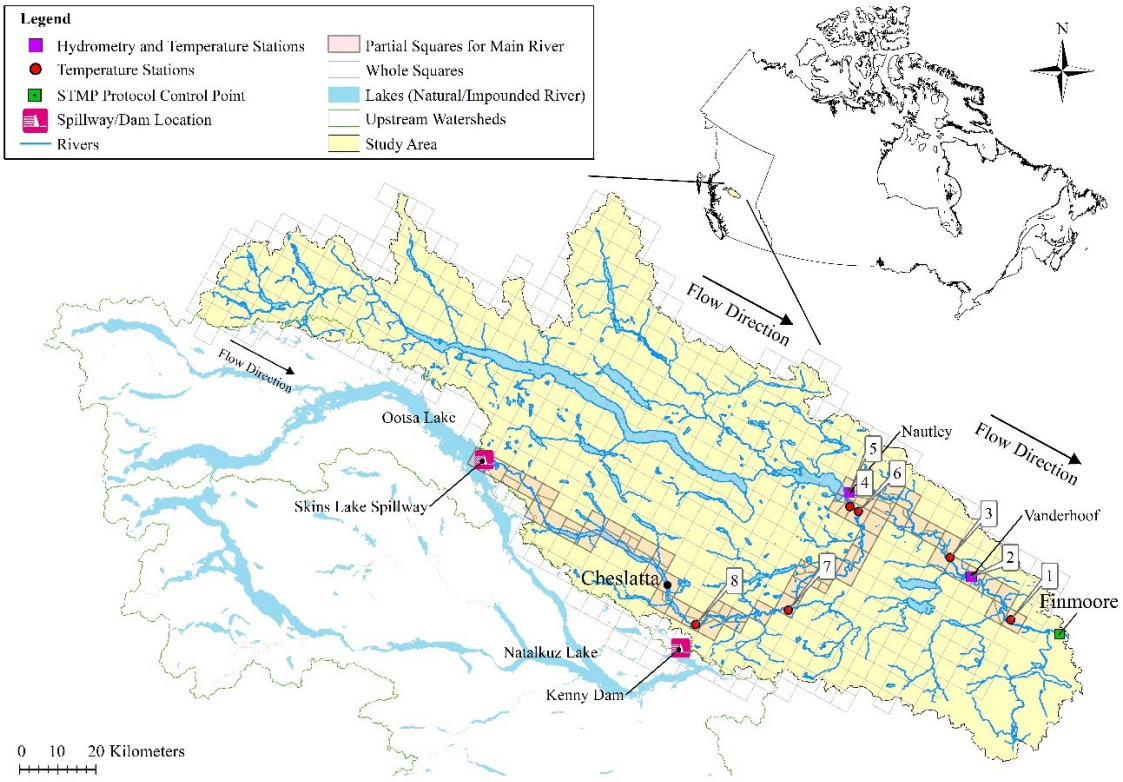

Figure 3.1 Bassin versant de la rivière Nechako et zone d'étude

3.2 Modèle CEQUEAU

CEQUEAU est un modèle hydrologique conceptuel semi-distribué utilisé pour la simulation et la prévision des débits et de la température de l'eau (Morin & Couillard, 1990). Le module hydrologique permet de simuler le débit à des pas de temps quotidiens ou intra-journaliers et a été utilisé pour la prévision du débit et/ou de la température de l'eau dans de nombreuses études de cas [e.g., Bâ *et al.* (2018), Brisson *et al.* (2015), et Ouellet-Proulx *et al.* (2019)]. La version actuelle du modèle CEQUEAU utilisée dans cette étude est basée sur des paramètres globaux pour l'ensemble du bassin versant.

La couverture terrestre et la topographie sont des données d'entrée physiographiques requises (Dugdale *et al.*, 2017b). Les données d'occupation des sols les plus récentes, avec une résolution de 10 m fournies par ESRI et l'Agence spatiale européenne (ESA), ont été utilisées (Karra *et al.*, 2021; Zanaga *et al.*, 2021). Le modèle numérique d'élévation NASA SRTM (DEM) avec une résolution de 30 m (Farr *et al.*, 2007) est utilisé pour calculer les élévations dans la structure d'entrée.

La spatialisation des processus hydrologiques se fait par une discrétisation en deux étapes du bassin versant. Premièrement, le bassin versant est divisé en zones égales. Ce sont des éléments carrés appelés «carreaux entiers» ou CE (Figure 7.2a). L'altitude, le pourcentage de superficie couverte par la forêt et le pourcentage convert par les lacs et les marais sont définis pour chaque CE. Deuxièmement, chaque CE est subdivisé en un maximum de quatre «carreaux partiels» (CP) en superposant les CE avec les limites des sous-bassins pour déterminer la direction du ruissellement (Figure 7.2a). Chaque carreau partiel est une unité de réponse hydrologique (HRU), dans laquelle le cheminement vertical et le stockage de l'eau sont représentés conceptuellement par trois réservoirs interconnectés. Un réservoir représente le stockage par les lacs et les marais, tandis que les deux autres représentent la zone de vadose (non saturée) et les couches inférieures du sol (Figure 7.2d). L'eau disponible dans les différents réservoirs est acheminée vers l'aval proportionnellement à la surface du carreau partiel à chaque pas de temps.

Le module hydrologique calcule un bilan hydrique, avec en intrant la fonte des neiges et/ou les précipitations liquides comme apports, le stockage dans des réservoirs et l'évapotranspiration comme terme négatif pour le bilan massique.

Ensuite, le module thermique calcule un bilan de chaleur de l'eau qui inclut le rayonnement direct à ondes courtes comme entrée principale et la chaleur latente comme perte principale. Le rayonnement net à ondes longues et la chaleur sensible peuvent être positifs ou négatifs, selon le gradient thermique de l'interface air-eau. L'advection de chaleur provenant de l'amont,

le ruissellement local, le ruissellement hypodermique et l'apport d'eau souterraine sont pris en compte dans le bilan thermique (Figure 7.2b,c).

Après avoir calculé l'eau disponible dans chaque CP, le module thermique calcule les sources et puits de chaleur suivants (St-Hilaire *et al.*, 2015):

$$\Delta H = H_{short} + H_{long} + H_{ev} + H_{sens} + H_{ups} + H_{local} + H_{ground} \tag{3.1}$$

où ΔH est la variation globale du budget thermique pour un CP (toutes les unités en MJ); H_{short} est le rayonnement entrant à ondes courtes; H_{long} est le gain ou la perte de chaleur par rayonnement à ondes longues; H_{ev} est la perte de chaleur latente (évaporation); H_{sens} est le gain ou la perte de chaleur sensible; H_{ups} est le gain de chaleur depuis l'amont; H_{local} est la chaleur provenant du ruissellement de surface et hypodermique local; et H_{ground} est le terme de chaleur lié à l'apport d'eau souterraine. Les termes du bilan thermique de surface sont estimés à l'aide des équations suivantes:

$$H_{short} = CRAYSO.A.R_{short} \tag{3.2}$$

$$H_{long} = CRAYIN.A.\sigma(\beta T_{air}^4 - T_{water}^4) \tag{3.3}$$

$$H_{ev} = -CEVAPO.L_{ev}.A.E \tag{3.4}$$

$$H_{sens} = CCONVE.A[0.2U(T_{air} - T_{water})] \tag{3.5}$$

où A est la surface d'eau dans le CP (m^2); $CRAYSO$, $CRAYIN$, $CEVAPO$, et $CCONVE$ sont des coefficients empiriques sans dimension pour le calage du module thermique (présentés dans le Tableau 6.2) ; R_{short} est le rayonnement solaire net (MJ/m^2); σ est la constante de Stefan-Boltzmann (5.6704×10^{-8} W/m^2/K^4 or 4.8992×10^{-9} MJ/m^2/K^4/day); β est l'émissivité du ciel (0-1); T_{air} et T_{water} sont respectivement la température de l'air et de l'eau (K); L_{ev} est la hauteur de l'eau évaporée du tronçon de la rivière, qui est calculée par le module hydrologique (m); E est la chaleur latente d'évaporation (2480 MJ/m^3; en supposant que la densité de l'eau est invariante, acceptable pour la plage de température de l'environnement) et U est la vitesse du vent (km/h).

Après avoir calculé tous les termes de chaleur, le module thermique calcule le changement de température de l'eau à l'aide de l'équation suivante:

$$\Delta T = \frac{\Delta H}{V.\theta} \tag{3.6}$$

où ΔT est la variation de la température de l'eau (°C); V est le volume d'eau estimé par le module hydrologique (m^3), et θ est la capacité thermique de l'eau (4.186 MJ/kg/°C). Tous les termes du bilan thermique et leur formulation sont expliqués en détail dans le manuel

CEQUEAU (Morin & Paquet, 2007). L'architecture modulaire du modèle permet d'ajouter et de modifier relativement facilement des algorithmes, des routines ou des composantes hydrologiques/thermiques (St-Hilaire *et al.*, 2015).

Le module hydrologique de CEQUEAU a 26 paramètres globaux (c'est-à-dire que la valeur du paramètre est constante pour tous les CP; Tableau I-A1 de l'annexe I-A). Le module thermique a huit paramètres globaux pour les processus de bilan thermique sur l'ensemble du bassin versant (Tableau 6.2). Une représentation schématique du module hydrologique avec identification des paramètres est disponible dans Morin and Couillard (1990) et Morin and Paquet (2007), également décrit dans la Figure 3.2.

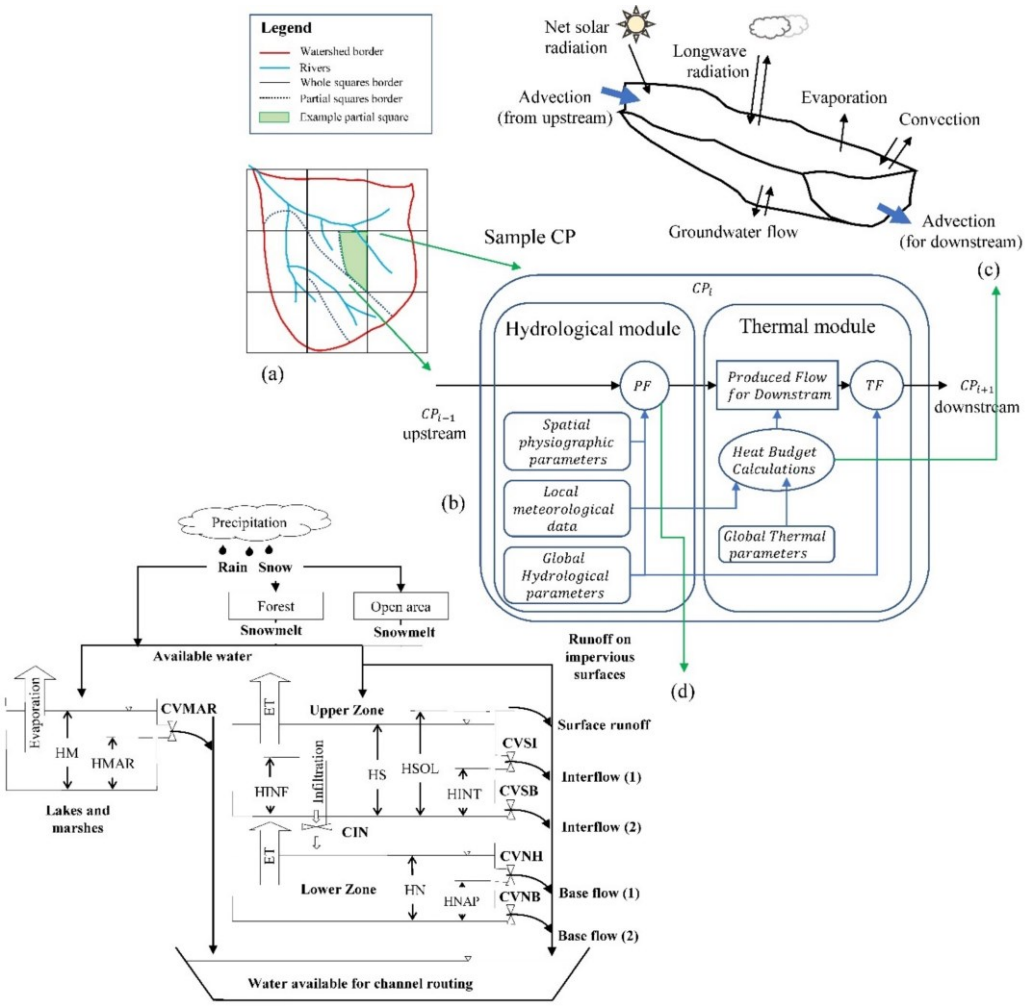

Figure 3.2 **Représentation schématique de (a) un bassin versant, de carrés entiers et de carrés partiels, (b) chaque carré partiel (CP) est une unité de réponse hydrologique pour laquelle les budgets hydrologiques et thermiques sont calculés. Le cœur du module hydrologique est constitué de la fonction de production (PF) et de la fonction de transfert (TF), qui calcule l'eau disponible à l'intérieur de chaque CP et procède au routage subséquent vers l'aval. (c) Le module thermique calcule le budget thermique sur chaque CP. L'énergie thermique est également transférée vers l'aval; (d) montre la PF schématiquement, qui se compose de trois réservoirs interconnectés pour modéliser les processus hydrologiques à l'intérieur de chaque CP.**

3.3 Données de réanalyse ERA5

Il est nécessaire de disposer de données météorologiques complètes pour calibrer et valider le modèle hydrologique pour la période historique comme référence à comparer avec les scénarios futurs. Cependant, il n'existe pas d'ensemble complet de données météorologiques observées pour le bassin dans le passé qui puisse être utilisé comme ensemble de données de référence. Par conséquent, une alternative consiste à utiliser des produits de réanalyse climatique comme suggérés par Gatien *et al.* (2022).

En plus des précipitations et de la température de l'air, le module de température de l'eau CEQUEAU nécessite la vitesse du vent, la pression de vapeur d'eau et le rayonnement solaire net à ondes courtes. Ces variables météorologiques sont fournies par le Centre européen de prévision météorologique à moyen terme (ECMWF) via leur *European Reanalysis 5th génération* (ERA5) (Hersbach *et al.*, 2020a), à l'exception de la pression de vapeur qui a été calculée à l'aide de la température du point de rosée de ERA5 à 2 m de hauteur et de l'équation de Teten (Monteith & Unsworth, 2013; Murray, 1967; Ouellet-Proulx, 2018). Ces données d'entrée maillées (avec une résolution de 31 × 31 km) ont été interpolées à tous les points de grille du bassin à l'aide de l'interpolateur CEQUEAU intégré basé sur l'approche des plus proches voisins (Ouellet-Proulx *et al.*, 2019). Les périodes de temps temporels initiales pour ERA5 et les modèles climatiques qui étaient horaires ont été converties en période de temps quotidiens pour exécuter le modèle CEQUEAU.

3.4 Algorithme de calage: CMA-ES

Le «Covariance Matrix Adaptation Evolution Strategy» (CMA-ES) est un algorithme évolutionnaire développé par Hansen and Ostermeier (1996). La supériorité de CMA-ES en tant que méthode d'optimisation globale capable de gérer une dimensionnalité élevée en fait un bon candidat pour l'optimisation (Hansen & Ostermeier, 1996). Cet algorithme a déjà été comparé à d'autres algorithmes d'optimisation connus pour le calage de modèles en hydrologie par Arsenault *et al.* (2014). Ces derniers auteurs ont montré que CMA-ES offre des performances supérieures à la plupart des algorithmes d'optimisation pour (1) trouver des optima globaux et (2) pour sa vitesse de convergence. L'annexe I-B explique que cet algorithme suit quatre étapes jusqu'à ce qu'un critère d'arrêt soit satisfait (Hansen, 2023). Les critères d'arrêt comprennent le nombre d'évaluations, le seuil de changement de la fonction objectif et le seuil de changement dans les variables de décision.

3.4.1 Calage du modèle hydrologique

La fonction objectif de l'algorithme d'optimisation est le coefficient d'efficacité de Kling-Gupta (KGE; Gupta et al., 2009) , décrit aux équations (3.7) à (3.11). Étant donné que chaque métrique d'efficacité a ses propres forces et limites, la présentation d'une seule métrique peut être biaisée ou ne pas refléter avec précision le succès du calage. En plus du KGE, les métriques d'efficacité du biais et de Nash-Sutcliffe [NSE; Nash and Sutcliffe (1970)] ont également été calculées.

$$KGE_j = 1 - ED_j \tag{3.7}$$

$$ED_j = \sqrt{(r_j - 1)^2 + (\alpha_j - 1)^2 + (\beta_j - 1)^2} \tag{3.8}$$

$$r_j = \frac{\sum_{i=1}^{N}\left(O_{i,j} - \bar{O}_j\right)\left(S_{i,j} - \bar{S}_j\right)}{\sqrt{\sum_{i=1}^{N}\left(O_{i,j} - \bar{O}_j\right)^2}\sqrt{\sum_{i=1}^{N}\left(S_{i,j} - \bar{S}_j\right)^2}} \tag{3.9}$$

$$\alpha_j = \frac{\sigma_{S\,j}}{\sigma_{O\,j}} \tag{3.10}$$

$$\beta_j = \frac{\bar{S}_j}{\bar{O}_j} \tag{3.11}$$

$$Bias_j = \bar{O}_j - \bar{S}_j \tag{3.12}$$

$$NSE_j = 1 - \frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{\sum_{i=1}^{N}\left(O_{i,j} - \bar{O}_j\right)^2} \tag{3.13}$$

où KGE_j, $Bias_j$ and NSE_j sont les métriques pour la j ème station; $O_{i\,j}$ et $S_{i\,j}$ sont respectivement les valeurs observées et simulées pour la j ème station au i ème pas de temps; N est le nombre de pas de temps (mesures); $\sigma_{S\,j}$ et $\sigma_{O\,j}$ sont l'écart-type des simulations et des observations pour la station j; la $\bar{S}_j$ et $\bar{O}_j$ sont la moyenne des valeurs simulées et observées.

3.4.2 Calage du modèle thermique

Les huit paramètres du modèle thermique ont été ajustés en utilisant la racine de l'erreur quadratique moyenne (RMSE) comme fonction objectif et l'algorithme CMA-ES pour l'optimisation. La RMSE a été sélectionnée parce qu'elle est la plus fréquemment utilisée dans la littérature [e.g., Ouellet-Proulx *et al.* (2017a) et Ouellet-Proulx *et al.* (2019)].

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{N}} \tag{3.14}$$

Comme pour le calage hydrologique, toutes les stations et toutes les données observées ont été utilisées pour le calage du modèle thermique. Les calages hydrologiques et thermiques ont été réalisés en deux étapes. Premièrement, les limites inférieures et supérieures des paramètres du modèle ont été examinées manuellement pour s'assurer que les simulations restaient dans des limites réalistes lorsque le paramètre avait une signification physique. Suite à ce calage manuel, l'algorithme d'optimisation CMA-ES a été utilisé pour affiner les valeurs des paramètres du modèle des modules hydrologiques et thermiques. Ces étapes sont expliquées en détail au Chapitre 6.

3.4.3 Calage des coefficients dans les équations DLI

Cette étude a utilisé la RMSE comme fonction objectif pour tous les calages de modèles DLI, visant à trouver les valeurs optimales pour les coefficients de chaque modèle. La RMSE a été utilisée pour calibrer les paramètres puisqu'il s'agit d'une mesure d'efficacité générale sans se concentrer sur une partie spécifique des données (Gupta *et al.*, 2009; Legates & McCabe Jr, 1999). En même temps, il a été démontré qu'aucun des critères d'efficacité ne fonctionne idéalement pour évaluer la qualité de l'ajustement (Krause *et al.*, 2005). Par conséquent, plusieurs mesures d'efficacité ont été calculées comme fortement recommandé par Legates and McCabe Jr (1999) et Gauch *et al.* (2023).

La racine de l'erreur quadratique moyenne relative ($RRMSE$) a été utilisée pour montrer la compétence de chaque équation sans dimension, à la fois pour les étapes de calage et de validation. L'efficacité de Nash–Sutcliffe (NSE) montre une efficacité d'ajustement générale, ainsi que la performance d'ajustement pour des valeurs élevées. La corrélation (r) et le coefficient de détermination (R^2) ont été présentés comme recommandé par Legates and McCabe Jr (1999) en même temps que les autres mesures. Étant donné que l'utilisation de plus de paramètres et de variables d'entrée peut entraîner un surajustement, il est nécessaire de choisir de meilleurs modèles tout en gardant les modèles parcimonieux. Dans ce but, le critère d'information d'Akaike (AIC) (Akaike, 1974) et le critère d'information bayésien (BIC) (Schwarz, 1978) ont également été calculés pour aider à sélectionner les meilleurs modèles tout en considérant le nombre de paramètres. Enfin, une simple métrique de biais a été calculée pour montrer le niveau de biais entre les simulations et les données observées.

3.5 Calage multisite d'un modèle hydrologique et thermique semi-distribué dans un grand bassin versant canadien

Dans la présente étude, trois méthodes ont été comparées pour le calage du modèle:

1) Méthode 1: Calage sur un site unique et calcul de la métrique de performance pour ce site et les autres sites indépendants avec le même jeu de paramètres;
2) Méthode 2: Optimisation d'une métrique multisite en utilisant toutes les observations de tous les sites comme un seul problème d'optimisation objectif. Puis, à partir de l'optimum multisite, la métrique est recalculée sur chaque site pour validation; et
3) Méthode 3: Cette méthode déduit les paramètres à l'échelle globale à l'aide des estimations à l'échelle du sous-bassin versant de la méthode 1. Cette méthode met à l'échelle les valeurs des paramètres de sites uniques et utilise les valeurs mises à l'échelle à tous les sites. Pour cette méthode, différents opérateurs sont possibles et sont expliqués dans la section suivante.

La métrique de performance KGE a été utilisée pour le calage hydrologique et le RMSE a été utilisé pour le calage du module thermique. La métrique de biais absolu est utilisée pour évaluer le biais éventuel des résultats.

Étant donné que le régime thermique d'une rivière dépend en partie des conditions d'écoulement qui prévalent, le module hydrologique a été calibré une fois, pour toutes les données disponibles de 2016 à 2019. Ensuite, à l'aide du module hydrologique calibré, les calages du module thermique ont été effectués. Les calages thermiques ont été effectués en utilisant la période allant d'avril à novembre en utilisant un pas de temps quotidien.

Sur la base de la «méthode de calage» et de la «station cible», 11 scénarios différents ont été définis. Pour chacune des huit variables de décision, nous avons utilisé la valeur moyenne qui a été calculée à partir des limites inférieure et supérieure. Ces valeurs moyennes ont été utilisées comme ensemble de paramètres initiaux pour tous les scénarios de calage. Le processus de calage et de validation hydrologique/thermique du modèle CEQUEAU est résumé dans la Figure 6.5.

Le calage du module hydrologique a été complété en comparant observations et simulations pour la station de Vanderhoof, qui est la seule station avec des mesures de débit à long terme. Le KGE de 0,96 et le NSE de 0,94 montrent que le calage est de haute qualité et peut reproduire avec précision l'hydrogramme de Vanderhoof.

Le Tableau 6.6 montre les paramètres thermiques après calage pour les 11 scénarios. Pour quantifier l'écart total des optima, les RMSE globaux ont été calculés dans la dernière colonne. Il convient de mentionner que des valeurs RMSE inférieures à 2°C sont typiquement considérées comme étant le reflet de simulations adéquates.

Comme prévu, pour chaque site individuel, la méthode 1 montre les meilleures performances. Cependant, le fait de se concentrer uniquement sur une station dégrade la mesure de performance sur les autres sites, donc nous avons choisi la $RMSE_{Globale}$ comme référence pour la comparaison entre les trois méthodes de calage. La $RMSE_{Globale}$ varie de 1,57°C (pour la méthode multisite) à 2,69°C (pour la simulation basée sur les paramètres obtenus à partir du calage à la station la plus en amont). L'ordre de grandeur des $RMSE_{Globale}$ pour les stations 1 à 4 (1,61 à 1,78°C) diffère des stations 5 à 7 (2,16 à 2,69°C).

En analysant l'emplacement de toutes les stations sur le bassin versant, il est évident que les stations 5 à 7 sont situées à proximité de plans d'eau importants en amont. La station 7 est immédiatement en aval du lac Cheslatta et la station 6 est proche du vaste tronçon en forme de méandre de la rivière Nechako, près du lac Copley. La station 5 est également susceptible d'être affectée par un grand tronçon sinueux situé en amont. Pour les stations 1 à 4, il n'y a pas de plan d'eau important avec écoulement lent immédiatement en amont des stations, de sorte que la température dans ces tronçons de la rivière est mieux représentée.

Les principaux flux de chaleur sur le bassin versant sont le rayonnement à ondes courtes, le rayonnement à ondes longues et les flux de chaleur sensible. La Figure 6.7 montre que la principale source de chaleur est le rayonnement solaire à ondes courtes (comme prévu), tandis que les principaux puits de chaleur sont le rayonnement à ondes longues et la chaleur sensible.

Bien que le calage sur un seul site ait donné les meilleurs optima locaux pour chaque site, la méthode de calage multisite est supérieure sur la base des résultats de cette étude. Il est recommandé d'utiliser toutes les stations disponibles pour le calage thermique, car le regroupement de toutes les stations et l'exécution d'un seul calage multisite sont plus rapides, avec une meilleure métrique de performance globale. Ce résultat confirme les résultats de Shen *et al.* (2022) et Arsenault *et al.* (2018) pour le calage hydrologique. Nos résultats ont montré des résultats similaires pour la modélisation thermique, et il est recommandé d'utiliser toutes les données disponibles pour le calage du modèle (calage multisite) et la prise de décision.

3.6 Scénarios futurs de débit et de température de l'eau dans un bassin versant aménagé : implications pour la gestion du débit et de la température d'été en aval du barrage

Suite à son calage, le modèle CEQUEAU a été utilisé pour simuler la température de l'eau sur le bassin versant de la Nechako pour deux périodes futures, 2040-2070 et 2070-2100, à l'aide de huit modèles climatiques globaux (GCM) (Tableau II-A4). Pour montrer l'ampleur du changement par rapport à la période de référence (pour 1980-2019, en utilisant l'ensemble

de données ERA5) pour ces deux périodes futures, les scénarios de forçage radiatif de 4,5 et 8,5 W/m^2 ont été sélectionnés sur la base de la terminologie convenue des voies socio-économiques partagées (SSP) dans la version 6 du projet d'intercomparaison de modèles couplés (CMIP6) (Tableau II-A5). Parmi tous les modèles CMIP6, ces huit modèles ont été sélectionnés pour mener cette étude, car ce sont eux qui ont toutes les variables requises à des pas de temps journalier et sous-journalier. Les pas de temps sous-journaliers ont été utilisés pour procéder à la correction du biais, de manière similaire à Gatien *et al.* (2022) pour des études comparatives inter-modèles. Chaque modèle climatique global utilisant des forçages SSP a fourni toutes les données météorologiques pour exécuter le modèle, qui ont été débiaisés à l'aide de l'algorithme MBCn (Cannon, 2018) avec les données ERA5 comme ensemble de données de référence pour deux scénarios de forçage radiatif (4,5 et 8,5 W/m^2). Huit modèles climatiques ont été utilisés, comme présenté dans le Tableau II-A4.

Considérant que le modèle CEQUEAU simule la température de l'eau selon un pas de temps journalier, les variables sous-journalières ont été agrégées pour générer des séries temporelles météorologiques quotidiennes. Enfin, les projections météorologiques quotidiennes ont été forcées dans CEQUEAU, et le modèle a été exécuté sur les deux périodes mentionnées ci-dessus pour générer des projections futures de débit et de température de l'eau. Chaque GCM a un réseau maillé différent, qui a été interpolé pour les points de grille CEQUEAU en utilisant l'approche des plus proches voisins.

Des courbes de durée de température (ou températures classées; TDC) ont été utilisées pour comparer et analyser les simulations quotidiennes de température de l'eau pour les périodes passées et futures. La TDC trace la probabilité de dépassement de la valeur observée à partir de plusieurs seuils de manière continue (Karakoyun *et al.*, 2018). Cette étude a identifié la probabilité de dépassement du seuil thermique critique de 20°C (Macdonald, 2019) sur les courbes TDC parce que ce seuil est celui qui est utilisé dans le protocole de déversement à SLS pour atténuer les impacts thermiques durant la migration du saumon rouge (Sockeye).

Pour les stations avec des périodes d'observation plus longues et sans plans d'eau en amont, le modèle CEQUEAU peut fournir des simulations avec une RMSE < 2°C et un biais < 0,8°C (Stations 1, 2, 3, 4 et 5). Cependant, trois des huit stations (stations 6, 7 et 8) affichent des métriques de biais supérieures à 1,5°C et des RMSE supérieurs à 2°C. Encore une fois, ce biais est probablement dû à la présence de grands plans d'eau en amont de ces stations. L'impact des grands lacs sur les simulations de température de CEQUEAU est discuté plus en détail au chapitre 7 (Khorsandi *et al.*, 2022). La station de Vanderhoof (Station 2) est d'une plus grande importance car elle correspond à l'endroit identifié dans le Programme de gestion des températures estivales (STMP), où la température quotidienne moyenne de l'eau doit demeurer ≤ 20°C durant la migration du saumon rouge.

Les principaux résultats décrivent l'augmentation probable de la température de l'eau à la station de Vanderhoof. Les données de température de l'eau observées par le passé et les simulations historiques à Vanderhoof montrent des dépassements similaires du seuil de 20°C (1% pour les deux; avec des observations de 2005 à 2019).

Les résultats simulés pour cette station pour la période de référence montrent une RMSE = 1,27°C et un biais ≈ 0°C, qui sont des résultats acceptables lors de l'utilisation d'un modèle thermique à l'échelle du bassin versant. Bien que les valeurs simulées soient généralement sans biais (Figure 7.4), il existe une légère surestimation du nombre annuel de jours au-dessus de 20°C (Figure 7.5). Les simulations indiquent que le nombre historique le plus élevé de dépassements de 20°C est de neuf jours (2006 ; Figure 7.5b).

La Figure 7.4 (pour la station de Vanderhoof) montre que le modèle CEQUEAU peut reproduire avec précision des journées à haute température. La Figure 7.6b montre cette capacité à long terme au fil des ans par les valeurs moyennes interannuelles journalières (DOY). Par conséquent, les valeurs simulées par CEQUEAU pour les futurs scénarios de changement climatique peuvent fournir un aperçu fiable du moment et de l'ampleur des changements dans les températures élevées de l'eau.

Les données des scénarios climatiques sélectionnés ont été utilisées comme entrées du modèle CEQUEAU. De plus, le débit d'eau simulé et la température de l'eau à SLS à l'aide de CE-QUAL-W2 ont été utilisés comme conditions aux frontières.

Comme prévu, ces scénarios montrent une probabilité de dépassement plus élevée du seuil de 20°C par rapport à la période de référence. La probabilité croissante d'occurrence peut être vue en passant de SSP2-4.5 à SSP5-8.5 et en allant plus loin dans l'horizon temporel. La probabilité de dépassement commence à 3.8% pour SSP2-4.5 (horizon 2040-2069) comme valeur la plus basse en utilisant les données du modèle MPI-ESM1-2-HR, et la valeur de dépassement la plus élevée est de 36% pour SSP5-8.5 (horizon 2070-2099) à l'aide des données du modèle CMCC-ESM2.

Au fil du temps, la fréquence de dépassement augmente. Cette tendance croissante est également évidente dans les valeurs simulées d'ensemble. La Figure 7.8 montre les scénarios d'évolution de la chronologie des températures élevées pour les horizons temporels futurs. Par rapport à la Figure 7.8, on peut voir qu'août resterait le mois le plus chaud, mais le début septembre se réchaufferait plus que juillet. La Figure 7.6b indique que la période cruciale de l'année pour le réchauffement est de la mi-juillet à la mi-août, ce qui correspond à la période STMP. Cette période cruciale de réchauffement deviendrait plus longue et plus chaude à l'avenir selon ce scénario. Les futurs étés chauds devraient commencer début juillet et durer jusqu'à début septembre.

Le nombre de jours au-dessus de 20°C pendant l'été chaque année diffère selon les quatre situations. Les chiffres II-C1 à II-C4 montrent cette variabilité entre les modèles. D'autre part, les quatre combinaisons (SSP4.5 et 8.5 pour les horizons 2040-2069 et 2070-2099) montrent un nombre croissant de jours au-dessus de 20°C pendant l'été, allant de 38 à 62 jours (Figure 7.9), ce qui est bien au-dessus de l'incertitude du modèle de neuf jours pour ce critère. Le SSP8.5 pour la période 2070-2099 montre que la température moyenne quotidienne de l'eau pour la période STMP est au-dessus de 20°C. Les scénarios futurs de température de l'eau montrent systématiquement des dépassements du seuil plus élevés annuellement que le niveau maximal qui s'est produit pendant la période de référence, qui était de neuf jours.

À l'heure actuelle, aucune installation ne permet de gérer directement la température de l'eau en aval du réservoir Nechako. Il est donc nécessaire de tester d'autres règles de gestion du réservoir prenant explicitement en compte la température de l'eau en aval pour réduire le dépassement du seuil de 20°C. Des études futures pourront analyser les implications et l'efficacité de ces changements sur les conditions aux frontières en amont dans des scénarios de changement climatique.

Durant sa migration, le saumon rouge est l'une des espèces vulnérables à la hausse des températures de l'eau, avec 100% de mortalité au-delà de 21°C de température de l'eau après 72 heures et après un exercice exhaustif (Middleton *et al.*, 2018; Robinson *et al.*, 2015). La limite minimum-maximum utilisant huit modèles climatiques montre la possibilité d'une température de l'eau plus élevée que ce seuil pendant cette période. En utilisant séparément les séries chronologiques de modèles climatiques individuels, le pourcentage de jours de dépassement du seuil varie de 3.8 à 36% (Figure 7.7). Le modèle MPI-ESMI-2-HR présente la fréquence minimale et le modèle CMCC-ESM2 a la fréquence de dépassement la plus élevée.

Les trois autres combinaisons (SSP2-4.5 et SSP5-8.5 pour 2040-2069 et 2070-2099) montrent une augmentation de la température de l'eau et une augmentation de la fréquence des jours avec une température de l'eau > 20°C. Considérant le seuil létal du saumon rouge (Tableau II-A1), le nombre de jours avec une température supérieure à 20°C est un signal sévère et alarmant pour l'habitat thermique de cette espèce (Deinet *et al.*, 2020).

3.7 Évaluation des modèles d'irradiance à ondes longues descendantes avec des données d'entrée ERA5 dans les environnements canadiens

Cette étude focalise en grande partie sur la modélisation de l'irradiance de longue longueur d'onde descendante (DLI), de la collecte de données à l'identification de modèles optimaux, en discutant des différentes sources d'erreur et de leur impact sur la modélisation du DLI.

L'étude explore également les bases physiques des formulations du DLI et met en avant les performances remarquables de l'ensemble de données ERA5 dans l'estimation du DLI.

Les bases physiques des formulations du DLI ont été explorées. L'émissivité, $\beta_{Effective}$, un paramètre essentiel dans la modélisation du DLI, a été caractérisé en fonction de la température de l'air observée, du DLI et de la loi de Stefan-Boltzmann. La bonne performance de l'ensemble de données de réanalyses ERA5 a été soulignée, et sa capacité à fournir des valeurs précises de la température de l'air, de la pression de vapeur et de la couverture nuageuse a été démontrée.

Les conditions de ciel dégagé ont été analysées, révélant l'influence dominante de la température de l'air sur $\beta_{Effective}$. Les conditions de ciel dégagé dans différentes stations ont montré des schémas variables en raison des différences climatiques régionales. Les conditions de ciel nuageux, ont présenté des variations plus complexes de $\beta_{Effective}$, reflétant les effets combinés de la température, de la pression de vapeur et de la couverture nuageuse.

L'étude a examiné neuf modèles de DLI (Tableau 3.1), notamment ceux présentant des relations linéaires, exponentielles et non linéaires entre le DLI et ses variables influentes. Ces modèles ont été évalués pour leur capacité à estimer avec précision le DLI dans différentes conditions. On a constaté que les modèles plus simples, tels que ceux basés sur une relation linéaire entre β et la pression de vapeur (e_a), présentaient de bonnes performances par temps dégagé. Les modèles plus complexes introduisaient des relations non linéaires pour tenir compte des variations de e_a, de la température (T_{air}) et de la couverture nuageuse (B).

Le modèle utilisant une formulation simple d'émissivité constante pour toutes les conditions de ciel (Modèle 1), sert de référence pour la comparaison de l'efficacité entre les modèles de DLI. Pour le calage de la station SLS, le Modèle 2 présente une valeur RMSE plus élevée (37 W/m²) et un NSE plus bas (0,4) par rapport aux autres modèles, le classant de manière similaire au Modèle 1 (racine quadratique de l'erreur moyenne, ou RMSE = 37,6 W/m²).

Tableau 3.1 Formulation des neuf modèles d'estimation DLI pour cette étude

Modèle	Unité pour J	Unité pour e_a	Forme finale de l'équation ($J_{All-sky} = (1 + uB^v) \times J_{Clear-sky}$)	Référence pour $J_{Clear-sky}$	Référence pour $(1 + uB^v)$
1	W/m^2	–	$J = a\sigma T_{air}^4$	Stephen-Boltzmann law for grey body	(Duarte et al., 2006)
2*	mW/cm^2	–	$J = (1 + uB^v) \times 5.31 \times 10^{-14} T_{air}^6$	(Swinbank, 1963; Equation 2)	(Duarte et al., 2006)
3	W/m^2	–	$J = (1 + uB^v)\langle 1 - 0{,}261 \times \exp[-7{,}77 \times 10^{-4}(273 - T_{air})^2]\rangle \sigma T_{air}^4$	(Idso et al., 1969)	(Duarte *et al.*, 2006)
4	W/m^2	$mmHg$	$J = [(1 + 0{,}17B^2)0{,}97 \times (0{,}74 + 0{,}0065e_a)]\sigma T_{air}^4$	(Morin et al., 1990; Equations 11 and 12)	(Morin et al., 1990; Equation 12)
5	W/m^2	hPa	$J = (1 + uB^v)[1{,}24(\frac{e_a}{T_{air}})^{1/7}]\sigma T_{air}^4$	(Brutsaert, 1975; Equation 11)	(Duarte et al., 2006)
6	W/m^2	hPa	$J = (1 + uB^v)\langle 1{,}08[1 - exp\,(-e_a^{\frac{T_{air}}{2016}})]\rangle\, \sigma T_{air}^4$	(Satterlund, 1979; Equation 5)	(Duarte et al., 2006)
7	W/m^2	hPa	$J = (1 + uB^v)$ $(1 - \left[1 + 46{,}5\left(\frac{e_a}{T_{air}}\right)\right] exp\,\langle -\left\{1{,}2 + 3\left[46{,}5\left(\frac{e_a}{T_{air}}\right)\right]\right\}^{1/2}\rangle)\sigma {T_{air}}^4$	(Prata, 1996; Equation 20)	(Duarte et al., 2006)
8	W/m^2	hPa	$J = (1 + uB^v)\begin{cases}[0{,}72 + 0{,}009(e_a - 2)]\sigma {T_{air}}^4 & e_a \geq 2 \\ [0{,}72 - 0{,}076(e_a - 2)]\sigma {T_{air}}^4 & e_a < 2\end{cases}$	(Niemelä et al., 2001; Equation 18)	(Duarte et al., 2006)
9**	W/m^2	$mm\,Hg$	$J = \{(1 + uB^v)[exp(-\alpha \times LAI)]0{,}97 \times (0{,}74 + 0{,}0065e_a) + 0{,}98 \times \langle 1 - exp(-\alpha \times LAI)\rangle\}\sigma {T_{air}}^4$	(Pomeroy et al., 2009; Equation 3; Verseghy et al., 1993; Equation 7)	(Duarte et al., 2006)

* La constante pour la conversion de mW/cm² en W/m² est de $5{,}31 \times 10^{-13}$. De plus, en suivant la notation de la loi de Stefan-Boltzmann avec σT_{air}^4 the le terme peut être écrit comme $(9{,}3645 \times 10^{-6} T_{air}^2)\sigma T_{air}^4$.

** La transmissivité du couvert forestier $exp(-\alpha \times LAI)$ est basée sur la loi de Beer de la radiation (Verseghy *et al.*, 1993). Pomeroy *et al.* (2009) citent Oke (1987) pour $\beta_{canopy} = 0{,}98$.

où a, u, v et, α, sont des paramètres de calage sans dimension. L'autre alternative recommandée consiste à utiliser les valeurs maillées DLI de la base de données ERA5 nommées «mean surface down longwave radiation flux» ou «msdwlwrf».

Les huit autres options d'estimation du DLI à SLS présentent des valeurs de RMSE absolus et relatifs (RRMSE) variables mais somme toute assez rapprochées, allant d'un minimum de 32,0 W/m² (RRMSE = 11,8% pour ERA5) à un maximum de 35,9 W/m² (RRMSE = 13,3% pour le Modèle 5). À noter, le Modèle 4 et sa version mise à jour prenant en compte les effets de la canopée basés sur le LAI (Modèle 9) donnent des résultats similaires, suggérant que les effets de la canopée ont un impact limité sur cette station, ce qui se traduit par une performance comparable entre ces deux modèles.

Pour le calage de la station bassin versant expérimental de Sainte-Marthe (BVESM), les huit options d'estimation du DLI (ERA5, modèles 3, 4, 5, 6, 7, 8, 9) présentent des résultats de RMSE similaires, allant d'un minimum de 30,4 W/m² (RRMSE = 9,9% pour les données ERA5) à un maximum de 38,4 W/m² (RRMSE = 12,5% pour le Modèle 2). Cette similarité entre les modèles (Tableau 8.3) est attendue, compte tenu de leur fondement dans les lois de la radiation physique et de leur dépendance aux variables influençant l'émissivité de l'air. À noter, l'utilisation directe en intrant de la DLI d'ERA5 surpasse tous les modèles à la fois à SLS et BVESM. Cette supériorité de performance est attribuée à la modélisation complète de la radiation en 3D complétée afin de générer ce produit de réanalyse. La plage de RMSE observée (30,4-38,4 W/m²) correspond aux résultats de Long *et al.* (2021), qui ont rapporté un RMSE moyen de 26,4 W/m² (variant de 23,2 à 42 W/m²) pour toutes les conditions de ciel sur le plateau tibétain.

De plus, nous avons calculé des mesures d'efficacité supplémentaires pour chaque modèle, fournissant plusieurs critères de comparaison (Tableau 8.3). Le RRMSE, une mesure sans dimension semblable au RMSE, indique que toutes les options de DLI se comportent bien, avec des valeurs de RRMSE systématiquement inférieures à 15%, en accord avec les résultats précédents (Long *et al.*, 2021). Plus précisément, pour le calage à la station SLS, ERA5 affiche un RRMSE de 11,8%, tandis qu'à BVESM, il atteint 9,9%. Étant donné que les valeurs de RMSE et de RRMSE sont proches pour différentes méthodes, nous avons analysé des mesures supplémentaires (Tableau 8.3). Le NSE, qui pénalise les écarts dans les valeurs élevées au sein d'une série temporelle, révèle la performance supérieure d'ERA5 (NSE = 0,55 pour SLS et 0,76 pour BVESM), suivi par les méthodes 4 et 9 (SLS) et 6 et 9 (BVESM).

Les métriques qui tiennent compte simultanément de la complexité du modèle (plus ou moins grande parcimonie dans le nombre de paramètres) et de la performance sont l'AIC et le BIC, favorisant les modèles parcimonieux. Ces critères révèlent qu'ERA5 est la meilleure option pour les deux stations, avec les valeurs minimales. Après ERA5, le Modèle 4 se comporte le mieux pour SLS, tandis que le Modèle 9 excelle à BVESM (Tableau 8.3). Notamment, le Modèle 9, incorporant un terme supplémentaire lié au LAI par rapport au Modèle 4, présente des résultats supérieurs à BVESM, où l'impact de la canopée est plus prononcé. Pour SLS, les métriques des Modèles 4 et 9 sont presque identiques. Cependant, à la station BVESM, le Modèle 9 affiche des résultats améliorés, reflétant une plus grande influence de la canopée.

L'impact de LAI, en tant que représentation des effets de la canopée, a été pris en compte dans la modélisation du DLI. L'ensemble de données ERA5 fournissait des données de LAI, bien que les valeurs aient été dérivées de climatologies mensuelles et ne soient pas spécifiques aux sites de l'étude. Les relations entre le LAI, e_a et $\beta_{Effective}$ ont été analysées

dans des conditions de ciel dégagé et de ciel couvert, révélant des interactions complexes qui rendaient difficile l'isolement de l'influence du LAI sur le DLI.

L'étude a présenté le Modèle 9, une extension du Modèle 4, qui introduisait un facteur de correction basé sur le LAI (α) pour distinguer le DLI en zone dégagée du DLI en zone ombragée. La valeur d'α a été déterminée par calage et représente l'effet air-canopée modulé par le facteur de vue sur le ciel.

Les résultats ont indiqué que l'utilisation des données de DLI de l'ensemble de données ERA5 surpassait systématiquement tous les modèles de DLI pour les deux stations d'étude. Cela souligne la supériorité claire de l'ensemble de données ERA5 dans l'estimation du DLI et met en lumière l'importance de données climatiques précises pour la modélisation de la radiation solaire.

En résumé, cette étude offre des perspectives précieuses sur les complexités de la modélisation du DLI, en mettant en lumière l'impact de diverses variables climatiques et l'utilisation d'ensembles de données tels qu'ERA5 pour parvenir à des estimations précises. Elle souligne l'importance du choix de modèles appropriés en fonction des conditions locales et des données disponibles, mettant en avant la nécessité de progrès continus dans la modélisation de la radiation solaire.

3.8 Modélisation de l'impact de l'irradiance descendante à ondes longues sur la température de l'eau à l'échelle du bassin versant à l'aide des données d'entrée ERA5 dans l'ouest du Canada

La section précédente a permis de mettre en exergue les modèles de DLI les plus performants dans la zone d'étude, tout en validant ces choix à l'aide d'une station dans l'ouest, et d'une station dans l'est du Canada. La présente section évalue la performance de CEQUEAU lorsque ces neuf modèles sélectionnés y sont incorporés. Les neuf formulations alternatives pour le développement ultérieur de CEQUEAU se définissent comme Tableau 3.1. L'autre alternative recommandée consiste à utiliser les valeurs maillées DLI de la base de données ERA5 nommées «mean surface down longwave radiation flux» ou «msdwlwrf».

Dans cette section, nous évaluons les performances de CEQUEAU avec neuf modèles de DLI sélectionnés dans le contexte du calage multisite à l'échelle du bassin versant. Chacun de ces neuf modèles de DLI représente une formulation distincte pour l'estimation de la DLI. Après le processus de calage, les performances de ces modèles sont évaluées sur plusieurs sites à l'intérieur du bassin versant de Nechako.

Le Modèle 9, qui repose sur les bases du Modèle 4 (le modèle actuel utilisé dans CEQUEAU), se révèle être le choix le plus précis et le plus adapté pour l'estimation de la température de

l'eau à l'échelle du bassin versant. Le Modèle 9 apporte une amélioration significative par rapport à son prédécesseur en incorporant LAI en tant que variable distribuée, établissant ainsi un lien entre la couverture de la canopée et la DLI. Cette approche innovante permet une représentation plus précise de la manière dont la DLI évolue dans le bassin versant, en mettant l'accent sur la conversion du rayonnement à ondes courtes en rayonnement à ondes longues et sa contribution au bilan thermique global.

Un des facteurs clés contribuant à l'amélioration de la précision de l'estimation de la température de l'eau à l'échelle du bassin versant est la meilleure répartition du bilan thermique, réalisée grâce à des processus physiquement réalistes. Le Modèle 9 reconnaît la présence de la canopée en incorporant la variable LAI. De plus, il établit une relation mathématique entre le LAI et la variable de vue sur le ciel et introduit l'émissivité de la canopée en tant que paramètre, ce qui permet d'obtenir des estimations plus précises.

Le choix du Modèle 9 pour la modélisation de la DLI dans le cadre de CEQUEAU renforce la précision des prédictions de la température de l'eau à l'échelle du bassin versant. Cette amélioration est attribuée à une répartition plus réaliste du bilan thermique, prenant en compte l'interaction complexe de divers paramètres et processus physiques. L'intégration des données LAI renforce la capacité du modèle à fournir des estimations précises de la température de l'eau, en faisant un outil mieux adapté pour les études hydrologiques-thermiques au sein du bassin versant de la rivière Nechako.

4 CONCLUSION ET PERSPECTIVES

Le premier objectif de cette étude était de comparer trois méthodes de calage de modèles hydrothermiques pour le bassin versant de la rivière Nechako en Colombie-Britannique, au Canada. Une simulation fiable du débit et de la température de ce bassin versant est essentielle car il comprend un habitat clé pour les espèces de poissons vulnérables. Un calage du module hydrologique pour le débit a d'abord été effectué, suivi d'un autre pour la température de l'eau. Les paramètres du module thermique du modèle CEQUEAU ont d'abord été optimisés de manière mono-site puis selon la méthode multisite en regroupant les sept sites. Enfin, des simulations ont été effectuées à l'aide de paramètres mis à l'échelle (par exemple, en utilisant comme valeurs mise à l'échelle, des moyennes de paramètres) pour les comparer aux deux autres méthodes. Nos résultats ont montré que les méthodes multisites conduisent à des optima globaux pour la RMSE globale pour toutes les stations simultanément. De plus, les méthodes de mise à l'échelle des paramètres n'aient pas de valeur ajoutée par rapport à le calage multisite, mais ils ont des valeurs de RMSE globales inférieures à celles de la plupart des calages mono-sites.

Pour le bassin versant de la Nechako, bien qu'un calibrage hydrologique du modèle soit essentiel, la température de l'eau est fortement influencée par les forçages et conditions atmosphériques locaux. Ainsi, pour cette étude, les principaux termes de contribution de la chaleur contrôlant la température de l'eau sont le rayonnement à ondes courtes, le rayonnement à ondes longues et la chaleur sensible. Les stations 1 à 4 ont montré de meilleures mesures de performance pour le calage thermique, peut-être parce que le tronçon de la rivière n'est pas affecté par de grands plans d'eau ou des changements importants comme une confluence avec un affluent majeur. Cette étude montre que le modèle CEQUEAU fournit des métriques de performance globale et mono-site compétitives. Les résultats ont confirmé les études précédentes sur les avantages de l'utilisation de toutes les données disponibles (méthode multisite) pour le calage au lieu d'un équivalent spatial de l'échantillonnage fractionné. Des études futures sont recommandées pour étudier l'impact des lacs sur le régime thermique des rivières à l'aide de CEQUEAU.

Les résultats du deuxième objectif de cette étude ont fourni des simulations de température de l'eau pour la période de référence historique de 1980-2019, pour le futur proche (2040-2069) et le futur lointain (2070-2099) pour la rivière Nechako en Colombie-Britannique, Canada. L'étude a utilisé des modèles climatiques issus de l'ensemble CMIP6 et les scénarios de changement climatique SSP2-4.5 et SSP5-8.5. Le modèle CEQUEAU a été forcé avec des données ERA5 pour le passé et huit modèles climatiques CMIP6 pour le futur. La principale conclusion de l'étude était la capacité du modèle à simuler avec précision les températures

de l'eau pendant l'été sur les huit sites d'observation. Le modèle CEQUEAU a fourni des valeurs fiables de débit d'eau et de température pour l'ensemble du bassin versant de la rivière Nechako. Les résultats de l'étude fournissent des estimations d'ensemble de la température de l'eau selon les scénarios de changement climatique, qui sont nécessaires aux décideurs du bassin versant de la rivière Nechako. Les scénarios de changement climatique indiquent qu'il y aura de 3.8 à 36% plus de jours avec des températures de l'eau supérieures à 20°C, ce qui pourrait constituer une grave menace pour la population de saumon rouge en raison des changements dans le régime thermique de la rivière Nechako. Pour les mêmes scénarios de relâchées d'eau que dans le passé récent, la fréquence des jours avec une température moyenne de l'eau supérieure à 20°C devrait être de 0 à 62 jours en juillet-août, contre 0 à 9 jours dans le passé. Les résultats de la simulation de la température de l'eau ont indiqué une probabilité relativement élevée de dépassement de la limite thermique de 20°C à Vanderhoof dans la rivière Nechako. Les scénarios indiquent que des conditions potentiellement très stressantes pour les espèces d'eau froide comme le saumon rouge lors d'épisodes de haute température se produiront probablement plus fréquemment. Des études futures sont nécessaires pour évaluer l'impact possible de l'exploitation du barrage en tant que stratégie d'adaptation pour y faire face et pour mettre en œuvre d'autres solutions d'atténuation de l'augmentation de la température de l'eau.

Le troisième objectif était d'améliorer les estimations des flux thermiques d'ondes longues dans le modèle CEQUEAU. Cette étude visait à améliorer la précision des modèles DLI pour le bassin versant de la rivière Nechako en Colombie-Britannique et au Québec, Canada, où des modèles DLI précis sont essentiels pour comprendre leur impact sur le bilan thermique des écosystèmes aquatiques. Le calage de chaque modèle DLI aux stations SLS et BVESM a été suivi de leur validation à ces emplacements. Notre enquête a révélé que l'équation de Stephan-Boltzmann peut fournir des estimations raisonnables de la DLI lorsque la formulation appropriée pour estimer β est appliquée. Parmi les modèles évalués, les modèles 4, 6 et 9 ont systématiquement démontré des performances supérieures à la fois à SLS et à BVESM, ce qui en fait de solides candidats pour une intégration ultérieure dans les modèles hydrologiques du bassin versant de la Nechako. Leur supériorité a été confirmée par le biais de multiples mesures de performance.

De plus, notre analyse a indiqué que l'utilisation directe des données de DLI de la base de données de réanalyse ERA5 surpassait de manière constante les méthodes alternatives, en se basant sur diverses mesures de performance. Cette constatation souligne la recommandation d'employer les données ERA5 comme approche privilégiée pour fournir des entrées aux modèles hydrologiques et écologiques.

Les modèles de DLI proposés, en particulier les modèles 4, 6 et 9, offrent des avantages potentiels pour l'intégration dans les modèles hydrologiques tels que CEQUEAU, fournissant des outils précieux pour les décideurs au sein du bassin versant de la Nechako. Compte tenu de l'impact local substantiel des données météorologiques dans cette région et du rôle crucial du rayonnement à ondes longues dans le bilan thermique de la température de l'eau pendant l'été, l'adoption de formulations de DLI plus optimales devrait renforcer les capacités prédictives des modèles de température de l'eau. Cette amélioration des compétences prédictives sera précieuse pour les études sur le changement climatique et diverses applications. Les futures recherches devraient se concentrer sur l'application des modèles recommandés à l'échelle du bassin versant en utilisant des modèles hydrologiques déterministes, évaluant leur efficacité dans des contextes hydrologiques et écologiques réels.

Le quatrième objectif de cette étude visait à comparer dix formulations (neuf modèles et les données DLI d'ERA5) pour la DLI, ainsi que l'ensemble de données ERA5 en remplacement de ces modèles. Le bassin versant de la Nechako et le modèle hydrothermique CEQUEAU ont été utilisés. Le calage de CEQUEAU pour chaque équation a été réalisé sur quatre sites de mesure de la température de l'eau à l'aide d'une méthode de calage multisite. Nos résultats ont montré que l'équation de Stephan-Boltzmann et une formulation appropriée pour l'émissivité de l'air pouvaient fournir les meilleures estimations pour la DLI, ce qui se traduit par une amélioration des estimations de la température de l'eau. Le Modèle 9 a montré les meilleurs résultats pour le bassin versant, et cette méthode est une option proposée pour l'utilisation future du modèle CEQUEAU dans le bassin versant de la rivière Nechako. La supériorité de cette méthode est due à l'incorporation de l'Indice de Surface Foliaire (Leaf Area Index; LAI) pour la répartition de la DLI entre l'air libre et la canopée. Bien que la méthode supérieure pour estimer la DLI soit une variante de l'équation de Stephan-Boltzmann, les résultats suggèrent d'utiliser directement la variable DLI à partir du produit de réanalyse ERA5, l'un des candidats proposés. Nos conclusions indiquent que l'utilisation de la DLI à partir de l'ensemble de données ERA5, associée à la formulation correcte pour modéliser les processus terre-eau, améliore les performances du modèle à l'échelle du bassin versant. De plus, l'utilisation des données de réanalyse ERA5 fournit des données sur une période prolongée pour la prise de décision à long terme, ce qui intéresse les parties prenantes du bassin versant de la rivière Nechako. Cette amélioration des performances confirme les nombreuses études précédentes recommandant les données ERA5 comme une source fiable pour les régions disposant de peu de données.

Perspectives futures

En conclusion, cette étude a présenté des informations précieuses sur l'impact du changement climatique et de l'exploitation des barrages sur la modélisation de la température de l'eau. De plus, le modèle CEQUEAU a été développé pour la modélisation du rayonnement à ondes longues. Cependant, comme pour toute recherche, il reste encore des aspects qui nécessitent une investigation approfondie. Les travaux futurs devraient se concentrer sur l'impact des grands plans d'eau sur l'écoulement et la température de l'eau dans le bassin versant de la rivière Nechako, ainsi que sur l'exploration de l'utilisation de la spatialisation des paramètres du modèle CEQUEAU pour améliorer la précision du modèle. De plus, une analyse de sensibilité thermique du modèle fournirait des informations importantes sur les paramètres du modèle à spatialiser. Aborder ces lacunes dans la modélisation de la température de l'eau en utilisant le modèle CEQUEAU contribuera à une compréhension plus approfondie de la relation complexe entre les plans d'eau et la température de l'eau, ce qui permettra en fin de compte de mettre en place des stratégies de gestion des ressources en eau plus efficaces.

5 RÉFÉRENCES BIBLIOGRAPHIQUES

Ahmad SK, Hossain F, Holtgrieve GW, Pavelsky T, Galelli S (2021). Predicting the likely thermal impact of current and future dams around the world. *Earth's Future* 9(10):e2020EF001916.

Akaike H (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6):716-723.

Algera DA, Kamal R, Ward TD, Pleizier NK, Brauner CJ, Crossman JA, Leake A, Zhu DZ, Power M, Cooke SJ (2022). Exposure risk of fish downstream of a hydropower facility to supersaturated total dissolved gas. *Water Resources Research* 58(6):e2021WR031887.

Arsenault R, Brissette F, Chen J, Guo Q, Dallaire G (2020). NAC2H: The North American Climate Change and Hydroclimatology Data Set. *Water Resources Research* 56(8):e2020WR027097.

Arsenault R, Brissette F, Martel J-L (2018). The hazards of split-sample validation in hydrological model calibration. *Journal of hydrology* 566:346-362.

Arsenault R, Poulin A, Côté P, Brissette F (2014). Comparison of stochastic optimization algorithms in hydrological model calibration. *Journal of Hydrologic Engineering* 19(7):1374-1384.

Bâ KM, Diaz V, Gómez-Albores MA, Díaz-Delgado C, Nájera-Mota N, Seidou O, Ortiz F (2018). Spatially distributed hydrological modelling of a Western Africa basin. *EPiC Ser. Eng* 3:343-350.

Bárbaro EW, Oliveira AP, Soares J, Ferreira MJ, Božnar MZ, Mlakar P, Escobedo JF (2009). Downward Atmospheric longwave radiation in the city of São Paulo. *AIP Conference Proceedings.* AIP, p 447-450.

Benyahya L, Caissie D, St-Hilaire A, Ouarda TB, Bobée B (2007). A review of statistical water temperature models. *Canadian Water Resources Journal* 32(3):179-192.

Bernhardt ES, Savoy P, Vlah MJ, Appling AP, Koenig LE, Hall RO, Arroita M, Blaszczak JR, Carter AM, Cohen M, Harvey JW, Heffernan JB, Helton AM, Hosen JD, Kirk L, McDowell WH, Stanley EH, Yackulic CB, Grimm NB (2022). Light and flow regimes regulate the metabolism of rivers. *Proceedings of the National Academy of Sciences* 119(8):e2121976119.

Bérubé S, Brissette F, Arsenault R (2022). Optimal Hydrological Model Calibration Strategy for Climate Change Impact Studies. *Journal of Hydrologic Engineering* 27(3):04021053.

Bosmans J, Wanders N, Bierkens MFP, Huijbregts MAJ, Schipper AM, Barbarossa V (2022). FutureStreams, a global dataset of future streamflow and water temperature. *Scientific Data* 9(1):307.

Brisson C, Boucher MA, Latraverse M (2015). Illustration of the added value of using a multi-site calibration and correction approach to reconstruct natural inflows and inter-catchment transfer flow: A case study. *Canadian Journal of Civil Engineering* 42(5):342-352.

Brown J, Rosolem R, Woods R, Rocha H, Roberti D (2021). Evaluation of high-resolution meteorological global data products using flux tower observations across Brazil.), p EGU21-15387.

Brunt D (1932). Notes on radiation in the atmosphere. I. *Quarterly Journal of the Royal Meteorological Society* 58(247):389-420.

Cannon AJ (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics* 50(1):31-49.

Craig JR, Brown G, Chlumsky R, Jenkinson RW, Jost G, Lee K, Mai J, Serrer M, Sgro N, Shafii M, Snowdon AP, Tolson BA (2020). Flexible watershed simulation with the Raven hydrological modelling framework. *Environmental Modelling & Software* 129:104728.

Cucchi M, Weedon GP, Amici A, Bellouin N, Lange S, Müller Schmied H, Hersbach H, Buontempo C (2020). WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth Syst. Sci. Data* 12(3):2097-2120.

Dalcin AP, Brêda JPLF, Marques GF, Tilmant A, de Paiva RCD, Kubota PY (2023). The Role of Reservoir Reoperation to Mitigate Climate Change Impacts on Hydropower and Environmental Water Demands. *Journal of Water Resources Planning and Management* 149(4):04023005.

Dallaire G, Poulin A, Arsenault R, Brissette F (2021). Uncertainty of potential evapotranspiration modelling in climate change impact studies on low flows in North America. *Hydrological Sciences Journal* 66(4):689-702.

Deinet S, Scott-Gatty K, Rotton H, Twardek WM, Marconi V, McRae L, Baumgartner LJ, Brink K, Claussen JE, Cooke SJ (2020). *The living planet index (LPI) for migratory freshwater fish: Technical report.* 30 p

Duarte HF, Dias NL, Maggiotto SR (2006). Assessing daytime downward longwave radiation estimates for clear and cloudy skies in Southern Brazil. *Agricultural and Forest Meteorology* 139(3):171-181.

Dugdale SJ, Hannah DM, Malcolm IA (2017a). River temperature modelling: A review of process-based approaches and future directions. *Earth-Science Reviews* 175:97-113.

Dugdale SJ, St-Hilaire A, Allen Curry R (2017b). Automating drainage direction and physiographic inputs to the CEQUEAU hydrological model: sensitivity testing on the lower Saint John River watershed, Canada. *Journal of Hydroinformatics* 19(3):469-492.

Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D (2007). The shuttle radar topography mission. *Reviews of Geophysics* 45(2):33.

Ficklin DL, Luo Y, Stewart IT, Maurer EP (2012). Development and application of a hydroclimatological stream temperature model within the Soil and Water Assessment Tool. *Water Resources Research* 48(1):16.

Fortin J-P, Turcotte R, Massicotte S, Moussa R, Fitzback J, Villeneuve J-P (2001). Distributed Watershed Model Compatible with Remote Sensing and GIS Data. I: Description of Model. *Journal of Hydrologic Engineering* 6(2):91-99.

Fullerton AH, Sun N, Baerwalde MJ, Hawkins BL, Yan H (2022). Mechanistic simulations suggest riparian restoration can partly counteract climate impacts to juvenile salmon. *JAWRA Journal of the American Water Resources Association* 58(4):525-546.

Gatien P, Arsenault R, Martel J-L, St-Hilaire A (2022). Using the ERA5 and ERA5-Land reanalysis datasets for river water temperature modelling in a data-scarce region. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 10.1080/07011784.2022.2113917:1-18.

Gauch M, Kratzert F, Gilon O, Gupta H, Mai J, Nearing G, Tolson B, Hochreiter S, Klotz D (2023). In Defense of Metrics: Metrics Sufficiently Encode Typical Human Preferences

Regarding Hydrological Model Performance. *Water Resources Research* 59(6):e2022WR033918.

Grooten M, Almond RE (2018). *Living planet report-2018: aiming higher.* WWF international, (978-2-940529-90-2) p

Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of hydrology* 377(1-2):80-91.

Hansen N (2023). The CMA evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772* https://doi.org/10.48550/arXiv.1604.00772.

Hansen N, Ostermeier A (1996). Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. *Proceedings of IEEE international conference on evolutionary computation.* (Nagoya, Japan, IEEE, p 312-317.

Herrero J, Polo M (2012). Parameterization of atmospheric longwave emissivity in a mountainous site for all sky conditions. *Hydrology and Earth System Sciences* 16(9):3139.

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J-N (2020a). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146(730):1999-2049.

Huang Y, Ramaswamy V, Soden B (2007). An investigation of the sensitivity of the clear-sky outgoing longwave radiation to atmospheric temperature and water vapor. *Journal of Geophysical Research: Atmospheres* 112(D5).

Islam SU, Hay RW, Déry SJ, Booth BP (2019). Modelling the impacts of climate change on riverine thermal regimes in western Canada's largest Pacific watershed. *Scientific Reports* 9(1):11398.

Jiang H, Yang Y, Bai Y, Wang H (2020). Evaluation of the Total, Direct, and Diffuse Solar Radiations From the ERA5 Reanalysis Data in China. *IEEE Geoscience and Remote Sensing Letters* 17(1):47-51.

Karakoyun Y, Yumurtacı Z, Dönmez AH (2018). Chapter 4.9 - Environmental fow assessment methods: A case study. *Exergetic, Energetic and Environmental Dimensions*, Dincer I, Colpan CO,Kizilkan O (Édit.) Academic Press, https://doi.org/10.1016/B978-0-12-813734-5.00060-3. p 1061-1074.

Karra K, Kontgis C, Statman-Weil Z, Mazzariello JC, Mathis M, Brumby SP (2021). Global land use/land cover with Sentinel 2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS.* (Brussels, Belgium, IEEE, p 4704-4707.

Khorsandi M, St-Hilaire A, Arsenault R (2022). Multisite calibration of a semi-distributed hydrologic and thermal model in a large Canadian watershed. *Hydrological Sciences Journal* 67(14):2147-2174.

Koll DD, Cronin TW (2018). Earth's outgoing longwave radiation linear due to H2O greenhouse effect. *Proceedings of the National Academy of Sciences* 115(41):10293-10298.

Krause P, Boyle D, Bäse F (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in geosciences* 5:89-97.

Legates DR, McCabe Jr GJ (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research* 35(1):233-241.

Lewis SL, Maslin MA (2015). Defining the Anthropocene. *Nature* 519(7542):171-180.

Li M, Jiang Y, Coimbra CFM (2017). On the determination of atmospheric longwave irradiance under all-sky conditions. *Solar Energy* 144:40-48.

Long Z, Ding L, Zhou J, Zhou T (2021). Toward the Estimation of All-Weather Daytime Downward Longwave Radiation over the Tibetan Plateau. *Atmosphere* 12(12):1692.

Macdonald JS (2019). *Saik'uz First Nation and Stellat'en First Nation v. Rio Tinto Alcan, et al. BCSC No.116524, Vancouver Registry, Sockeye Salmon and Water Temperature.* Department of Justice Canada, Canada. 1-48 p

McFarlane V, Clark SP (2021). A detailed energy budget analysis of river supercooling and the importance of accurately quantifying net radiation to predict ice formation. *Hydrological Processes* 35(3):e14056.

Middleton CT, Hinch SG, Martins EG, Braun DC, Patterson DA, Burnett NJ, Minke-Martin V, Casselman MT (2018). Effects of natal water concentration and temperature on the behaviour of up-river migrating sockeye salmon. *Canadian Journal of Fisheries and Aquatic Sciences* 75(12):2375-2389.

Monteith J, Unsworth M (2013). *Principles of environmental physics: plants, animals, and the atmosphere.* Academic Press, USA. 403 p

Morin G, Couillard D (1990). Predicting river temperatures with a hydrological model. *Encyclopedia of fluid mechanic, surface and groundwater flow phenomena,* Volk Gulf Publishing Company, Houston, Texas Vol 10. p 171-209.

Morin G, Paquet P (2007). *Modèle hydrologique CEQUEAU.* INRS, Centre Eau, Terre et Environnement, Quebec, Canada. 458 p

Murray FW (1967). On the computation of saturation vapor pressure. *Journal of Applied Meteorology and Climatology* 6(1):203-204.

Nash JE, Sutcliffe JV (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of hydrology* 10(3):282-290.

Nash KL, Cvitanovic C, Fulton EA, Halpern BS, Milner-Gulland EJ, Watson RA, Blanchard JL (2017). Planetary boundaries for a blue planet. *Nature Ecology & Evolution* 1(11):1625-1634.

Niemelä S, Räisänen P, Savijärvi H (2001). Comparison of surface radiative flux parameterizations: Part I: Longwave radiation. *Atmospheric research* 58(1):1-18.

Noël P, Rousseau AN, Paniconi C, Nadeau DF (2014). Algorithm for Delineating and Extracting Hillslopes and Hillslope Width Functions from Gridded Elevation Data. *Journal of Hydrologic Engineering* 19(2):366-374.

Norton GE, Bradford A (2009). Comparison of two stream temperature models and evaluation of potential management alternatives for the Speed River, Southern Ontario. *Journal of Environmental Management* 90(2):866-878.

Oke TR (1987). *Boundary layer climates.* Routledge,

Ouellet-Proulx S (2018). *Prévision thermique d’ensemble en rivière avec assimilation de données*. Doctoral (Université du Québec, Institut national de la recherche scientifique, Quebec, Canada). 222 p. https://espace.inrs.ca/7621/1/T750.pdf

Ouellet-Proulx S, Chiadjeu OC, Boucher M-A, St-Hilaire A (2017a). Assimilation of water temperature and discharge data for ensemble water temperature forecasting. *Journal of hydrology* 554:342-359.

Ouellet-Proulx S, St-Hilaire A, Boucher M-A (2017b). Water temperature ensemble forecasts: Implementation using the CEQUEAU model on two contrasted river systems. *Water* 9(7):457.

Ouellet-Proulx S, St-Hilaire A, Boucher MA (2019). Implication of evaporative loss estimation methods in discharge and water temperature modelling in cool temperate climates. *Hydrological Processes* 33(22):2867-2884.

Ouellet V, Secretan Y, St-Hilaire A, Morin J (2014). Water temperature modelling in a controlled environment: comparative study of heat budget equations. *Hydrological Processes* 28(2):279-292.

Ouellet V, St-Hilaire A, Dugdale SJ, Hannah DM, Krause S, Proulx-Ouellet S (2020). River temperature research and practice: Recent challenges and emerging opportunities for managing thermal habitat conditions in stream ecosystems. *Science of The Total Environment* 736:139679.

Pomeroy JW, Marks D, Link T, Ellis C, Hardy J, Rowlands A, Granger R (2009). The impact of coniferous forest temperature on incoming longwave radiation to melting snow. *Hydrological Processes* 23(17):2513-2525.

Rajesh M, Rehana S (2022). Impact of climate change on river water temperature and dissolved oxygen: Indian riverine thermal regimes. *Scientific Reports* 12(1):9222.

Robinson KA, Hinch SG, Raby GD, Donaldson MR, Robichaud D, Patterson DA, Cooke SJ (2015). Influence of postcapture ventilation assistance on migration success of adult sockeye salmon following capture and release. *Transactions of the American Fisheries Society* 144(4):693-704.

Schwarz G (1978). Estimating the dimension of a model. *The Annals of Statistics* 6(2):461-464.

Sheedy B (2005). *Analysis of a cold water release facility in the Nechako Reservoir*. Master of Business Administration (Faculty of Business Administration-Simon Fraser University). 90 p. https://summit.sfu.ca/_flysystem/fedora/sfu_migrate/10292/etd2058.pdf

Shen H, Tolson BA, Mai J (2022). Time to update the split-sample approach in hydrological model calibration. *Water Resources Research* 58(3):e2021WR031523.

Sicart JE, Hock R, Ribstein P, Chazarin JP (2010). Sky longwave radiation on tropical Andean glaciers: parameterization and sensitivity to atmospheric variables. *Journal of glaciology* 56(199):854-860.

Simpson NP, Mach KJ, Constable A, Hess J, Hogarth R, Howden M, Lawrence J, Lempert RJ, Muccione V, Mackey B, New MG, O'Neill B, Otto F, Pörtner H-O, Reisinger A, Roberts D, Schmidt DN, Seneviratne S, Strongin S, van Aalst M, Totin E, Trisos CH (2021). A framework for complex climate change risk assessment. *One Earth* 4(4):489-501.

St-Hilaire A, Boucher M-A, Chebana F, Ouellet-Proulx S, Zhou QX, Larabi S, Dugdale S, Latraverse M (2015). Breathing a new life to an older model: The CEQUEAU tool for flow and water temperature simulations and forecasting. *Proceedings of the 22nd Canadian Hydrotechnical Conference.* (Montréal, Canada, April 29–May 2), p 10.

Stratton Garvin LE, Rounds SA, Buccola NL (2022). *Updates to models of streamflow and water temperature for 2011, 2015, and 2016 in rivers of the Willamette River Basin, Oregon.* U.S. Department of the Interior, U.S. Geological Survey, USA, (2022-1017). 88 p

Sullivan AB, Rounds SA (2021). *Modeling water temperature response to dam operations and water management in Green Peter and Foster Lakes and the South Santiam River, Oregon.* U.S. Department of the Interior, U.S. Geological Survey, USA, (2020-5145). 37 p

Taghizadeh E, Ahmadi-Givi F, Brocca L, Sharifi E (2021). Evaluation of satellite/reanalysis precipitation products over Iran. *International Journal of Remote Sensing* 42(9):3474-3497.

Tang W, Qin J, Yang K, Zhu F, Zhou X (2021). Does ERA5 outperform satellite products in estimating atmospheric downward longwave radiation at the surface? *Atmospheric research* 252:105453.

Tarek M, Brissette F, Arsenault R (2020a). Uncertainty of precipitation reference dataset for climate change impact studies. *EGU General Assembly.*

Tarek M, Brissette FP, Arsenault R (2020b). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.* 24(5):2527-2544.

Verseghy DL, McFarlane NA, Lazare M (1993). Class—A Canadian land surface scheme for GCMS, II. Vegetation model and coupled runs. *International Journal of Climatology* 13(4):347-370.

Voortman B, Bartholomeus R, Van der Zee S, Bierkens MF, Witte JPM (2015). Quantifying energy and water fluxes in dry dune ecosystems of the Netherlands. *Hydrology and Earth System Sciences* 19(9):3787-3805.

Xiong YJ, Yin J, Paw U KT, Zhao SH, Qiu GY, Liu Z (2020). How the three Gorges Dam affects the hydrological cycle in the mid-lower Yangtze River: a perspective based on decadal water temperature changes. *Environmental Research Letters* 15(1):014002.

Yao B, Liu C, Yin Y, Liu Z, Shi C, Iwabuchi H, Weng F (2020). Evaluation of cloud properties from reanalyses over East Asia with a radiance-based approach. *Atmos. Meas. Tech.* 13(3):1033-1049.

Zanaga D, Van De Kerchove R, De Keersmaecker W, Souverijns N, Brockmann C, Quast R, Wevers J, Grosu A, Paccini A, Vergnaud S, Cartus O, Santoro M, Fritz S, Georgieva I, Lesiv M, Carter S, Herold M, Li L, Tsendbazar N-E, Ramoino F, Arino O (2021). ESA WorldCover 10 m 2020 v100.).

Zhou Z, Lin A, Wang L, Qin W, Zhong Y, He L (2020). Trends in downward surface shortwave radiation from multi-source data over China during 1984–2015. *International Journal of Climatology* 40(7):3467-3485.

Zhu S, Piotrowski AP (2020). River/stream water temperature forecasting using artificial intelligence models: a systematic review. *Acta Geophysica* 68(5):1433-1442.

6 ARTICLE I

Multisite calibration of a semi-distributed hydrologic and thermal model in a large Canadian watershed

Calage multisite d'un modèle hydrologique et thermique semi-distribué dans un grand bassin versant canadien

Auteurs :

Mostafa Khorsandi[1], André St-Hilaire[1], Richard Arsenault[2]

[1]Canadian Rivers Institute and Institut National de la Recherche Scientifique, Centre Eau Terre Environnement (INRS-ÉTÉ), 490 De la Couronne St., Québec City, Quebec, G1K 9A9, Canada

[2]Hydrology, climate and climate change laboratory, École de technologie supérieure (ÉTS), Montréal, Quebec, Canada

L'article a été publié dans la revue Hydrological Sciences Journal

Date de publication : Novembre 2022

Un erratum soumis pour cet article en octobre 2023

Contribution des auteurs :

Mostafa Khorsandi: Curation des données, méthodologie, conceptualisation, logiciel, analyse formelle, validation, visualisation, rédaction – ébauche originale, rédaction – révision et édition.

André St-Hilaire: Curation des données, méthodologie, conceptualisation, acquisition de financement, administration du projet, ressources, supervision, validation, rédaction - révision et édition.

Richard Arsenault: Acquisition de financement, ressources, supervision, validation, rédaction – révision & édition.

Résumé

Cette étude compare les méthodes de calage monosite et multisite avec des paramètres mis à l'échelle pour la modélisation thermique. Le calage multisite et la mise à l'échelle des paramètres sont nouvelles pour calibrer un modèle hydrologique-thermique. Pour cela, le calage du modèle et les simulations sont réalisés à l'aide d'un modèle hydrologique-thermique semi-distribué (CEQUEAU). Dans un premier temps, un calage hydrologique est effectué. Les mêmes sorties du modèle hydrologique étalonné sont utilisées pour tous les calages thermiques. L'algorithme d'optimisation à objectif unique de la stratégie d'évolution de l'adaptation de la matrice de covariance (CMA-ES) est appliqué pour caler le module thermique à l'aide des températures de l'eau observées sur sept sites de surveillance dans le bassin versant de Nechako en Colombie-Britannique, au Canada. Le calage sur un seul site a fourni les meilleures performances pour les sites individuels, mais ces ensembles de paramètres ont donné de mauvais résultats lorsqu'ils ont été appliqués aux autres sites. Le calage multisite et les méthodes de mise à l'échelle fonctionnent de manière adéquate pour tous les sites avec un RMSE global < 2°C comme seuil acceptable. Cependant, la méthode multisite a donné les meilleures mesures de performance.

Mots-clés

CEQUEAU, Multisite, Modèle hydrologique, Modèle de température de l'eau, Calage, Température de la rivière.

Abstract

This study compares single-site and multisite calibration methods with upscaled parameters for water temperature modeling. Multisite calibration and parameter upscaling are novel for calibrating a hydrological-thermal model. To this end, model calibration and simulations are carried out using a semi-distributed hydrological-thermal model (CEQUEAU). First, a hydrological calibration is carried out. The same calibrated hydrological model outputs are used for all thermal calibrations. The Covariance matrix adaptation evolution strategy (CMA-ES) single-objective optimization algorithm is applied to calibrate the thermal module using observed water temperatures at seven monitoring sites within the Nechako Watershed in British Columbia, Canada. Single-site calibration provides the best performance for individual sites, but these sets of parameters perform poorly when applied to the other sites. The multisite calibration and the upscaled methods perform adequately for all sites with overall RMSE < 2°C as an acceptable threshold. However, the multisite method results in the best performance metrics.

Key Words

CEQUEAU, Multisite, Hydrologic model, Water temperature model, Calibration, River temperature.

6.1 Introduction

Forecasting or simulating river temperature is increasingly recognized as a key element of water resources management (Ouellet *et al.*, 2020). Each of the main steps in producing such simulations includes some level of uncertainty, namely the model structure, algorithms, data input measurement error, and model calibration.

Many hydrologic and thermal models are available, including statistical and more physically-based semi- or fully-distributed models (Dugdale *et al.*, 2017a; Ouellet *et al.*, 2020). Due to the lack of available observed data in the past, hydrological models were commonly calibrated using observations at a single station (Duan *et al.*, 2003). However, new data acquisition tools provide opportunities to increase the number of locations on a river where observations become available. The knowledge gap in using spatially distributed data to calibrate and validate such models was addressed by several studies (Franco *et al.*, 2020; Odusanya *et al.*, 2019; Pandey *et al.*, 2020a; Pandey *et al.*, 2020b).

The calibration aims to match the simulated temperatures with the observations by finding the optimal (or near-optimal) parameter values. Cao *et al.* (2006) developed multisite calibration/validation using the SWAT model for hydrological variables: potential

evapotranspiration (PET), annual water yield, daily streamflow, baseflow, and soil moisture. Their main focus was on mountainous regions and the separation of sub-basin contribution to overall basin runoff. Ricard *et al.* (2013) developed multisite calibration/validation using the HYDROTEL model for streamflow in the St. Lawrence River watershed. Their results show some limitations in high-flow estimation. However, their results provide an acceptable trade-off between local performance and regional consistency for parameter estimations. Hasan and Pradhanang (2017) focused on the flow regime in the Himalayan Mountains in western Nepal. They used a multisite calibration scheme to improve snow-driven flow simulations successfully. Shrestha *et al.* (2016) used multisite calibration for flow and nutrient load simulations in semi-arid areas. They used monthly sampling to calibrate the SWAT model using streamflow, total suspended sediment (TSS), total nitrogen (TN), and total phosphorus (TP) for five monitoring stations within the Onkaparinga watershed in South Australia. Their results showed that using a multisite method for flow and sediment simulations provided no improvement compared to single-site calibration, but they did show improvements in TN and TP simulations.

In another study, researchers used a multisite calibration method to study the impact of extreme flows in an urbanized watershed (Dong et al., 2019). They analyzed water movement in the Hamilton Harbour watershed (Canada) by focusing on season-specific processes. Multisite calibration and its ability to simulate reality more efficiently than single-site calibration helped them analyze water flows in pastures, croplands, and forests (Dong *et al.*, 2019). Leta *et al.* (2017) found that using multisite calibration helps to better account for the spatial variability of hydrological processes than single-site calibration. To this end, Leta et al. (2017) identified sensitive parameters using the Latin hypercube one-factor-at-a-time (LH-OAT) technique. Then, SWAT was calibrated by using: (1) single-site calibration (SSC) at the watershed outlet; (2) step-by-step calibration starting at the upstream station and subsequently moving downstream (sequential calibration or SC); and (3) simultaneous multisite calibration (SMSC), where flow data of two stations were simultaneously used in a single calibration. Their results suggested that the SC and the SMSC techniques provided excellent results and improved the model performance. This improvement was observed in the more realistic form of flow hydrographs and an increase in the Nash-Sutcliffe Efficiency metric (NSE) from 0.5 to 0.8. Odusanya *et al.* (2019) performed multisite calibration using satellite-based evapotranspiration in a data-sparse watershed in southwestern Nigeria using a monthly time step. This new calibration method estimated actual evapotranspiration (AET) more precisely, and the calibrated model captured seasonal variability in several water budget components.

The biggest obstacle to multisite calibration is the paucity of data. Forcing models with reanalysis climate data as model inputs, or remote-sensed observations as required observed variables, leads to increasingly more opportunities to study data-sparse regions. With reliable global-scale reanalysis data like ERA5 and remotely sensed observations, there are several opportunities for using them for thermal modelling at the watershed scale. Tarek et al. (2019) used the ERA5 reanalysis as a reference data set for hydrological modelling over North America. Gatien *et al.* (2022) evaluated ERA5 as the reference input data for river thermal modelling using HEC-RAS.

Most studies and implementations have traditionally used single-site observations to perform model calibration for water temperature modelling because, in many instances, only one monitoring station exists in the drainage basin (Duan *et al.*, 2003; Ouellet-Proulx *et al.*, 2019). However, as computationally efficient methods and more diverse data sources have become more available, studies have increasingly implemented multisite calibration for streamflow simulations (Bai *et al.*, 2017; Hasan & Pradhanang, 2017; Nkiaka *et al.*, 2018; Shrestha *et al.*, 2016; Xue *et al.*, 2016). However, few specialized models for water temperature simulation have been calibrated using multisite calibration methods.

Although many hydrological models exist, few have a water temperature module (Ficklin *et al.*, 2012). Typically, these models first simulate streamflow to obtain water volumes and then calculate heat fluxes to determine the water temperature. Ficklin *et al.* (2012) reviewed several models for water temperature. Some of these models are mechanistic and are generally based on calculating a water heat budget. These include BasinTEMP (Allen *et al.*, 2007), SHADE-HSPF (Becknell *et al.*, 1993), HEAT SOURCE (Boyd & Kasper, 2003), QUAL2K (Chapra *et al.*, 2008), CE-QUAL-W2 (Cole & Wells, 2006), CrUSTe (LeBlanc *et al.*, 1997), CEQUEAU (Morin & Paquet, 2007), Streamline (Rutherford *et al.*, 1997), and SNTEMP (Theurer *et al.*, 1984), to name a few.

Although several studies focus on multisite hydrological calibration, to our knowledge, only one study has investigated the concept of multisite thermal calibration. The only such study is that of Barnhart *et al.* (2014). They added a stream temperature module to the SWAT model and used multisite calibration methods for water temperature for the first time. They showed that this method performs better for water temperature simulations basin-wide than the original model calibration. Multisite calibration is relatively unexplored for other water temperature models, leaving a gap in our understanding of thermal model capabilities, limitations, and uncertainties (Ricard *et al.*, 2013). The present study focuses on the CEQUEAU model, designed explicitly for coupled hydrological-thermal modeling (Morin & Couillard, 1990; St-Hilaire *et al.*, 2015). For hydrological-thermal models such as CEQUEAU, a thermal calibration often involves a two-step approach, starting with the hydrological module calibration and then

the thermal module calibration (Ouellet-Proulx *et al.*, 2019). Brisson *et al.* (2015) showed the benefits of the CEQUEAU model multisite calibration for the spatial reconstruction of water flows over two watersheds in Quebec, Canada. However, they did not implement the thermal module in their study.

Some of the thermal models, like CEQUEAU, use global parameters for the whole watershed. The motivation for this simplification is to prevent over-parameterization and to keep the model parsimonious and computationally efficient. The current version of CEQUEAU uses global parameters for the thermal module, while it has the ability to spatialize some of the parameters (16 out of 26) in the hydrological module. However, relating the multisite observations and calibration to the global basin-scale parameters for the thermal module requires specific attention. This problem is often referred to upscaling in hydrology (Blöschl, 2001; Blöschl & Sivapalan, 1995). Upscaling for parsimonious models with global parameters means estimating model parameters on coarser spatial scales using parameter sets estimated at finer scales (Blöschl & Sivapalan, 1995). Upscaling was used in several studies for multiple state variables in hydrology, for example, latent heat flux (Van Niel *et al.*, 2012), soil moisture (Qin *et al.*, 2013), and temporal upscaling of evapotranspiration (Xu *et al.*, 2015). Also, several arithmetic operators were introduced for upscaling water flow/state in hydrology (Samaniego *et al.*, 2010). However, using upscaling methods to estimate parameters for a thermal model in hydrology is novel.

Our specific objective is to use reanalysis data in model input and then study the performance of upscaling model parameters for simulating water temperature using the CEQUEAU model. Therefore, this study aims to quantify the efficiency of three methods in the thermal calibration of CEQUEAU. The first approach is calibrating and validating the model thermally for each site. The second method calibrates/validates the model in a multisite manner, performed by pooling all the data from several stations together. In the third method, the upscaled sets of parameters were calculated by averaging the parameter values obtained from single-site calibrations using different upscaling operators (i.e., different types of means). Then, the thermal module performances using upscaled parameters were compared against the single-site and multisite calibrations.

6.2 Methods

6.2.1 Study area

The Nechako River sub watershed (downstream of Skins Lake Spillway to Finmoore) has a 12,805-km^2 drainage area. It is located in the central part of British Columbia, Canada (Figure 6.1).

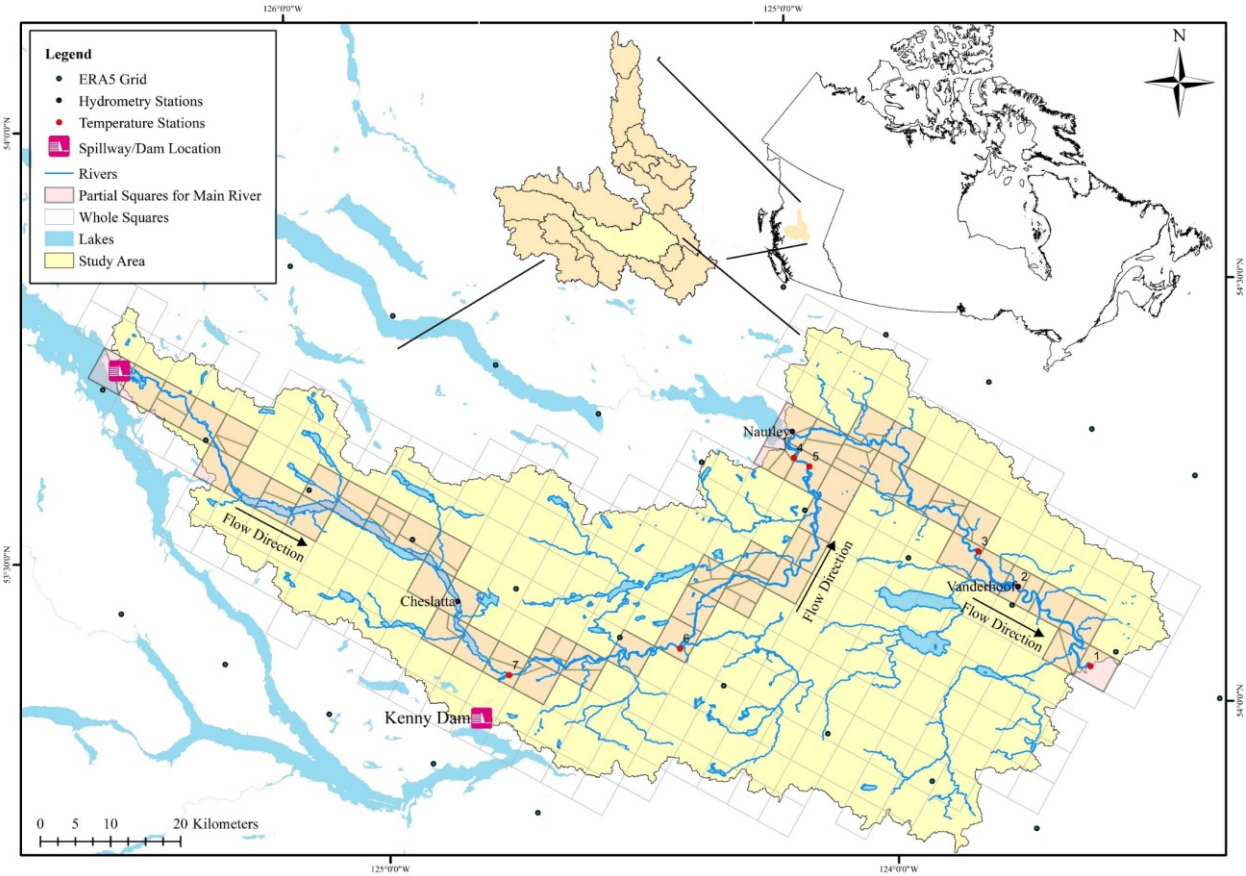

Figure 6.1 Map of the Nechako watershed

This watershed's daily temperature varies from -44°C in winter to 35°C in summer. Also, the elevation in this watershed varies from 628 to 2240 m above sea level. The mean precipitation over the watershed is around 690 mm/year using all available stations for the last 40 years, varying from 460 mm/year to 2177 mm/year for available stations. Flow in the river is fully regulated by the Skins Lake Spillway (SLS) between the Nechako Reservoir and the confluence with the Nautley River, after which the flows combine and continue past the town of Vanderhoof and eventually flow to the Fraser River. Three hydrometric stations measure the discharge along this path (data sources: Rio Tinto, Environment and Climate Change Canada, data access: March 2022). Those are at the Skins Lake Spillway, in the Nautley River (just upstream of the confluence with the Nechako river), and at Vanderhoof (see Figure 6.1 and Table 6.1).

Table 6.1 Water flow station locations and data availability period for each station

Station	Longitude (West)	Latitude (North)	Data availability	
			Start	End
Skins Lake Spillway	-125.986373	53.781751	01 Jan 2016	31 Dec 2019
Nautley	-124.59792	54.085499	01 Jan 2016	31 Dec 2019
Vanderhoof	-124.005287	54.024968	01 Jan 2016	31 Dec 2019

The Skins Lake Spillway and the Nautley flows and water temperatures are used to provide the boundary conditions to the hydrological and thermal modules and are, therefore, not used as target stations for hydrological calibration. Seven stations measured the water temperature at daily steps, labelled from 7 to 1 from upstream to downstream. Rio Tinto operates the spillway for environmental flows and flood mitigation purposes. During the Sockeye salmon (*Onchorhynchus Nerka*) migration season, which lasts from mid-July to mid-August, Rio Tinto releases water volumes to maintain mean daily water temperatures below 20°C at Finmoore, a few kilometres upstream of the confluence of the Nechako and Stuart rivers. The Summer Temperature Management Program (STMP) water release protocol was developed between 1980 and 1983 (Macdonald, 2019). The present paper focuses on the Nechako River's reach between the Skins Lake Spillway and Station 1 (about 16 km upstream of the confluence with the Stuart River; Figure 6.1). The drainage area at Station 1 is approximately 5,433 km^2. It is important to recall that flows are mainly controlled by releases at the Skins Lake spillway for much of the river length, up to the target location of Vanderhoof.

The required meteorological data for the CEQUEAU model simulation are precipitation, maximum and minimum air temperatures, wind speed, water vapour pressure, and net solar shortwave radiation. They are provided by the European Center for Medium-Range Weather Forecasting (ECMWF) through their European Reanalysis 5th generation (ERA5) (Hersbach *et al.*, 2020b). Since vapor pressure is not one of the ERA5 variables for near-surface, it is calculated using Tetan's equation and dew point temperature at 2 m height (Ouellet-Proulx, 2018), which is available on ERA5 database (Ouellet-Proulx, 2018). ERA5 gridded data were interpolated on all basin grid points using the built-in CEQUEAU interpolator, based on the nearest neighbours approach (Ouellet-Proulx *et al.*, 2019).

Landcover and topography are other required data to create the CEQUEAU model input structure. The global landcover provided by European Space Agency (ESA) with a 10 m resolution is used (Karra *et al.*, 2021; Zanaga *et al.*, 2021). Furthermore, the NASA SRTM

Digital Elevation Model (DEM) with 30 m resolution is used to calculate elevations for the input structures (Farr *et al.*, 2007).

6.2.2 CEQUEAU model

CEQUEAU (Morin & Couillard, 1990) is a conceptual, semi-distributed hydrological model used for flow and water temperature simulation and forecasting. The hydrological module allows streamflow simulation at daily or sub-daily time steps and has been used for streamflow and/or water temperature forecasting in many case studies (Bâ *et al.*, 2018; Brisson *et al.*, 2015; Ouellet-Proulx *et al.*, 2019).

The spatialization of hydrological processes is done through a two-step discretization of the watershed. First, the watershed is divided into equal areas. These are square elements called “Whole Squares” or CE (Figure 6.2.a). The altitude, the percentage of area covered by forest, and the percentage covered by lakes and marshes are defined for each CE. Secondly, each CE is further subdivided into a maximum of four “Partial Squares” (CP) by overlapping CEs with sub-basin boundaries to determine the runoff direction (Figure 6.2.b). Each partial square is a hydrologic response unit (HRU), in which vertical routing and water storage are represented conceptually by three interconnected reservoirs. One reservoir represents storage by lakes and marshes, while the other two represent the vadose (unsaturated) and lower soil layers (Figure 6.2.b). Available water in the three reservoirs is routed downstream proportionally to the partial square area at each timestep.

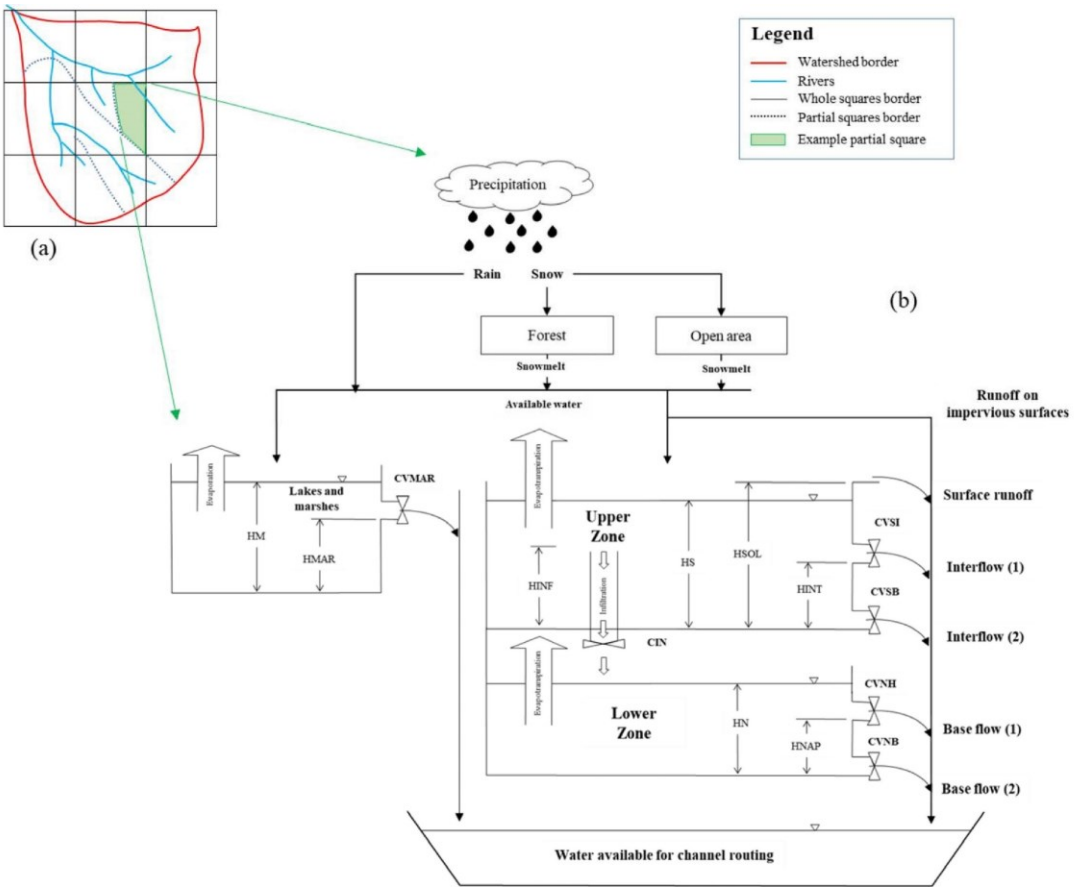

Figure 6.2 **Schematic representation of (a) Watershed, whole squares and partial squared, (b) each partial square as a hydrological response unit or HRU, which shows production function variables, water reservoirs, and CEQUEAU hydrological processes for each partial square (St-Hilaire et al. 2015). These variables are defined in Table I-A1 in Appendix I-A**

The hydrological module computes a water balance, including snowmelt and liquid precipitation as inputs, storage in reservoirs, and evapotranspiration as a negative term for the thermal heat budget. The thermal module computes a water heat budget that includes direct shortwave radiation as the primary input and latent heat as the primary loss. Net longwave radiation and sensible heat can be positive or negative, depending on the thermal gradient of the air-water interface. Heat advection from upstream, local runoff, interflow, and groundwater inflow are accounted for in the simulation.

After calculating available water in each CP, the thermal module calculates the following heat sources and sinks (St-Hilaire *et al.*, 2015):

$$\Delta H = H_{short} + H_{long} + H_{ET} + H_{sens} + H_{ups} + H_{local} + H_{ground} \quad (6.1)$$

where ΔH is the overall change of heat budget for a CP (MJ, the unit for other heat terms is the same); H_{short} is incoming shortwave radiation; H_{long} is heat gain or loss by longwave

radiation; H_{ET} is the latent heat loss (evaporation); H_{sens} is the sensible heat gain or loss; H_{ups} is the heat gain from upstream; H_{local} is the heat coming from local runoff and interflow; and H_{ground} is the heat term related to groundwater input. The surface heat budget terms are estimated using the following equations:

$$H_{short} = CRAYSO.A.R_{short} \tag{6.2}$$

$$H_{long} = CRAYIN.A.\sigma(\beta T_{air}^4 - T_{water}^4) \tag{6.3}$$

$$H_{ET} = -CEVAPO.L_{ET}.A.E \tag{6.4}$$

$$H_{sens} = CCONVE.A[0.2U(T_{air} - T_{water})] \tag{6.5}$$

where A is the water surface area in the CP (m^2); $CRAYSO$, $CRAYIN$, $CEVAPO$, and $CCONVE$ are empirical dimensionless coefficients for thermal module calibration (presented in Table 6.2); R_{short} is net solar radiation (MJ/m^2); σ is Stefan-Boltzmann constant (5.6704×10^{-8} W/m^2/K^4 or 4.8992×10^{-9} MJ/m^2/K^4/day); β is the sky emissivity (0-1); T_{air} and T_{water} are air and water temperature, respectively (K); L_{ET} is the depth of evaporated water from river reach (m), which is calculated by the hydrological module (m); E is the latent heat of evaporation (2480 MJ/m^3 considering constant water density) and U is wind speed (km/h). After calculating all the heat terms, the thermal module calculates water temperature change using the following equation:

$$\Delta T = \frac{\Delta H}{V.\theta} \tag{6.6}$$

where ΔT is the change in the water temperature (°C); V is the estimated water volume by the hydrological module (m^3), and θ is the heat capacity of water (4.186 MJ/kg/°C). All heat budget terms and their formulation are explained in detail in the CEQUEAU manual (Morin & Paquet, 2007).

Table 6.2 Parameters used in the CEQUEAU thermal module

No.	Description	Lower Bound	Upper Bound	Unit	Symbol used in CEQUEAU
1	The coefficient for shortwave radiation (ratio)	0.05	2	dimensionless	CRAYSO
2	The coefficient for longwave radiation (ratio)	-1	2	dimensionless	CRAYIN
3	The coefficient for evaporation (ratio)	0.05	2	dimensionless	CEVAPO
4	The coefficient for sensible heat (ratio)	0.05	2	dimensionless	CCONVE
5	Threshold snow stock for controlling water temperature	0	1000	mm	CRIGEL
6	Temperature of groundwater	5	10	°C	TNAP
7	Amount of minimum precipitation that will identify days with low solar radiation	5	20	mm	BASSOL
8	Correction factor for bassol (ratio)	0	1	dimensionless	CORSOL

The model's architecture allows for relatively easy addition and modification of algorithms, routines, or hydrological/thermal components (St-Hilaire *et al.*, 2015). The current model for the Nechako watershed has 271 CE and 615 CP.

CEQUEAU has 26 global parameters (i.e., the parameter value is constant for all CPs) for the hydrological module (Table I-A1 in Appendix I-A). The thermal module has eight global parameters for heat budget processes over the whole watershed (Table 6.2). In this study, all hydrological and thermal parameters are global. A schematic representation of the hydrological module with parameter identification is available in Morin and Couillard (1990) and Morin and Paquet (2007), briefly described in Figure 6.2.

6.2.3 CMA-ES calibration algorithm

Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is an evolutionary algorithm developed by (Hansen, 2023). The superiority of CMA-ES as a global optimization method able to handle high dimensionality makes it a good candidate for optimization (Hansen & Ostermeier, 1996). This algorithm is already compared extensively with other famous optimization algorithms for model calibration in hydrology (Arsenault *et al.*, 2014). Arsenault *et al.* (2014) showed CMA-ES has higher performance than most optimization algorithms in (1) finding global optima and (2) convergence speed for hydrological purposes. Appendix I-B explains that this algorithm follows four steps until a stopping criterion is met (Hansen, 2023).

6.2.4 Single-site and multisite calibration methods

In the present study, three methods were compared for model calibration:

Calibration at a single site and calculation of the performance metric for that and the other independent sites (henceforth referred to as “Method 1”);

Optimization of a multisite performance metric using all observations from all sites as a single objective optimization problem. Then using the multisite optimum, the performance metric is re-calculated at each site for validation (“Method 2”); and

Method 3: This method infers global scale parameters using sub-watershed scale estimates from Method 1. This method upscales the parameter values of single sites and utilizes the upscaled values at all sites. For this method, several operators are possible, which are explained in the following section.

The Kling–Gupta Efficiency (KGE) metric was used for hydrological calibration. KGE was introduced to provide a more balanced alternative to the Nash–Sutcliffe Efficiency (NSE) metric (which is often criticized as putting too much emphasis on high flows) by Gupta *et al.* (2009). In this study, the Root Mean Square Error (RMSE) was used for the thermal module calibration because it is used in many thermal modelling studies favouring interstudy comparability (Ouellet-Proulx *et al.*, 2019). The absolute $Bias$ performance metric is used to show the bias of the results. A summary of these performance metrics is presented in Gupta *et al.* (2009), and Krause *et al.* (2005), and the following equations define them:

$$KGE_j = 1 - ED_j \tag{6.7}$$

$$ED_j = \sqrt{(r_j - 1)^2 + (\alpha_j - 1)^2 + (\beta_j - 1)^2} \tag{6.8}$$

$$r_j = \frac{\sum_{i=1}^{N}\left(O_{i,j} - \bar{O}_j\right)\left(S_{i,j} - \bar{S}_j\right)}{\sqrt{\sum_{i=1}^{N}\left(O_{i,j} - \bar{O}_j\right)^2}\sqrt{\sum_{i=1}^{N}\left(S_{i,j} - \bar{S}_j\right)^2}} \tag{6.9}$$

$$\alpha_j = \frac{\sigma_{S_j}}{\sigma_{O_j}} \tag{6.10}$$

$$\beta_j = \frac{\bar{S}_j}{\bar{O}_j} \tag{6.11}$$

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{N}} \tag{6.12}$$

$$Bias_j = \bar{O}_j - \bar{S}_j \tag{6.13}$$

$$NSE_j = 1 - \frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{\sum_{i=1}^{N}\left(O_{i,j} - \bar{O}_j\right)^2} \tag{6.14}$$

where KGE_j, $RMSE_j$, $Bias_j$ and NSE_j are the performance metrics for the jth station; $O_{i,j}$ and $S_{i,j}$ respectively are observed and simulated values for the jth station at i^{th} time step; N is the number of time steps (measurements); σ_{S_j} and σ_{O_j} are the standard deviation of simulations and observations for station j, and $\bar{S}_j$ and $\bar{O}_j$ are mean of simulated and observed values.

6.2.5 Upscaling operators to derive parameters at spatial scales

Following thermal calibration of the model at finer scales using single sites, several sets of parameters are generated, namely one for each specific sub-watershed associated to each station. Since the CEQUEAU model is based on global parameters for the whole watershed, the next step is to derive watershed-scale parameters using the sub-watershed's parameters. Samaniego et al. (2010) provided several upscaling operators for this goal, including Arithmetic, Geometric, and Harmonic means (Table 6.3).

Table 6.3 Several methods for upscaling proposed by Samaniego *et al.* (2010) to derive parameters at level 1 using the parameters from level 0

Name	Notation	Formula
Arithmetic mean	$\mathcal{A}$	$\rho^1 = \mathcal{A}\langle\rho_i^0\rangle = \frac{1}{n}\sum_{i=1}^{n}\rho_i^0$
Geometric mean	$\mathcal{G}$	$\rho^1 = \mathcal{G}\langle\rho_i^0\rangle = \left[\prod_{i=1}^{n}\rho_i^0\right]^{\frac{1}{n}}$
Harmonic mean	$\mathcal{H}$	$\rho^1 = \mathcal{H}\langle\rho_i^0\rangle = \frac{n}{\sum_{i=1}^{n}\frac{1}{\rho_i^0}}$

Each operator derives effective parameters at spatial level 1 based on local scale parameters at spatial level 0 (Samaniego *et al.*, 2010), as shown in Table 6.3. Upscaling in this study can be described using Figure 6.3 and Equation (6.15).

$$\rho_k^1 = \mathcal{U}_k\langle\rho_{k,i}^0\rangle,\ k = 1, \dots, K \tag{6.15}$$

where $\mathcal{U}_k\langle\cdot\rangle$ is the upscaling method ($\mathcal{A}\langle\cdot\rangle$, $\mathcal{G}\langle\cdot\rangle$ and $\mathcal{H}\langle\cdot\rangle$ for Arithmetic, Geometric, and Harmonic means, respectively) for the kth parameter, and K is the number of parameters of

the model; ρ^1 denotes derived parameter value for level 1, ρ_i^0 denotes the parameter value of the sub-watershed i at level 0, and n is the number of sub-watersheds inside the whole watershed.

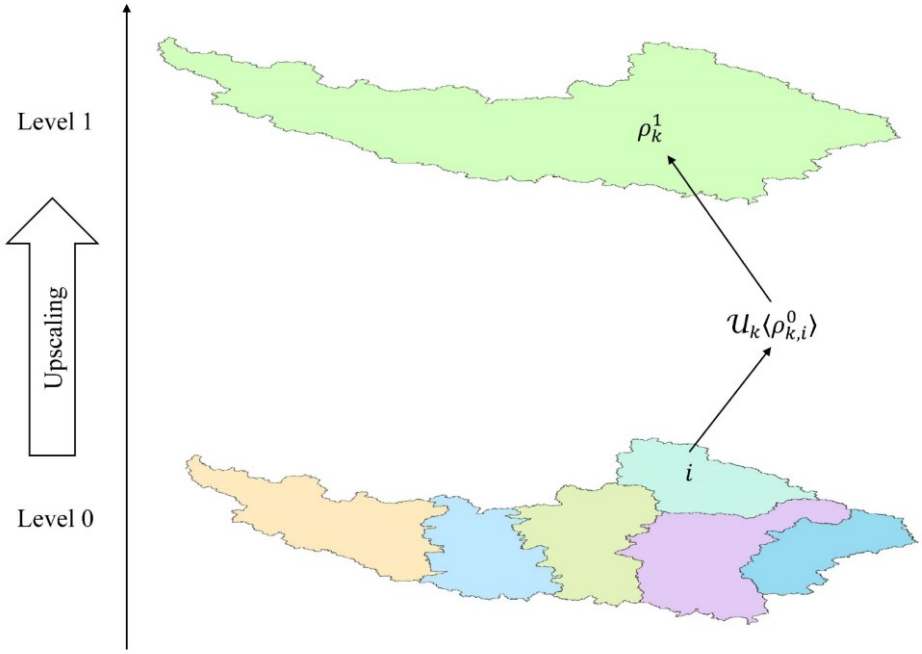

Figure 6.3 **Estimation of parameter ρ_k^1 at level 1 for k^{th} parameter of the model using upscaling operator $\mathcal{U}_k\langle\cdot\rangle$ on $\rho_{k,i}^0$ at level 0 on all sub watersheds i ($i = 1, \dots, n$)**

6.2.6 Model calibration setup and implementation

Given that the thermal regime of a river depends partly on the prevailing flow conditions, the hydrological module was calibrated once for all the available data from 2016 to 2019. Input data cover 2016 to 2019, which was selected based on available measured water temperatures (Table 6.4).

Table 6.4 Water temperature measurement locations and data availability period for each station

Station	Upstream drainage area (km^2)**	Longitude (West)	Latitude (North)	Data availability	
				Start	End
Skins Lake Spillway (CP=1109)	13821.6	-125.986373	53.781751	22 Jun 2017*	31 Dec 2019
Nautley (CP=181)	6369.8	-124.59792	54.085499	01 Jan 2016	31 Dec 2019
Station 1 (CP = 1)	25700.3	-123.78855	53.969796	27 Jun 2016	5 Oct 2016
				20 Jun 2017	30 Oct 2017
				234 days measurements	
Station 2 (CP = 23) (Vanderhoof)	25174.1	-124.005287	54.024968	01 Jan 2016	31 Dec 2019
				1461 days measurements	
Station 3 (CP = 43)	24128.5	-124.11658	54.045529	27 Jun 2016	29 Sep 2016
				25 May 2017	19 Oct 2017
				243 days measurements	
Station 4 (CP = 180)	17071.6	-124.570916	54.055841	23 Jun 2017	19 Dec 2017
				10 Apr 2018	27 May 2018
				01 Jan 2019	18 Dec 2019
				623 days measurements	
Station 5 (CP = 186)	17056.3	-124.532011	54.05397	27 Jun 2016	29 Sep 2016
				25 May 2017	13 Aug 2017
				168 days measurements	
Station 6 (CP = 352)	16103.7	-124.621282	53.769992	28 Jun 2016	29 Sep 2016
				24 May 2017	20 Oct 2017
				244 days measurements	
Station 7 (CP = 491)	15246.8	-124.933627	53.645354	28 Jun 2016	28 Sep 2016
				24 May 2017	17 Oct 2017
				240 days measurements	

* The gap period (01 Jan 2016-21 Jun 2017) was filled using the corresponding day-of-year mean values

* The drainage areas include the upstream area of Kenny dam and Skins Lake Spillway

Then, using the hydrologically calibrated model, the thermal module calibration was performed. The thermal calibrations were done using the period ranging from April to November using daily timestep. This period was identified by looking at the mean interannual hydrograph at each site (see Figure 6.4) and the observed water temperature periods (Table 6.4).

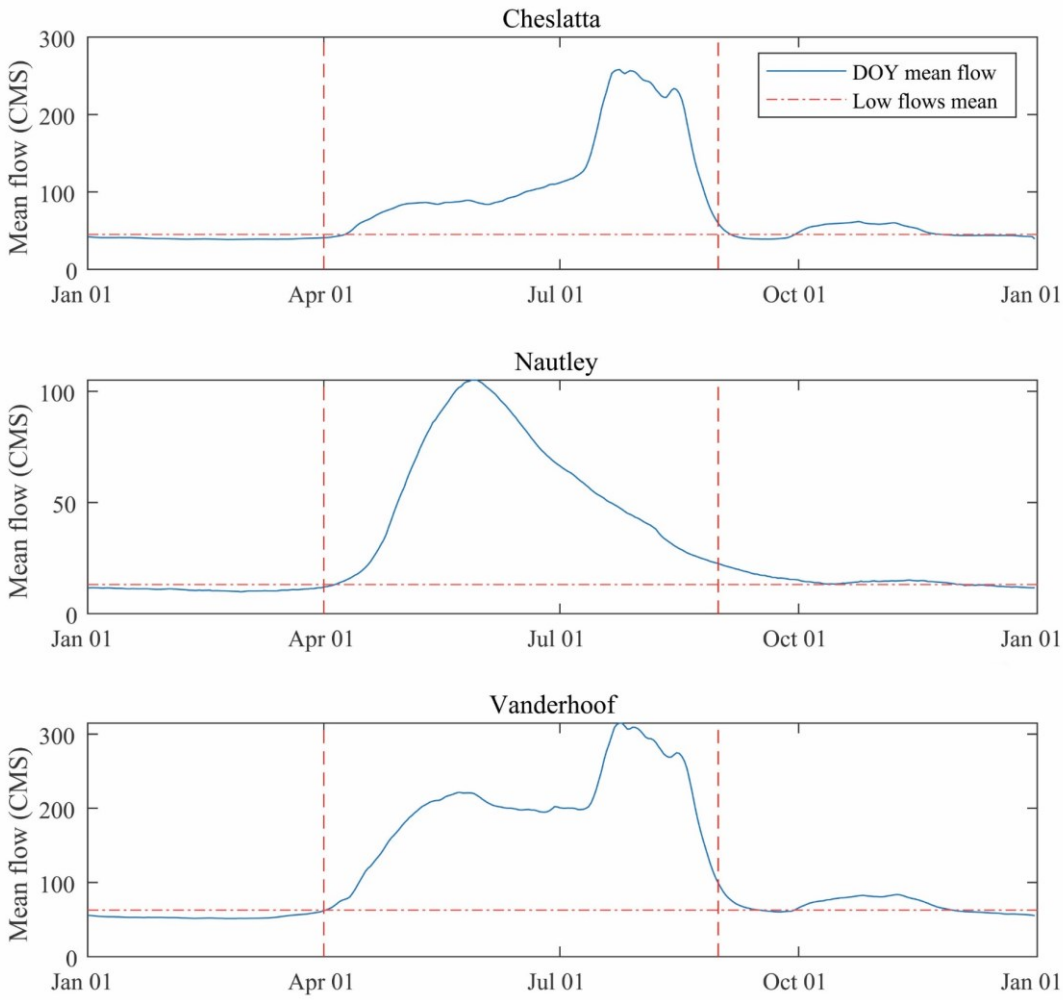

Figure 6.4 Identification of low and high flow periods for three hydrometric stations

Low flow months cover the period from September to March, inclusively, while the others (April through August) are the high flow months. On average, 66% (range from 40% to 86%) of flow at Vanderhoof comes from the Cheslatta sub-watershed, including releases from the Skins Lake Spillway, 24% (range from 10% to 50%) from the Nautley sub-watershed, and the remaining discharge comes from other parts of the watershed. At the Nautley station, flow is natural, and the peak flow occurs between May and June. Therefore, flows downstream of the Nechako-Nautley confluence are greatly influenced by those originating from the spillway.

Based on the “Calibration Method” and the “Target Station,” 11 different scenarios were defined in Table 6.5. The mean of lower and upper boundaries for each decision variable were used as initial parameter sets for all calibration scenarios.

Table 6.5 Different Scenarios defined for calibration

Calibration Method	Target station(s)	Scenario No.
Single-Site	1	1
	2	2
	3	3
	4	4
	5	5
	6	6
	7	7
Multisite	1-7	8
Upscaling 1 ($\mathcal{A}$)	1-7	9
Upscaling 2 ($\mathcal{G}$)	1-7	10
Upscaling 3 ($\mathcal{H}$)	1-7	11

Since observed water temperatures began in April, the period of January to March 2016 was of sufficient length to act as the spin-up period for all scenarios; all the available data for each station were used for the calibration, and the validation was done using the other stations. Historical measured water releases and water temperatures are imposed at the spillway and the Nautley station for each day. For each scenario, the same hydrologically calibrated module was used. Then, the thermal module was calibrated using the flows generated by the calibrated hydrological module. All parameters were calibrated using the aforementioned CMA-ES optimization method using a stopping criterion of a maximum of 5000 model evaluations. The hydrological/thermal calibration and validation process for the CEQUEAU model is summarized in Figure 6.5.

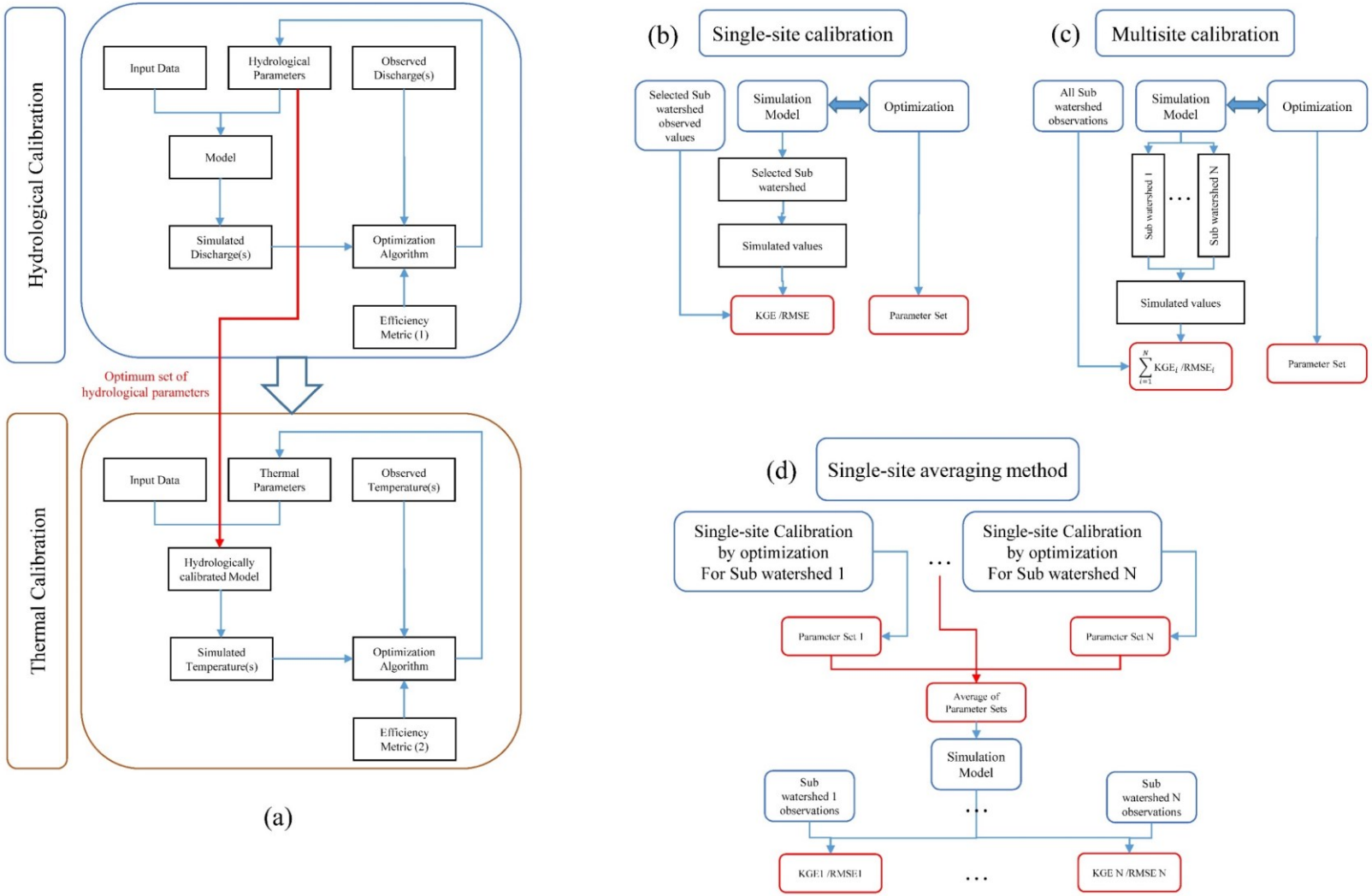

Figure 6.5 **(a) Hydrological and thermal calibration process diagram; (b) Single-site calibration process (Method 1); (c) Multisite calibration method (Method 2) and (d) single-site averaging method (Method 3)**

6.3 Results

6.3.1 Hydrological module parameters and performance metrics

The hydrological calibration method was completed for the Vanderhoof station (Station 2 in Figure 6.1), which is the only station with long-term water flow measurements. The primary performance metric used for hydrological calibration was the KGE, calculated and presented in Figure 6.6.

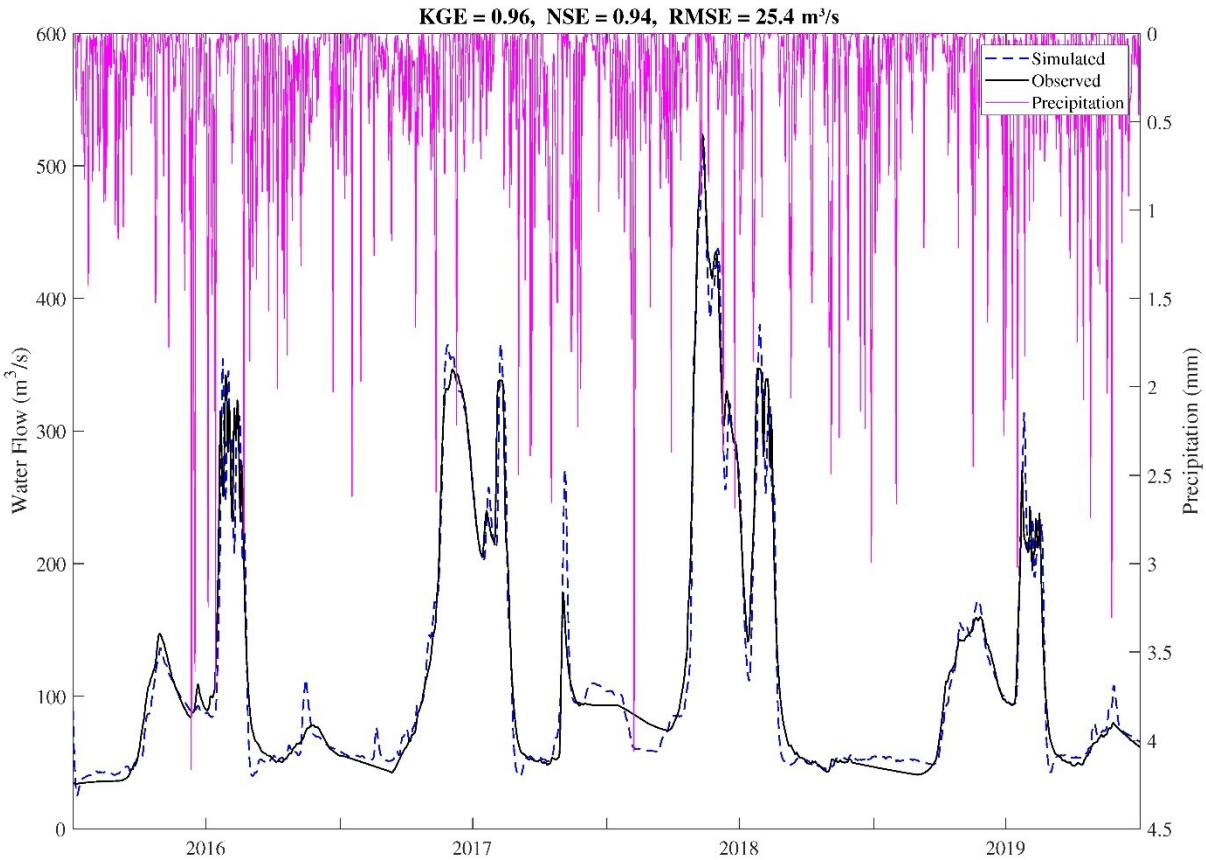

Figure 6.6 **Simulated flow using hydrological calibrated parameters for the Vanderhoof station. This set of parameters and simulation for the whole basin was used for thermal calibration and simulations**

Table I-A1 in Appendix I-A shows the 26 calibrated parameter values for the hydrological module. Using these parameters, the simulated discharge values at Vanderhoof station are presented in Figure 6.6. Since hydrological calibration is not the main focus of this paper, the maximum performance metric for simulated water flows was targeted by using all data without a validation step. The KGE = 0.96, and NSE = 0.94 show that the calibration is of high quality and can accurately reproduce the hydrograph at Vanderhoof. This is not surprising since the main flow of the watershed is imposed using two points (SLS and the Nautley station). High flows were of particular interest in this calibration because a difference between simulated and

observed high flows typically corresponds to high temperatures and is one of the possible sources of error for heat budget calculations. The hydrological model RMSE over the four years is 25.4 m^3/s (mean annual observed flow = 116.2 m^3/s). This relatively good hydrological calibration will allow studying the errors for thermal modelling and the calibration method selection impact on overall error while minimizing the impacts caused by biased or inadequate discharge simulations.

6.3.2 Thermal module parameters and performance metrics

Table 6.6 shows thermal parameters after calibration for 11 scenarios.

Table 6.6 Thermal parameters values for different calibration scenarios

Scenario No.	CRAYSO	CRAYIN	CEVAPO	CCONVE	CRIGEL (mm)	TNAP (°C)	BASSOL (mm)	CORSOL
1	2.00	2.00	0.64	2.00	605.58	8.24	14.67	0.00
2	1.37	1.96	0.05	1.72	0.01	9.28	12.28	0.00
3	2.00	2.00	0.49	2.00	999.72	9.05	8.46	0.02
4	1.37	2.00	0.05	1.27	92.30	9.91	13.40	0.00
5	2.00	1.69	0.05	2.00	291.36	6.80	17.61	0.65
6	2.00	1.52	0.05	2.00	419.92	6.89	7.36	0.15
7	2.00	1.44	0.05	2.00	606.97	9.90	11.05	0.68
8	1.69	2.00	0.05	2.00	0.00	5.81	11.21	0.00
9 [μ = A$\langle\rho_k^1\rangle$]	1.82	1.80	0.20	1.86	430.84	8.58	12.12	0.21
10 [G$\langle\rho_k^1\rangle$]	1.80	1.79	0.10	1.83	89.24	8.49	11.66	0.00
11 [H$\langle\rho_k^1\rangle$]	1.77	1.77	0.07	1.81	0.08	8.40	11.19	0.00
σ_μ	0.31	0.25	0.26	0.28	342.78	1.31	3.55	0.31

Also, Table 6.7 shows performance metrics for the thermal module calibration using three methods.

Table 6.7 **Performance metrics by calibrating/validating the thermal module parameters of CEQUEAU using temperature measurements at different locations for: [1] focused site for calibration (Method 1) and [2] the other two sites which are not the focus of calibration**

Scenario Number	Target Station	[1] RMSE (°C) for each station in Method 1		[2] RMSE (°C) at each site using derived parameters from calibration/upscaling (°C)							
		RMSE (°C)	Days	1	2	3	4	5	6	7	Overall RMSE
1	1	0.74	234	0.74	1.81	0.89	1.80	1.81	2.07	2.14	1.75
2	2	1.26	1461	1.24	1.26	1.39	1.48	2.47	2.43	2.41	1.61
3	3	0.89	243	0.75	1.87	0.89	1.84	1.76	2.04	2.12	1.78
4	4	1.35	623	1.28	1.37	1.46	1.35	2.59	2.47	2.44	1.66
5	5	1.46	168	1.66	2.86	1.50	2.72	1.46	1.96	1.99	2.50
6	6	1.93	244	1.29	2.44	1.21	2.28	1.53	1.93	2.01	2.16
7	7	1.99	240	1.93	3.10	1.71	2.92	1.47	1.99	1.99	2.69
8	1-7	-	3213	0.83	1.40	1.01	1.60	2.03	2.19	2.23	1.57
9	1-7	-	-	0.91	1.99	0.99	1.90	1.79	2.07	2.13	1.86
10	1-7	-	-	0.85	1.79	0.99	1.71	1.87	2.09	2.16	1.74
11	1-7	-	-	0.86	1.63	1.00	1.68	1.90	2.10	2.17	1.66

Method 1 results are presented in Column [1], Scenarios 1-7. By using the single-site method (Method 1), RMSEs are calculated at each station, starting with the station used for calibration (Column [1]). The resulted time series for each station are shown in Figures I-C1 to I-C7 in Appendix I-C. Column [2] in Table 6.7 shows the RMSE values obtained at the other stations using the same parameters. Figures I-D1 to I-D7 in Appendix I-D show the resulting time series for all stations using each scenario. Since single-site calibration is a local optimum just for one station, the other six stations show deviations of their optima. To quantify the total deviation of optima, the overall RMSEs were calculated in the last column. The formula for this overall RMSE is as follows:

$$RMSE_{Overall} = \sqrt{\frac{\sum_{i=1}^{N} RMSE_i^2 \times M_i}{\sum_{i=1}^{N} M_i}} \tag{6.16}$$

where $RMSE_{Overall}$ is the overall $RMSE$ for a set of parameters over all stations, $RMSE_i$ is the $RMSE$ value for station i, using the total days of observation at station i (M_i), and N is the number of stations.

For the scenario 8 (multisite calibration), the objective function was the $RMSE$ calculated using all stations. In this scenario, using the resulting parameters, $RMSE_i$, $i = 1, \ldots, 7$ were calculated (Scenario 8, Column [2]). Resulted time series for each station using this scenario is depicted in Figure I-D8 in Appendix I-D. Finally, upscaled parameters were calculated using the seven parameter sets from seven single-site optimizations. These parameters resulted in single site $RMSE$s and $RMSE_{Overall}$, presented in Scenarios 9-11 in Table 6.7. It is worth mentioning that RMSE values less than 2°C are required to consider the simulations adequate to provide insight into management perspectives on the Nechako watershed.

As expected, for each individual site, Method 1 shows the best performances. However, since focusing just on one station degrades the performance metric at the other sites, we chose $RMSE_{Overall}$ as the benchmark for comparison between the three calibration methods. The $RMSE_{Overall}$ ranges from 1.57°C (for the multisite method) to 2.69°C (for simulation based on the parameters obtained from calibration at the most upstream station). The order of magnitude of $RMSE_{Overall}$ values for stations 1 to 4 (1.61 to 1.78°C) differ from stations 5 to 7 (2.16 to 2.69°C).

Figures I-C1 to I-C7 in Appendix I-C show the simulated versus observed time series for stations 1-7. The resulting time series for stations 1-4 show unbiased simulations while resulting simulations for stations 5-7 are biased in a way that underestimates the observations. By analyzing the location of all stations on the watershed, it is evident that stations 5-7 are located in the proximity of sizeable upstream water bodies. Station 7 is immediately downstream of Cheslatta lake, and Station 6 is close to the Nechako River's vast meander-shaped reach, close to Copley Lake. Station 5 is also likely affected by a large meandering reach located upstream. For stations 1-4, there is no large, slow-moving volume of water immediately upstream of the stations, so the river reach has a more stable depth and width.

6.3.3 Energy flux contribution to water temperature and the data availability impact on calibration

Figure 6.7 presents the time series of the main heat fluxes at the Vanderhoof station (Station 2) to show the contribution of different local heat budget terms on water temperature.

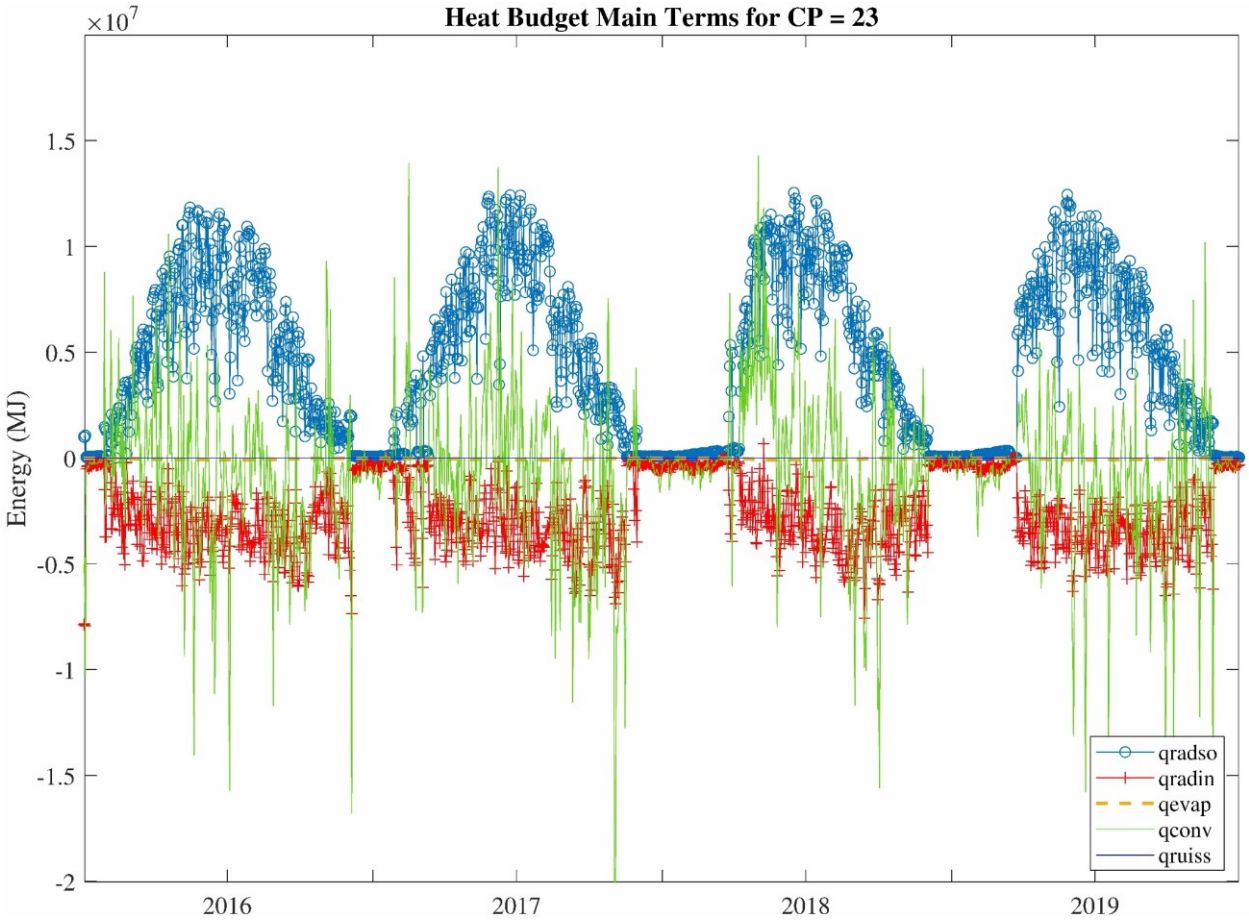

Figure 6.7 **Heat budget component for Vanderhoof station (Station 2). Heat terms are: qradso = solar shortwave radiation, qradin = longwave radiation, qevap = evapotranspiration, qconv = sensible heat, qruiss = local advection**

The primary local heat components are shortwave radiation, longwave radiation, and sensible heat fluxes. Figure 6.7 shows that the primary local heat source is solar shortwave radiation (as expected), while the leading heat sinks are longwave radiation and sensible heat. The latter shows variability during the cold period (December to March), and its actual impact on water temperature during this period becomes nil once freezing has occurred.

6.4 Discussion

Many studies still use single-site calibration for thermal calibration of semi-distributed hydrological models. With growing monitoring spatial density, multisite calibration is becoming increasingly possible. With three stations for long-term measurements on the Nechako watershed and seven stations with partial time series (during four summers from 2016 to 2019), this study compares three possible thermal module calibration methods for the CEQUEAU model using multisite measurements.

6.4.1 The logic of calibration strategies for global parameters

The use of global parameters for model calibration implies that some spatial heterogeneity may not be fully accounted for. While a relatively large number of parameters could vary spatially, this variability would have to be based on some knowledge of the physical process that is being modelled. For instance, the snowmelt threshold varies with altitude and canopy cover, as well as basin azimuth. Using a basin-wide mean for the snowmelt temperature threshold is in agreement with a desire to abide by the principle of parsimony, thereby sacrificing some representativeness of snowmelt heterogeneity which is valid for a number of thermal parameters as well.

There is no clear answer to which averaging method is the best for the parameter's upscaling. To this goal, three upscaling methods were used. Among the three methods, the harmonic mean resulted in the smallest value for overall RMSE compared to the other two methods. However, the difference with the optimal method (multisite calibration) is visible, especially for Station 2 at Vanderhoof, which has more importance from a water management perspective and has the most prolonged observation period.

6.4.2 Hydrological module calibration impacts on thermal calibration

Flow influences the water temperature in two main ways: firstly, by defining the volume of water to be heated or cooled, and secondly, by advecting heat from upstream to downstream. Hence, adequate multisite calibration efficiency for hydrological modelling is necessary for successful thermal calibration. Because flows are imposed as boundary conditions at two points upstream of the study area, this study de facto eliminates a large portion of uncertainty/bias that could have been associated with simulated discharge used as inputs for thermal modelling. This approach offers the best performance metrics for the third flow measuring site at Vanderhoof. Nkiaka *et al.* (2018) stated that NSE > 0.6 is the minimum threshold for a good calibration. The KGE and NSE performance metrics close to 1 (best possible value) and RMSE = 25.4 m^3/s (Mean flow during the period is 116.2 m^3/s) for the four years ensure that the flow uncertainty has low impacts on thermal modelling. Compared to previous studies, the KGE > 0.9 obtained in this study shows higher performance compared to a similar study by Barnhart *et al.* (2014) with KGE ≈ 0.8. Moreover, the heat term budget associated with the advection of heat from local water flow in each CE (Q_{ruiss} in Figure 6.7) is not one of the primary heat budget terms. The RMSE performance metric in this study is comparable and/or lower than previous studies on the Nechako watershed. For example, Ouellet-Proulx *et al.* (2017b) presented RMSE under 1.6°C for calibration and validation at the Vanderhoof. Ouellet-Proulx *et al.* (2019), improved the CEQUEAU model performance for latent heat budget calculation from 1.6°C to 0.99-1.51°C using several evaporation equations

for summertime at the Vanderhoof. Our study showed that single site calibration for Vanderhoof, similar to Ouellet-Proulx *et al.* (2019), is 1.26°C.

Gatien *et al.* (2022) presented RMSEs less than 1°C at Vanderhoof for daily and sub-daily time steps for summer periods, which was made possible by calibrating each year independently, using hydraulic modelling. Their model simulates water temperatures using the HEC-RAS 1D model and the same ERA5 data used in the present study as inputs. However, their results focused on a river reach compared to a watershed scale analysis in this study.

6.4.3 The calibrated parameter values and their impacts

One of the advantages of multisite calibration is finding a more globally optimum set of parameters for all sub-watersheds using multisite observations (Brown & Craig, 2020; Ricard *et al.*, 2013). This calibration method is helpful for semi-distributed models like CEQUEAU since the resulting set of parameters accounts for some of the spatial variability to be expected in large watersheds. The current version of CEQUEAU uses global parameters for the thermal module, while it has the ability to spatialize some of the parameters (16 out of 26) in the hydrological module. One study analyzed the sensitivity of some CEQUEAU parameters for hydrological modelling, while no previous study delved into the sensitivity analysis of CEQUEAU to thermal parameters. Brisson *et al.* (2015) showed that CIN, CVNB, CVSB, CVSI, HINF, HINT, HNAP, and HPOT are the parameters to which the model is most sensitive in their Quebec (Canada) case study.

Based on Figure 6.7, Nechako's main surface heat flux terms are shortwave radiation, longwave radiation, and sensible heat which are adjusted by CRAYSO, CRAYIN, and CCONVE parameters, respectively, during calibration. As discussed earlier, finding more accurate meteorological data could improve the model performance and results, especially for these variables. The spatialization of these parameters should be tested in future studies.

Other than the parameters mentioned above, it can be seen that CEVAPO is showing a slight variation close to 0, which means lower estimations for calculated latent heat in adjusting the water temperature. Figure 6.7 confirms that this heat budget term is not dominant for cooling water in the Nechako, using the ERA5 dataset. This relative impact among heat budget terms should be tested using observed meteorological data or other reanalysis databases.

On the other hand, although CRIGEL is showing considerable variation among different scenarios, the overall RMSE is close for them. CRIGEL is an estimate of the amount of snowpack on the ground above which the water temperature is 0. So, this parameter has no impact on water temperature in the summer. Since summer months were used to calibrate the

model, this variation shows the insensitivity of summer calibration to a parameter related to the onset of the spring snowmelt season.

Three calibration parameters show less variability: TNAP, BASSOL, and CORSOL. There are physical reasons behind their close values between stations. TNAP is the groundwater temperature. Since surface phenomena have less impact on groundwater, it shows minor variation over the year. This physical fact follows our results which show TNAP values between 6 to 10°C. The BASSOL parameter refers to the precipitation depth, which results in a full cloud cover and for rainy days, subsequently, less solar radiation. Since the cloud thickness and precipitation directly correlate, it is a parameter with physical bases. The values for this parameter vary from 7 to 16 mm, which is a reasonable estimate for cloudy/rainy days. The CORSOL is the correction factor for BASSOL active for rainy days, and its value close to 0 indicates unbiased estimations for BASSOL.

As can be seen, CEQUEAU successfully simulates the hydrological/thermal response despite the fact that the parameters of the thermal module are global. The thermal module includes eight global (i.e., spatially constant) parameters in its current version. Comparing overall RMSE for different calibration methods (Table 6.7) suggests that the multisite calibration method is better for thermal module calibration, considering performance metrics (overall RMSE = 1.57°C for all observed data) and computational demand. The possible reason for this superiority is the amount of information available for the calibration algorithm.

On the other hand, for thermal module calibration, the single-site average methods (scenarios 9-11) showed better overall RMSEs than the stations upstream of the Nautley confluence (scenarios 5-7). This finding can be related to the impact of large water bodies upstream of the Nautley station, which likely impacts the river's thermal regime. Since the large lakes warm up the water during summer, and the current structure of the CEQUEAU does not fully consider this effect, systematic underestimations of observed temperature were found in simulated temperatures at stations 5-7. From a physical perspective, for a given constant water volume, more water surface area in summer means greater solar radiation heat flux and a positive heat budget term, leading to higher observed temperatures than smaller exposed water surface areas. This bias, which is also present for single-site calibrated parameters, also affected the average-based methods, which is clear for the overall RMSE performance metric for scenarios 9-11 compared to scenario 8. In its current configuration, the CEQUEAU model does not model the thermal impact of big lakes in a detailed manner (e.g., lake bathymetry, width, and depth are not explicitly considered, and only area is accounted for). At the stations downstream of the Nautley confluence, averaging methods are not superior in performance metrics and computational demand. This is most likely due to the fact that lakes have a lesser impact in those reaches.

This study showed that, with overall RMSEs ranging from 0.74 to 3.1°C for seven measuring stations. Barnhart *et al.* (2014) reported RMSE values ranged from 1.64 to 3.18°C for five stations using the SWAT model. Although inter-model/inter-basin comparisons with a long time difference have obvious limits, considering a more distributed model like SWAT compared to a semi-distributed model like CEAUEAU with parsimonious parametrization confirms that the CEQUEAU model provides acceptable output with a simpler structure.

6.4.4 Final selection for the calibration method

Although single-site calibration resulted in the best local optima for each site, the multisite calibration method is superior based on the results in this study. Each method has its advantages and disadvantages. The geographical location of stations has an impact on thermal calibration. Local processes often have more impacts on thermal calibration, including spatial variability on meteorological input data. Using all available stations for thermal calibration is recommended since pooling all stations and performing a single multisite calibration is more time-efficient, with a better overall performance metric. The time required for multisite calibration for distributed models such as CEQUEAU is almost equal to that for a single-site calibration. The single-site average method required N times more computing resources, with N being the number of stations. For the Nechako watershed, the most crucial time of the year is the high flow days in summer, during which solar radiation is at its peak. Excluding stations 5-7, the results of this study show that the multisite method provides RMSE ≤ 1.6°C for all stations and the overall RMSE=1.57°C. Although the average parameter methods have better overall RMSEs than single-site calibrations, using them shows more deviations at every site with an overall RMSE ≥ 1.66°C. Given the stated importance of good model performance at Finmoore (≈ 15 km downstream of Station 1) and that strong model performance is desirable at other stations, the multisite calibration method is recommended for thermal calibration in this watershed, where it provides RMSE ≤ 1°C for Station 1. This result confirms the results of Shen *et al.* (2022) for hydrological calibration. Excluding one or several stations in calibration is the spatial equivalent of a “split sample approach” in hydrology. Arsenault *et al.* (2018) and Shen *et al.* (2022) recommended that this method is likely flawed and should be replaced with other methods. Our results showed similar results for thermal modelling, and it is recommended to use all available data for model calibration (multisite calibration) and decision-making.

6.5 Conclusion

This study’s main objective was to compare three thermal model calibration methods for the Nechako watershed in British Columbia, Canada. Reliable streamflow and temperature

simulation for this watershed are vital because it includes key habitat for vulnerable fish species. A hydrological module calibration for streamflow was first performed, followed by one for water temperature. The parameters of the thermal module of the CEQUEAU model were first optimized in a single-site manner and then using the multisite method by pooling all seven sites together. Lastly, simulations were done using upscaled parameters (e.g., single-site parameter averages) to compare to the other two methods. Our results showed that multisite methods lead to the global optima for overall RMSE for all stations simultaneously, resulting in negligible loss of performance at each site compared to the single-site method as a benchmark. Also, it seems upscaled parameters do not have added value compared to multisite calibration, but they have lower overall RMSE values than most single-site calibrations.

For the Nechako watershed, while a hydrologic calibration of the model is essential, the water temperature is highly impacted by local atmospheric forcings and conditions. Therefore, this study shows major heat contribution terms control the water temperature. Stations 1 to 4 showed better performance metrics for thermal calibration, possibly because the river reach is not impacted by large water bodies or dramatic changes like a confluence. Compared with basin-scale models in the literature, this study shows that the CEQUEAU model provides competitive overall and single-site performance metrics while using a more straightforward parameterization approach. The results confirmed previous studies about the benefits of using all available data (multisite method) for calibration instead of a spatial equivalent to split sampling. Future studies are recommended to study the impact of lakes on the thermal regime of rivers using CEQUEAU.

Acknowledgement

This work was funded by the Canadian Natural Sciences and Engineering Research Council (NSERC) and Rio Tinto as part of a Collaborative Research and Development grant (Grant Number: CRDPJ 523640-18).

ARTICLE I BIBLIOGRAPHIE

Allen D, Dietrich W, Baker P, Ligon F, Orr B (2007). Development of a mechanistically based, basin-scale stream temperature model: applications to cumulative effects modeling. *Proceedings of the redwood region forest science symposium: What does the future hold? .* p 11-24.

Arsenault R, Brissette F, Martel J-L (2018). The hazards of split-sample validation in hydrological model calibration. *Journal of hydrology* 566:346-362.

Arsenault R, Poulin A, Côté P, Brissette F (2014). Comparison of stochastic optimization algorithms in hydrological model calibration. *Journal of Hydrologic Engineering* 19(7):1374-1384.

Bâ KM, Diaz V, Gómez-Albores MA, Díaz-Delgado C, Nájera-Mota N, Seidou O, Ortiz F (2018). Spatially distributed hydrological modelling of a Western Africa basin. *EPiC Ser. Eng* 3:343-350.

Bai J, Shen Z, Yan T (2017). A comparison of single- and multi-site calibration and validation: a case study of SWAT in the Miyun Reservoir watershed, China. *Frontiers of Earth Science* 11(3):592-600.

Barnhart BL, Whittaker GW, Ficklin DL (2014). Improved stream temperature simulations in SWAT using NSGA-II for automatic multi-site calibration. *Transactions of the ASABE* 57(2):517-530.

Becknell B, Imhoff J, Kittle J, Donigian A, Johanson R (1993). *Hydrological simulation program: Fortran. User's manual for release 10.* AQUA TERRA Consultants, United States. 667 p

Blöschl G (2001). Scaling in hydrology. *Hydrological Processes* 15(4):709-711.

Blöschl G, Sivapalan M (1995). Scale issues in hydrological modelling: A review. *Hydrological Processes* 9(3-4):251-290.

Boyd M, Kasper B (2003). Analytical methods for dynamic open channel heat and mass transfer: Methodology for heat source model Version 7.0.).

Brisson C, Boucher MA, Latraverse M (2015). Illustration of the added value of using a multi-site calibration and correction approach to reconstruct natural inflows and inter-catchment transfer flow: A case study. *Canadian Journal of Civil Engineering* 42(5):342-352.

Brown G, Craig JR (2020). Structural calibration of an semi-distributed hydrological model of the Liard River basin. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 45(4):287-303.

Cao W, Bowden WB, Davie T, Fenemor A (2006). Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrological Processes* 20(5):1057-1073.

Chapra S, Pelletier G, Tao H (2008). QUAL2K: A modeling framework for simulating river and stream water quality, version 2.11. *Documentation and Users Manual*.

Cole TM, Wells SA (2006). *CE-QUAL-W2: A two-dimensional, laterally averaged, hydrodynamic and water quality model, version 3.5.* US Army Engineering and Research Development Center. 679 p

Dong F, Neumann A, Kim DK, Huang J, Arhonditsis GB (2019). A season-specific, multi-site calibration strategy to study the hydrological cycle and the impact of extreme-flow events along an urban-to-agricultural gradient. *Ecological Informatics* 54:100993.

Duan Q, Gupta HV, Sorooshian S, Rousseau AN, Turcotte R (2003). *Calibration of watershed models.* American Geophysical Union, Washington, USA. 345 p

Dugdale SJ, Hannah DM, Malcolm IA (2017a). River temperature modelling: A review of process-based approaches and future directions. *Earth-Science Reviews* 175:97-113.

Environment and Climate Change Canada (2022). Historical Hydrometric Data. (Government of Canada, Canada).

Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D (2007). The shuttle radar topography mission. *Reviews of Geophysics* 45(2):33.

Ficklin DL, Luo Y, Stewart IT, Maurer EP (2012). Development and application of a hydroclimatological stream temperature model within the Soil and Water Assessment Tool. *Water Resources Research* 48(1):16.

Franco ACL, Oliveira DYd, Bonumá NB (2020). Comparison of single-site, multi-site and multi-variable SWAT calibration strategies. *Hydrological Sciences Journal* 65(14):2376-2389.

Gatien P, Arsenault R, Martel J-L, St-Hilaire A (2022). Using the ERA5 and ERA5-Land reanalysis datasets for river water temperature modelling in a data-scarce region. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 10.1080/07011784.2022.2113917:1-18.

Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of hydrology* 377(1-2):80-91.

Hansen N (2023). The CMA evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772* https://doi.org/10.48550/arXiv.1604.00772.

Hansen N, Ostermeier A (1996). Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. *Proceedings of IEEE international conference on evolutionary computation.* (Nagoya, Japan, IEEE, p 312-317.

Hasan MA, Pradhanang SM (2017). Estimation of flow regime for a spatially varied Himalayan watershed using improved multi-site calibration of the Soil and Water Assessment Tool (SWAT) model. *Environmental Earth Sciences* 76(23):787.

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D (2020b). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146(730):1999-2049.

Karra K, Kontgis C, Statman-Weil Z, Mazzariello JC, Mathis M, Brumby SP (2021). Global land use/land cover with Sentinel 2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS.* (Brussels, Belgium, IEEE, p 4704-4707.

Krause P, Boyle D, Bäse F (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in geosciences* 5:89-97.

LeBlanc RT, Brown RD, FitzGibbon JE (1997). Modeling the Effects of Land Use Change on the Water Temperature in Unregulated Urban Streams. *Journal of Environmental Management* 49(4):445-469.

Leta OT, Griensven Av, Bauwens W (2017). Effect of single and multisite calibration techniques on the parameter estimation, performance, and output of a SWAT model of a spatially heterogeneous catchment. *Journal of Hydrologic Engineering* 22(3):05016036.

Macdonald JS (2019). *Saik'uz First Nation and Stellat'en First Nation v. Rio Tinto Alcan, et al. BCSC No.116524, Vancouver Registry, Sockeye Salmon and Water Temperature.* Department of Justice Canada, Canada. 1-48 p

Morin G, Couillard D (1990). Predicting river temperatures with a hydrological model. *Encyclopedia of fluid mechanic, surface and groundwater flow phenomena,* Volk Gulf Publishing Company, Houston, Texas Vol 10. p 171-209.

Morin G, Paquet P (2007). *Modèle hydrologique CEQUEAU.* INRS, Centre Eau, Terre et Environnement, Quebec, Canada. 458 p

Nkiaka E, Nawaz NR, Lovett JC (2018). Effect of single and multi-site calibration techniques on hydrological model performance, parameter estimation and predictive uncertainty: a case study in the Logone catchment, Lake Chad basin. *Stochastic Environmental Research and Risk Assessment* 32(6):1665-1682.

Odusanya AE, Mehdi B, Schürz C, Oke AO, Awokola OS, Awomeso JA, Adejuwon JO, Schulz K (2019). Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data-sparse catchment in southwestern Nigeria. *Hydrology and Earth System Sciences* 23(2):1113-1144.

Ouellet-Proulx S (2018). *Prévision thermique d'ensemble en rivière avec assimilation de données*. Doctoral (Université du Québec, Institut national de la recherche scientifique, Quebec, Canada). 222 p. https://espace.inrs.ca/7621/1/T750.pdf

Ouellet-Proulx S, St-Hilaire A, Boucher M-A (2017b). Water temperature ensemble forecasts: Implementation using the CEQUEAU model on two contrasted river systems. *Water* 9(7):457.

Ouellet-Proulx S, St-Hilaire A, Boucher MA (2019). Implication of evaporative loss estimation methods in discharge and water temperature modelling in cool temperate climates. *Hydrological Processes* 33(22):2867-2884.

Ouellet V, St-Hilaire A, Dugdale SJ, Hannah DM, Krause S, Proulx-Ouellet S (2020). River temperature research and practice: Recent challenges and emerging opportunities for managing thermal habitat conditions in stream ecosystems. *Science of The Total Environment* 736:139679.

Pandey VP, Dhaubanjar S, Bharati L, Thapa BR (2020a). Spatio-temporal distribution of water availability in Karnali-Mohana Basin, Western Nepal: Climate change impact assessment (Part-B). *Journal of Hydrology: Regional Studies* 29:100691.

Pandey VP, Dhaubanjar S, Bharati L, Thapa BR (2020b). Spatio-temporal distribution of water availability in Karnali-Mohana Basin, Western Nepal: Hydrological model development using multi-site calibration approach (Part-A). *Journal of Hydrology: Regional Studies* 29:100690.

Qin J, Yang K, Lu N, Chen Y, Zhao L, Han M (2013). Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia. *Remote Sensing of Environment* 138:1-9.

Ricard S, Bourdillon R, Roussel D, Turcotte R (2013). Global calibration of distributed hydrological models for large-scale applications. *Journal of Hydrologic Engineering* 18(6):719-721.

Rutherford JC, Blackett S, Blackett C, Saito L, Davies-Colley RJ (1997). Predicting the effects of shade on water temperature in small streams. *New Zealand Journal of Marine and Freshwater Research* 31(5):707-721.

Samaniego L, Kumar R, Attinger S (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research* 46(5):25.

Shen H, Tolson BA, Mai J (2022). Time to update the split-sample approach in hydrological model calibration. *Water Resources Research* 58(3):e2021WR031523.

Shrestha MK, Recknagel F, Frizenschaf J, Meyer W (2016). Assessing SWAT models based on single and multi-site calibration for the simulation of flow and nutrient loads in the semi-arid Onkaparinga catchment in South Australia. *Agricultural Water Management* 175:61-71.

St-Hilaire A, Boucher M-A, Chebana F, Ouellet-Proulx S, Zhou QX, Larabi S, Dugdale S, Latraverse M (2015). Breathing a new life to an older model: The CEQUEAU tool for flow and water temperature simulations and forecasting. *Proceedings of the 22nd Canadian Hydrotechnical Conference.* (Montréal, Canada, April 29–May 2), p 10.

Theurer FD, Voos KA, Miller WJ (1984). *Instream water temperature model. Instream flow information paper 16.* US Fish and Wildlife Service, p

Van Niel TG, McVicar TR, Roderick ML, van Dijk AIJM, Beringer J, Hutley LB, van Gorsel E (2012). Upscaling latent heat flux for thermal remote sensing studies: Comparison of alternative approaches and correction of bias. *Journal of hydrology* 468-469:35-46.

Xu T, Liu S, Xu L, Chen Y, Jia Z, Xu Z, Nielson J (2015). Temporal upscaling and reconstruction of thermal remotely sensed instantaneous evapotranspiration. *Remote Sensing* 7(3):3400-3425.

Xue X, Zhang K, Hong Y, Gourley JJ, Kellogg W, McPherson RA, Wan Z, Austin BN (2016). New multisite cascading calibration approach for hydrological models: Case study in the red River Basin Using the VIC model. *Journal of Hydrologic Engineering* 21(2):05015019.

Zanaga D, Van De Kerchove R, De Keersmaecker W, Souverijns N, Brockmann C, Quast R, Wevers J, Grosu A, Paccini A, Vergnaud S, Cartus O, Santoro M, Fritz S, Georgieva I, Lesiv M, Carter S, Herold M, Li L, Tsendbazar N-E, Ramoino F, Arino O (2021). ESA WorldCover 10 m 2020 v100.).

7 ARTICLE II

Future flow and water temperature scenarios in an impounded drainage basin: Implications for summer flow and temperature management downstream of the dam

Scénarios futurs de débit et de température de l'eau dans un bassin versant avec retenue d'eau : implications pour la gestion du débit et de la température d'été en aval du barrage

Auteurs :

Mostafa Khorsandi[1*] (Corresponding author), Andre St-Hilaire[1], Richard Arsenault[2], Jean-Luc Martel[2], Samah Larabi[3], Markus Schnorbus[3], Francis Zwiers[3]

[1]Canadian Rivers Institute and Institut National de la Recherche Scientifique, Centre Eau Terre Environnement (INRS-ÉTÉ), 490 De la Couronne St., Québec City, Quebec, G1K 9A9, Canada

[2]Hydrology, climate and climate change laboratory, École de technologie supérieure (ÉTS), Montréal, Quebec, Canada

[3]Pacific Climate Impacts Consortium (PCIC), University of Victoria (UVIC), Victoria, British Columbia, Canada

L'article a été publié dans la revue Climatic Change Journal

Date de publication : Novembre 2023

Contribution des auteurs :

Mostafa Khorsandi: Conceptualisation, curation des données, analyse formelle, enquête, méthodologie, logiciel, validation, visualisation, rédaction - ébauche originale, rédaction - révision et édition; André St-Hilaire: Conceptualisation, acquisition de financement, administration de projet, ressources, supervision, validation, rédaction – révision & révision; Richard Arsenault: Conceptualisation, acquisition de financement, ressources, supervision, validation, rédaction – révision & édition; Jean-Luc Martel: Curation des données, visualisation, rédaction – révision & édition; Samah Larabi: Conservation des données, rédaction – révision et édition; Markus Schnorbus: Ressources, rédaction – révision et édition; Francis Zwiers: Ressources, rédaction – révision et édition.

Résumé

La température de l'eau est une variable clé affectant l'habitat des poissons dans les rivières. Le saumon sockeye (*Oncorhynchus nerka*), une espèce clé dans les écosystèmes aquatiques du nord-ouest de l'Amérique du Nord, est profondément affecté par les changements du régime thermique des rivières, et il joue un rôle essentiel dans les contextes écologiques et économiques en raison de son histoire de vie, de sa distribution étendue et de sa pêche commerciale. Dans cette étude, nous explorons les effets du changement climatique sur le régime thermique de la rivière Nechako (Colombie-Britannique, Canada), une rivière relativement grande partiellement contrôlée par le déversoir du lac Skins. Le modèle hydrologique-thermique CEQUEAU a été calibré à l'aide d'observations de débit et de température de l'eau. Le modèle a été forcé à l'aide des données de la Cinquième génération de la Réanalyse atmosphérique de l'ECMWF pour le passé et des projections météorologiques (réduites en échelle et corrigées des biais) des modèles climatiques pour les scénarios futurs. Le calage hydrologique a été réalisée pour la période 1980-2019 à partir de données provenant de deux stations hydrométriques, et le calage de la température de l'eau a été mise en œuvre à l'aide d'observations pour la période 2005-2019 provenant de huit stations de température de l'eau. Les changements de température de l'eau ont été évalués pour deux périodes futures (2040-2069 et 2070-2099) à l'aide de huit modèles climatiques du Projet d'intercomparaison des modèles couplés Phase 6 et en utilisant deux scénarios de trajectoire socio-économique partagée (4,5 et 8,5 W/m^2 d'ici 2100) pour chaque période. Les résultats montrent que les températures de l'eau au-dessus de 20°C [un seuil supérieur pour un habitat thermique adéquat pour la migration du saumon sockeye dans cette rivière] à la station de Vanderhoof augmenteront en fréquence quotidienne. Alors que la fréquence de survenue de ce phénomène est de 1 % (0 à 9 jours/été) sur la base des observations de 2005-2019, cette plage de chiffres est de 3,8 à 36 % (0 à 62 jours/été) selon l'ensemble des scénarios de changement climatique. Ces résultats montrent la diminution de la disponibilité de l'habitat pour le saumon sockeye en raison du changement climatique et l'importance de la gestion de l'eau pour résoudre ce problème.

Mots-clés

Modèle CEQUEAU, Changement climatique, Rivière Nechako, Température de la rivière.

Abstract

Water temperature is a key variable affecting fish habitat in rivers. The Sockeye salmon (*Oncorhynchus nerka*), a keystone species in north western aquatic ecosystems of North America, is profoundly affected by thermal regime changes in rivers, and it holds a pivotal role in ecological and economic contexts due to its life history, extensive distribution, and commercial fishery. In this study, we explore the effects of climate change on the thermal regime of the Nechako River (British Columbia, Canada), a relatively large river partially controlled by the Skins Lake Spillway. The CEQUEAU hydrological-thermal model is calibrated using discharge and water temperature observations. The model is forced using the Fifth generation of ECMWF Atmospheric Reanalysis data for the past and meteorological projections (downscaled and bias-corrected) from climate models for future scenarios. Hydrological calibration is completed for the 1980-2019 period using data from two hydrometric stations, and water temperature calibration is implemented using observations for 2005-2019 from eight water temperature stations. Changes in water temperature are assessed for two future periods (2040–2069 and 2070–2099) using eight Coupled Model Intercomparison Project Phase 6 climate models and using two Shared Socioeconomic Pathway scenarios (4.5 and 8.5 W/m^2 by 2100) for each period. Results show that water temperatures above 20°C [an upper threshold for adequate thermal habitat for Sockeye salmon migration in this river] at the Vanderhoof station will increase in daily frequency. While the frequency of occurrence of this phenomenon is 1% (0-9 days/summer) based on 2005-2019 observations, this number range is 3.8-36% (0-62 days/summer) according to the ensemble of climate change scenarios. These results show the decreasing habitat availability for Sockeye salmon due to climate change and the importance of water management in addressing this issue.

Key Words

CEQUEAU model, Climate change, Nechako River, River temperature.

7.1 Introduction

The aquatic systems are threatened by natural and anthropogenic pressures, including land cover change, flow regime change, and global climate change (Algera *et al.*, 2022; Bosmans *et al.*, 2022). Of all these threats, climate change may have the most prolonged and largest-scale impacts on various ecosystems, particularly in freshwater (Nash *et al.*, 2017). Habitat loss already has severe consequences for aquatic life like salmonids (Deinet *et al.*, 2020), with a 76% reduction in their global population and even more in some regions (93% in Europe) (Deinet *et al.*, 2020). It remains unclear how much of this habitat loss is due to climate change.

However, it has been shown that increased air temperature has changed North America's flow regime and water temperatures (Islam *et al.*, 2019). The river reaches located downstream of dam facilities have managed flows, which can increase (via inadequate environmental flow prescriptions) or reduce (through water releases) the impact of climate change (Ahmad *et al.*, 2021; Algera *et al.*, 2022; Fullerton *et al.*, 2022; Sheedy, 2005; Sullivan & Rounds, 2021; Xiong *et al.*, 2020). There are several ways to mitigate the impact of climate change through man-made efforts on the local scale. The analysis of dam operations as a potential mitigation method to counteract climate change impacts on juvenile salmon was conducted by Sullivan and Rounds (2021) and Stratton Garvin *et al.* (2022). They successfully utilized hydrodynamic models to demonstrate that changes in dam operations, particularly at the upstream dam, can effectively alter water temperatures released from the dam, with implications for seasonal temperature patterns and downstream river temperature variations. Climate change affects juvenile salmon through changes in water temperature, habitat availability, and food resources, influencing their growth and survival. Dam operation simulations aimed to investigate scenarios that can improve conditions for endangered anadromous fish by incorporating a temperature target.

A study conducted in the Nechako River watershed in Western Canada (Macdonald, 2019) focused on assessing the approaches to alleviate the challenges posed by reservoir operation and the resulting reduced water flows, both leading to unfavorable conditions for the Sockeye salmon population. River water temperature plays a crucial role in aquatic life, and previous studies have shown that temperature is one of the dominant factors influencing the Nechako River watershed's aquatic habitats (Macdonald *et al.*, 2007) because water temperature shapes the conditions essential for aquatic organisms' survival. It directly influences the metabolic rates of aquatic organisms, affecting their growth, reproduction, and overall physiological functions. Furthermore, temperature governs the solubility of gases in water, impacting oxygen levels vital for aquatic species' survival (Macdonald, 2019). For example, Islam *et al.* (2019) showed that summer water temperatures in the Fraser River watershed, which includes the Nechako River as a sub-watershed, rose by 1°C during 1950-2015. This rise in summer water temperature has doubled the number of days where the daily mean water temperature exceeded 20°C. Therefore, previous studies related to Sockeye salmon (*Onchorhynchus Nerka*) habitat in the Nechako River resulted in a water temperature management protocol during the summertime (Macdonald, 2019).

The Summer Temperature Management Program (STMP) water release protocol was developed between 1980 and 1983 (Macdonald, 2019) to mitigate water temperature increases downstream of the Skins Lake Spillway (SLS), the hydraulic structure to control the river flow downstream. Based on STMP, during the Sockeye salmon migration season, which

lasts from mid-July to mid-August, the SLS owner/manager (Rio Tinto, the mining company that owns the Kenny Dam and SLS, also manages the water flow from SLS) releases volumes of water to maintain mean daily water temperatures below 20°C at Finmoore town, approximately 35 km upstream of the confluence of the Nechako and Stuart rivers. Moreover, understanding the specific temperature ranges within which Sockeye salmon thrive is crucial. The known thermal tolerances for Sockeye salmon (Middleton *et al.*, 2018; Robinson *et al.*, 2015) are summarized in Table II-A1 in Appendix II-A.

In recent decades, several models have been used to assess the impacts of climate change on river temperature, particularly on systems with salmonid populations around Canada (Ahmadi-Nedushan *et al.*, 2007; Dugdale *et al.*, 2018; Kwak *et al.*, 2017b; Wilson *et al.*, 2015). Existing water temperature models can be divided into three categories: (1) deterministic models (Dugdale *et al.*, 2017a), (2) classic statistical models (Benyahya *et al.*, 2007), and (3) artificial intelligence models as a subcategory of empirical models (Zhu & Piotrowski, 2020). Since deterministic models explicitly formulate physical processes, they are often considered a preferred tool for assessing possible shifts in water flow and temperature regimes under climate change scenarios (Ouellet *et al.*, 2020).

CEQUEAU, a hydrological-thermal model, is specifically designed for forecasting water temperature at the watershed scale (Dugdale *et al.*, 2017a; Ficklin *et al.*, 2012). CEQUEAU is a flexible modeling tool that allows adding new modules (e.g., options for different evapotranspiration and snowmelt algorithms) (St-Hilaire *et al.*, 2015), and the tool can be used jointly with modern algorithms to achieve model calibration or to conduct a sensitivity analysis [e.g., Khorsandi *et al.* (2022)]. CEQUEAU is also well-adapted for simulating dam release operations and thermal modeling of rivers. For the Nechako River watershed, CEQUEAU has been used for operational flow forecasting by Rio Tinto. Ouellet-Proulx *et al.* (2017a) studied ensemble water temperature forecasting using water temperature and flow assimilation. The model's source code for the evaporative heat loss module is improved (Ouellet-Proulx *et al.*, 2019). However, deterministic modeling using CEQUEAU is still needed to assess the combined impact of flow regime change and climate change on the Nechako. Therefore, this study attempts to provide guidance on this need by simulating the impacts of climate change on the Nechako River in the context of STMP implementation with CEQUEAU. Therefore, this study aims to:

1. Calibrate the CEQUEAU model hydrologically and thermally using the multisite water flow and temperature calibration method.

2. Establish upstream boundary conditions by coupling CEQUEAU with other models that simulate reservoir temperature and those found in tributaries emptying in the reservoir (i.e., VIC-GL, RBO, and CE-QUAL-W2 models).
3. Project future water temperature scenarios using downscaled, bias-corrected meteorological data, upstream model outputs as boundary conditions, and the calibrated CEQUEAU model.
4. Analyze these projected water temperature scenarios, particularly focusing on compliance with the mandatory 20°C threshold at the Vanderhoof station.

7.2 Methods

7.2.1 Study area

The Nechako River watershed is located in the central part of British Columbia, Canada. The area under study is a sub-watershed of Nechako River watershed with a 12,805 km^2 drainage area (Figure 7.1). Downstream of Ootsa Lake (Figure 7.1, the large lake immediately upstream of SLS), the flow of the river is fully regulated by the SLS between the Nechako Reservoir and its confluence with the Nautley River, after which the flows of the Nechako and Nautley rivers combine and continue past the town of Vanderhoof and eventually into the Fraser River. Three hydrometric stations measure the discharge (Environment and Climate Change Canada, https://www.canada.ca/en/environment-climate-change/, data access: May 2022). Those are at the SLS, the Nautley River (upstream of the confluence with the Nechako River), and the town of Vanderhoof (see Figure 7.1). The SLS flows and temperatures provide the boundary conditions to the hydrological and thermal modules and are not used as target stations for hydrological calibration. Eight stations measure water temperature (Rio Tinto, data access: May 2022) in the watershed (Figure 7.1). The study area and monitoring stations are explained in detail in Ouellet-Proulx et al. (2017a) and Khorsandi et al. (2022).

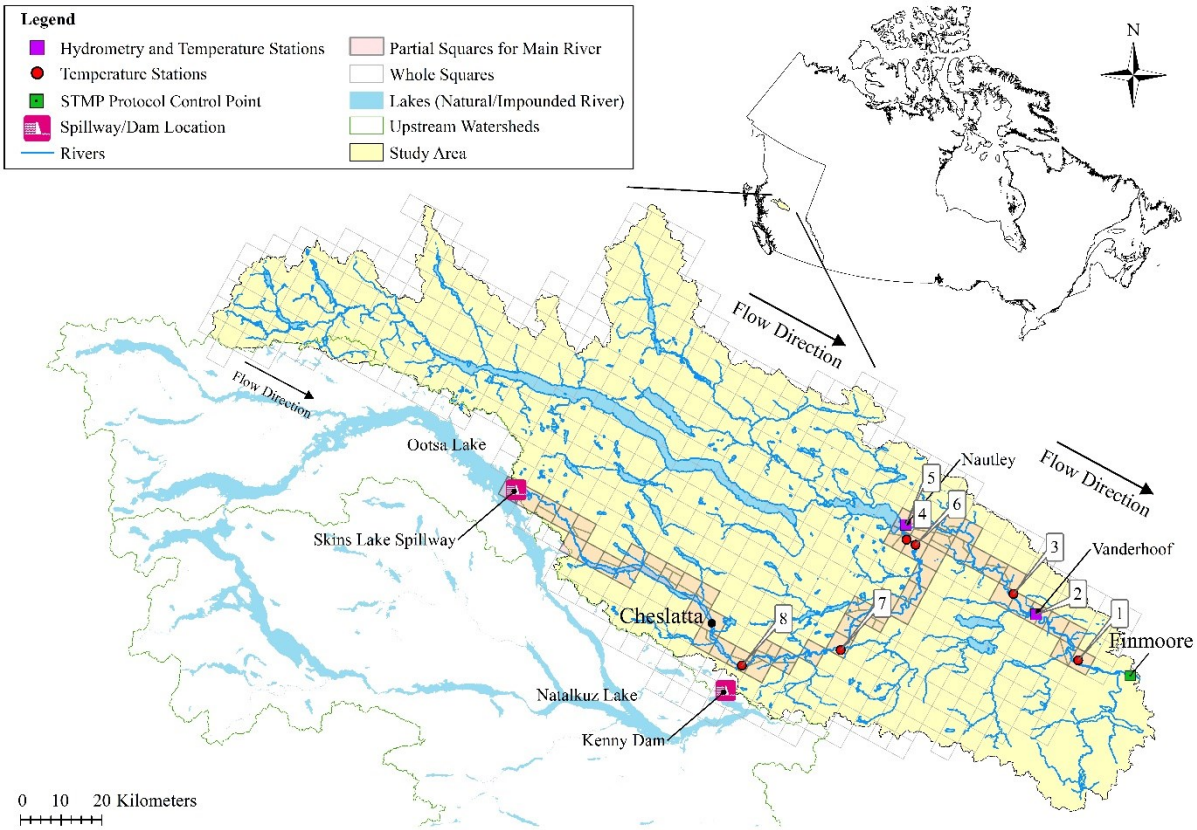

Figure 7.1 **The Nechako River watershed is located in the central part of British Columbia, Canada, with a 45,000 km^2 drainage area. Downstream of Ootsa Lake (the large lake immediately upstream of SLS), the flow of the river is fully regulated by the SLS between the Nechako Reservoir and its confluence with the Nautley River, after which the flows of the Nechako and Nautley rivers combine and continue past the town of Vanderhoof and eventually into the Fraser River. Three hydrometric stations measure the discharge (Environment and Climate Change Canada, https://www.canada.ca/en/environment-climate-change/, data access: May 2022). Those are at the SLS, the Nautley River (upstream of the confluence with the Nechako River), and the town of Vanderhoof. The SLS flows and temperatures provide the boundary conditions to the hydrological and thermal modules and are not used as target stations for hydrological calibration. Eight stations measure water temperature (Rio Tinto, data access: May 2022) in the watershed. The study area and monitoring stations are explained in detail in Ouellet-Proulx et al. (2017a) and Khorsandi et al. (2022).**

7.2.2 CEQUEAU model

CEQUEAU is a hydrological-thermal model designed explicitly for hydrological and surface water temperature modeling (Morin & Couillard, 1990; St-Hilaire *et al.*, 2015). The model has been extensively tested in several case studies (Dugdale *et al.*, 2018; Dugdale *et al.*, 2017b; Fniguire *et al.*, 2022; Kwak *et al.*, 2017a; Kwak *et al.*, 2017b; Ouellet-Proulx *et al.*, 2017a; Ouellet-Proulx *et al.*, 2017b).

Land cover and topography are required physiographic input data for the CEQUEAU model (Dugdale *et al.*, 2017b). We used the most up-to-date global land cover data, with a 10 m spatial resolution provided by Environmental Systems Research Institute (ESRI) and the

European Space Agency (ESA)(Karra *et al.*, 2021; Zanaga *et al.*, 2021). We used the National Aeronautics and Space Administration (NASA) SRTM Digital Elevation Model (SDEM) with 30 m spatial resolution (Farr *et al.*, 2007) to calculate elevations in the CEQUEAU's input structure.

The CEQUEAU model conceptualizes a drainage basin as an interconnected network of hydrological response units called partial squares (CP; based on the French acronym in the CEQUEAU manual), delineated as sub-components of square grid cells (Figure 7.2a). For each CP, the hydrological module calculates a simplified hydrological budget using a production function (PF) that simulates water routing into the surface runoff, interflow, and groundwater. Then a hydrological transfer function (TF) applies a routing scheme on the available surface water to calculate the water volume routed to the downstream CP (Figure 7.2d). The PF uses precipitation and air temperature (minimum and maximum) as meteorological inputs (Figure 7.2b). The hydrological module includes 26 global (i.e., one value for all CPs) parameters and then produces the output, i.e., simulated discharge (Figure 7.2d).

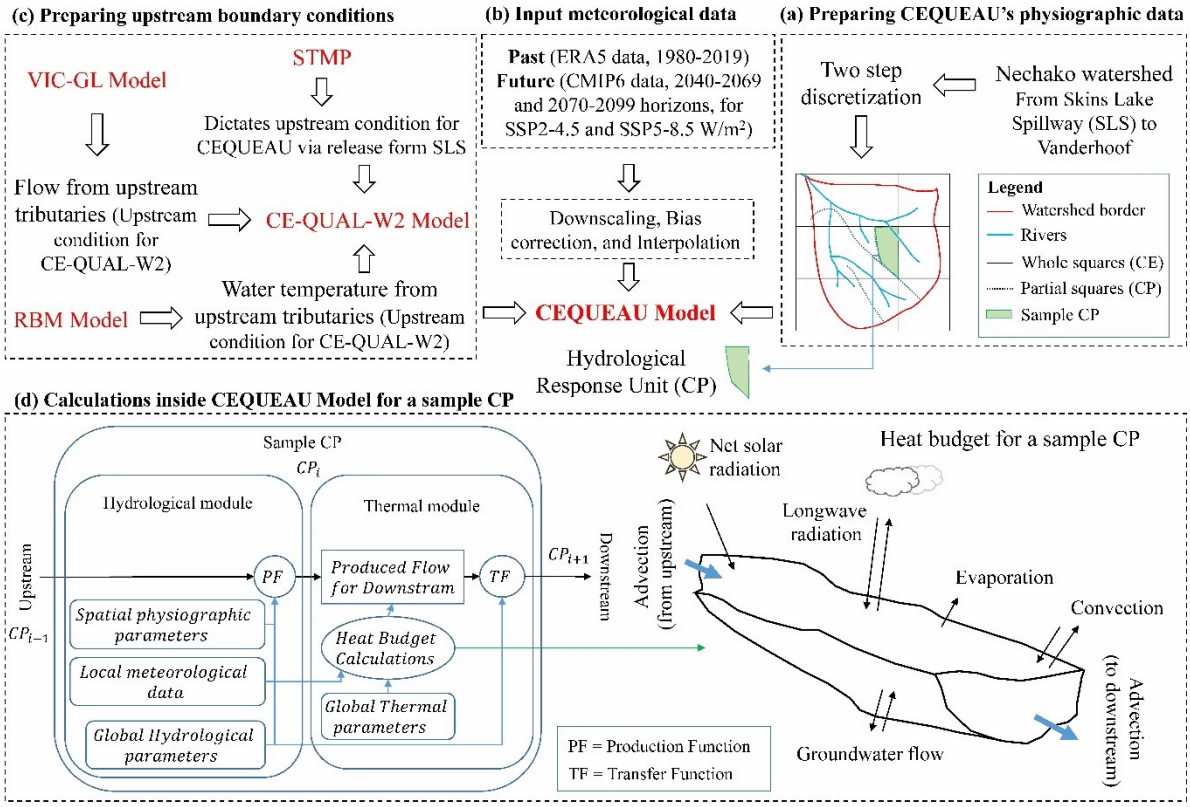

Figure 7.2 **Schematic representation of steps, concepts, and models to set up the CEQUEAU model in this study, including (a) physiographic data for the watershed, whole squares, and partial squares; (b) input meteorological data (precipitation, min and max air temperature, vapor pressure, cloud cover, net solar shortwave radiation), (c) the models incorporated to prepare upstream boundary condition for CEQUEAU; (d) structure of a sample partial square (CP) which is a hydrological response unit for which both hydrological and thermal budgets are computed. The core of the hydrological module is the Production Function (PF) and Transfer Function (TF), which calculates available water inside each CP and subsequent routing downstream. The thermal module calculates the heat budget using available water inside each CP.**

7.2.3 Modeling upstream boundary conditions

Flow and temperature at the SLS are the boundary conditions for the CEQUEAU model. These data are observed values at SLS (Environment and Climate Change Canada, https://www.canada.ca/en/environment-climate-change/, data access: May 2022) for the calibration period. In terms of temperature, one sensor is installed immediately downstream of SLS. The data for this sensor has been available since 2017. Therefore, for the calibration period, similar to previous studies on the Nechako using the CEQUEAU model (Ouellet-Proulx, 2018; Ouellet-Proulx *et al.*, 2017a; Ouellet-Proulx *et al.*, 2017b), we calculated the mean water temperature for each day of the year (DOY) using the observed data from 2017 to 2021.

A coupled hydrologic model simulates upstream boundary conditions for flow and temperature at SLS for future horizons (2040-2069 and 2070-2099): the Variable Infiltration Capacity (VIC-GL) (Liang *et al.*, 1996; Schnorbus, 2018), the River Basin Model (RBM) (Larabi *et al.*, 2022), the STMP reservoir operation program, and the CE-QUAL-W2 hydrodynamic model (Cole & Wells, 2006) (Figure 7.2). The coupled modeling platform aims to simulate the governing processes of water flow and temperature feeding the reservoir and reservoir hydrodynamics. The modeling platform allows for consideration of changes in timing and volume of water availability to simulate reservoir thermal stratification and temperature of water released at SLS.

The VIC-GL is an upgraded version of the VIC model, a spatially distributed land surface model that accounts for glacier processes (Schnorbus, 2018). The VIC-GL model was implemented at the upstream area of the Nechako Reservoir. The stream temperature model, RBM, is a gridded physically-based model that uses a one-dimensional mixed Eulerian-Lagrangian approach to simulate water temperature based on local air-water surface heat exchange and advected heat flux from upstream. VIC-GL was coupled with RBM to simulate discharge and water temperature at the main tributaries feeding the Nechako Reservoir (Larabi *et al.*, 2022). Both models were calibrated against observed discharge and water temperature at the six main tributaries of the Nechako Reservoir as identified by Canada Water Survey stations. They used input meteorological data for 1945-2018 at a 3-hour timestep for which both observed water flow and water temperature data are available for several periods depending on the station (Larabi *et al.*, 2022). These models (VIC-GL, RBM, and CE-QUAL-W2) did not need upstream conditions since each model provided upstream conditions for the next one. Therefore, the calibration and validation were performed using partial time series data availability and daily time steps for water flow and temperatures.

CE-QUAL-W2 is a mechanistic hydrodynamic model for water quality modeling (Cole & Wells, 2006). This model uses a two-dimensional scheme to differentiate water bodies along the river (longitudinal) and depth (vertical). This scheme makes this model suitable for studying water flow and quality studies in large water bodies like dam reservoirs (Afshar *et al.*, 2011; Kim & Kim, 2006). The model assumes water is laterally well mixed but can be vertically stratified. CE-QUAL-W2 can model water velocity and flow at different time scales with the hydraulic sub-model. Also, these variables are input for the water quality sub-model, which simulates temperature, dissolved oxygen, and several other variables required to study aquatic life. The model requires meteorological data as well as boundary conditions of inflows and outflows. Using VIC-GL and RBM to provide inflow boundary conditions and historical powerhouse intake and water release at SLS, the CE-QUAL-W2 model was first calibrated against historical reservoir water elevation for the period spanning 1986-2017. Then, the model was

calibrated against water temperature profiles at Kenney Dam (See Figure 7.1, in the Nechako River) and Natalkuz Lake (Upstream of Kenny Dam, Figure 7.1) for the summer of 1994, as well as outlet water temperature at SLS for the summer of 2016-2017 (See Figure 7.1). The deployment and calibration of the integrated modeling platform are discussed in detail by Larabi *et al.* (2022).

Implementing the Summer Temperature Management Program (STMP) represents a crucial facet of water flow regulation at the Nechako Reservoir (see Figure 7.2c). As a response to the high influence of reservoir regulation on water quantity and quality in the Nechako River, the STMP was introduced in 1983 with the primary objective of ameliorating conditions for Sockeye salmon migration. The STMP focuses on mitigating elevated water temperatures during the critical migration period from July 20 to August 20 at Finmoore. This program involves augmenting water releases at the SLS during the migration period, effectively curbing the frequency of water temperatures exceeding 20°C. The daily mean water release is approximately 32 m^3/s at SLS during fall and winter to support Sockeye salmon. The water release is increased during summer to a maximum limit (approximately 450 m^3/s) in response to high temperatures. By orchestrating these controlled releases, the STMP plays a pivotal role in preserving the ecological integrity of the Nechako River and sustaining the migratory patterns of vital salmon species (Larabi *et al.*, 2022; Macdonald, 2019).

To simulate future scenarios of temperatures and flows at SLS, the VIC-GL/RBM/CE-QUAL-W2 combination of models is forced with future climate model outputs using two Shared Socioeconomic Pathways (SSPs). These future hydrologic scenarios are then used as input to the reservoir operation model (e.g., STMP) provided by Rio Tinto to simulate associated scenarios of powerhouse intake and SLS discharge (Figure 7.2b,c).

7.2.4 Fifth generation of ECMWF atmospheric reanalysis data

There is a need for complete meteorological data to calibrate and validate the hydrological model for the historical period as the baseline to be compared with future scenarios. However, there is no complete observed meteorological dataset for the basin in the past that can be used as a reference dataset. Therefore, an alternative is to use climate reanalysis products, as Gatien *et al.* (2022) suggested.

In addition to precipitation and air temperature, the CEQUEAU water temperature module (Figure 7.2d) requires wind speed, water vapor pressure, and net solar shortwave radiation. These meteorological variables are provided by the European Center for Medium-Range Weather Forecasting (ECMWF) through their European Reanalysis 5th generation (ERA5) (Hersbach *et al.*, 2020a), with the exception of vapor pressure which was calculated using

ERA5 dew point temperature at 2 m height and with Teten's equation (Monteith & Unsworth, 2013; Murray, 1967). These gridded input data (with 31 × 31 km resolution) were interpolated to all basin whole squares using the built-in CEQUEAU interpolator based on the nearest neighbors approach (Ouellet-Proulx *et al.*, 2019). ERA5 and climate models had three hours of temporal time steps. After downscaling and bias correction for climate models, we converted these data to daily time steps for running the CEQUEAU model.

7.2.5 Model calibration and implementation

This study uses multisite model calibration, using the maximum available information to provide the best set of parameters for the whole watershed (Arsenault *et al.*, 2018; Bérubé *et al.*, 2022; Shen *et al.*, 2022). This study uses data from all available hydrometric (2) and water temperature stations (8). First, the hydrological module with 26 parameters was calibrated using the data from two hydrometric stations for June-September, which is the high-temperature period of the year (Table II-A2 in Appendix II-A). Then, using the calibrated hydrologic module outputs, the thermal module with eight parameters was calibrated using the data from eight water temperature monitoring stations for the same period.

7.2.5.1 Covariance matrix adaptation evolution strategy calibration algorithm

This study uses the Covariance Matrix Adaptation Evolution Strategy [CMA-ES; Hansen and Ostermeier (1996)] to find the optimal set of parameters for both hydrological and thermal modules for CEQUEAU during model calibration. CMA-ES is an evolutionary algorithm developed as a global optimization method (Hansen, 2023). This algorithm is frequently used for hydrological model calibration (Elshall *et al.*, 2015; Smaoui *et al.*, 2018; Yu *et al.*, 2012; Zouhri *et al.*, 2021). Arsenault *et al.* (2014) showed the superiority of CMA-ES in finding global optima and convergence speed compared to nine other well-known optimization algorithms used to calibrate hydrological models. Khorsandi *et al.* (2022) showed the efficiency of this method for thermal calibration of the CEQUEAU model for the Nechako watershed. This method can be summarized in four steps, which are explained in detail by Hansen (2023) and Khorsandi *et al.* (2022).

7.2.5.2 Hydrological model calibration

The objective function for the optimization algorithm is the Kling-Gupta Efficiency (KGE) (Gupta *et al.*, 2009) coefficient, which is in Equations (7.1)-(7.5). The objective function in the calibration process shows the goodness of fit between simulated and observed data. Since every performance metric has its own strengths and limitations, presenting only one

performance metric may be biased or may not accurately reflect the calibration's success. In addition to KGE, the Bias and Nash-Sutcliffe Efficiency [NSE; Nash and Sutcliffe (1970)] performance metrics were also calculated as follows:

$$KGE_j = 1 - ED_j \tag{7.1}$$

$$ED_j = \sqrt{(r_j - 1)^2 + (\alpha_j - 1)^2 + (\beta_j - 1)^2} \tag{7.2}$$

$$r_j = \frac{\sum_{i=1}^{N}\left(O_{i,j} - \overline{O}_j\right)\left(S_{i,j} - \overline{S}_j\right)}{\sqrt{\sum_{i=1}^{N}\left(O_{i,j} - \overline{O}_j\right)^2}\sqrt{\sum_{i=1}^{N}\left(S_{i,j} - \overline{S}_j\right)^2}} \tag{7.3}$$

$$\alpha_j = \frac{\sigma_{S_j}}{\sigma_{O_j}} \tag{7.4}$$

$$\beta_j = \frac{\overline{S}_j}{\overline{O}_j} \tag{7.5}$$

$$Bias_j = \overline{O}_j - \overline{S}_j \tag{7.6}$$

$$NSE_j = 1 - \frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{\sum_{i=1}^{N}\left(O_{i,j} - \overline{O}_j\right)^2} \tag{7.7}$$

where KGE_j, $Bias_j$ and NSE_j are the performance metrics for the jth station; $O_{i,j}$ and $S_{i,j}$ respectively are observed and simulated values for the jth station at the ith time step; N is the number of time steps (measurements); r_j is the pearson correlation coefficient between observed and simulated values; α_j is the ratio of the standard deviation of simulated values to the standard deviation of observed values; β_j is the ratio of the mean of simulated values to the mean of observed values; $\sigma_{S\,j}$ and $\sigma_{O\,j}$ are the standard deviation of simulations and observations for the jth station, and $\overline{S}_j$ and $\overline{O}_j$ are the mean of simulated and observed values. As previously recommended by Arsenault *et al.* (2018) and Shen *et al.* (2022), all the available data and stations were used for model calibration, which means that no data was reserved for validation.

7.2.5.3 Thermal model calibration

The eight thermal model parameters (Table II-A3 in Appendix II-A) - were adjusted using Root Mean Square Error (RMSE) as the objective function and the CMA-ES algorithm for optimization. The RMSE performance metric was used as recommended in the literature (Ouellet-Proulx *et al.*, 2017a; Ouellet-Proulx *et al.*, 2019).

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{N}} \tag{7.8}$$

where $RMSE_j$ is the performance metric for the jth station.

As with the hydrological calibration, all stations and all observed data were used for thermal model calibration. Both hydrological and thermal calibrations were completed in a two-step process. First, lower and upper boundaries for the model parameters were examined manually to ensure that the simulations stayed within realistic limits when the parameter had a physical meaning. Following this manual calibration, the CMA-ES optimization algorithm was used to refine the model parameter values of the hydrological and thermal modules. These steps are explained in detail by Khorsandi *et al.* (2022).

7.2.6 Future model projections

Following calibration, the calibrated CEQUEAU model was used to simulate water temperature over the Nechako watershed for two future periods, 2040-2070 and 2070-2100, using eight CMIP6's General Circulation Models (GCM) (Table II-A4 in Appendix II-A). The 4.5 and the 8.5 W/m^2 radiative forcing scenarios show the range of possible changes for the two future periods compared to the baseline period (for 1980-2019, using the ERA5 dataset). We selected the 4.5 and the 8.5 W/m^2 scenarios based on Shared Socioeconomic Pathways (SSP) terminology agreed upon in CMIP6 (Table II-A5 in Appendix II-A), with 8.5 W/m^2 being the more pessimistic scenario. The water temperature simulations resulted in performance metrics for the river's thermal regime. From all CMIP6 models, these eight models were selected to conduct this study because they had all the required variables at sub-daily time steps. We used the sub-daily time steps for bias correction, similar to Gatien et al. (Personal communication) for inter-model comparison studies.

Each global climate model using SSP forcings provided all meteorological data to run the model, which were bias-corrected using the N-dimensional Multivariate Bias Correction algorithm (MBCn) algorithm (Cannon, 2018) with the ERA5 data as the reference dataset for two radiative forcing scenarios (4.5 and 8.5 W/m^2). Eight climate models were used, as presented in Table II-A4 in Appendix II-A. The bias correction at the sub-daily timestep better represents the meteorological data's diurnal cycle and improves the daily means (Gatien *et al.*, 2022).

The MBCn algorithm is a novel approach in climatology that addresses limitations in traditional bias correction methods. While many existing algorithms focus on univariate time series and disregard the interdependencies between different variables, MBCn introduces a multivariate

perspective. Inspired by an image processing technique, it leverages the N-dimensional probability density function transform to correct climate model projections of several climate variables. MBCn generalizes quantile mapping, preserving the entire observed continuous multivariate distribution in the corresponding climate model distribution. This adapted technique accurately maintains the quantile changes between historical and projection periods (Cannon, 2018).

Because the CEQUEAU model simulates water temperature at daily time steps, the sub-daily variables were aggregated to generate daily meteorological time series (by averaging for most variables and summing for precipitation). Finally, the daily meteorological projections were forced into CEQUEAU, and the model simulated future discharge and water temperature projections for the two periods mentioned above. Each GCM has a different gridded network, which was interpolated for the CEQUEAU grid points using the nearest neighbors approach (Ouellet-Proulx *et al.*, 2019).

7.2.7 Duration curve method

All inputs and outputs for the CEQUEAU model have daily time steps in this study. For the simulation outputs, temperature duration curves (TDC) were used to show and analyze the daily water temperature simulations for the past and future periods. TDC plots the probability of the exceedance of the observed value from several thresholds in a continuous manner (Karakoyun *et al.*, 2018). This study identified the probability of exceedance of the critical environmental threshold of 20°C [an upper threshold for adequate thermal habitat for Sockeye salmon migration in this river at Finmoore; Macdonald (2019)] on TDC curves. The TDC curves in this study show the probability of exceedance for water temperatures higher than 5°C. The temperatures above the 5°C threshold was selected because the CEQUEAU model is not designed to model low (close to zero) water temperatures (e.g., there is no ice-formation algorithm in CEQUEAU).

7.3 Results

7.3.1 Hydrological module calibration and parameters

Figure 7.3 shows the simulated and observed discharge values for Vanderhoof and Nautley stations during the calibration period.

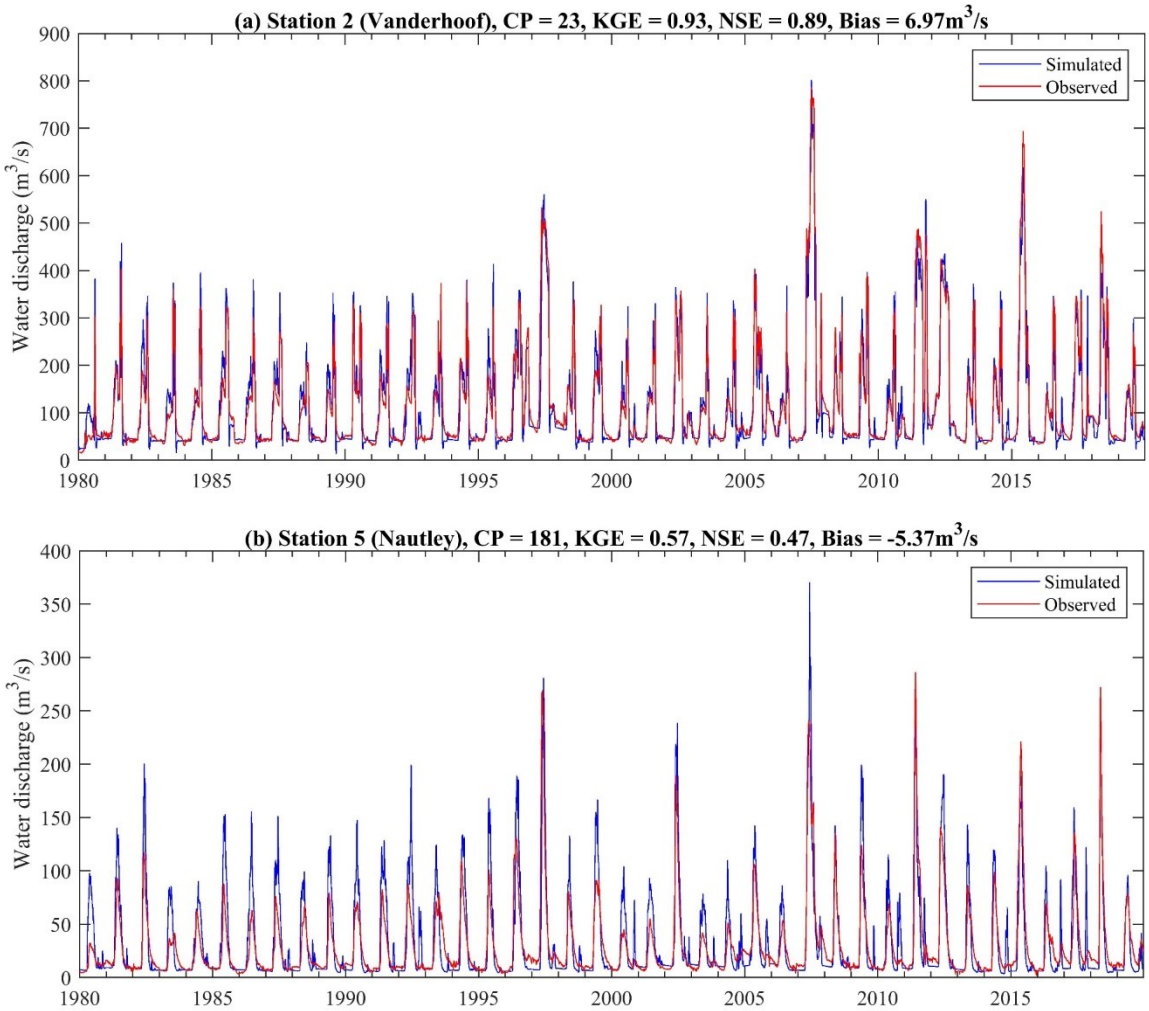

Figure 7.3 **Simulated and observed flows using the calibrated model for Vanderhoof and Nautley stations for the June-September period**

Performance indicators were excellent for the Vanderhoof station (KGE = 0.93, NSE = 0.89) (Figure 7.3a), and the KGE value was acceptable for the Nautley station (KGE = 0.57, NSE = 0.47) (Figure 7.3b). Table II-A2 in Appendix II-A shows the 26 calibrated hydrological parameters using the multisite approach using these two observation stations. However, the simulated values at the Nautley station show overestimations for high flows. This Bias for high flows is likely due to the presence of two large lakes upstream of the station. The current version of the CEQUEAU model conceptualizes the lake storage effect by using a single reservoir and a modified transfer coefficient. However, a single calibrated value for CVMAR (Lakes and marshes drainage coefficient in a CP) and HMAR (Lakes and marshes drainage threshold in a CP), the two parameters associated with lake water storage, is likely not well-adapted to the lake cascade configuration in the Nautley sub-watershed (See Table II-A2 in Appendix II-A). On the other hand, for the Nautley station, high flows are not the priority in this study. High flows at the Nautley station are inconsequential to this study since they occur in

May and June (Khorsandi *et al.*, 2022) instead of the warm July-August period, for which the simulations for the Nautley station show acceptable performance metrics (KGE = 0.61, NSE = 0.29) with unbiased flow estimations (Bias = 1.19 m^3/s, relative Bias = 0.03).

7.3.2 Thermal model calibration

Table II-A3 in Appendix II-A shows the thermal module parameter values using eight stations and a multisite approach for calibration. The simulated time series versus observed values are shown in Figure 7.4 for eight stations using the calibrated values for thermal calibration.

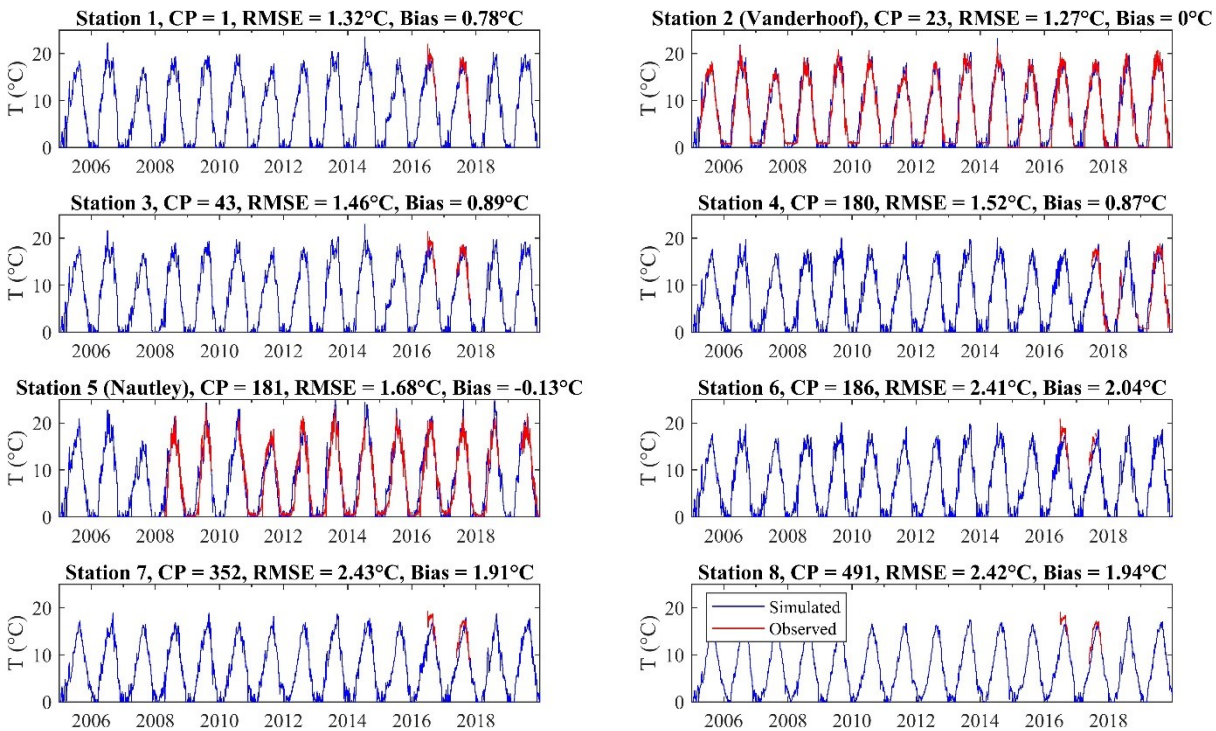

Figure 7.4 **The simulated (blue) and observed (red) water temperature time series for each station using the multisite calibration approach for the June-September period**

For the stations with more extended observation periods and without water bodies upstream, the CEQUEAU model can provide low Bias simulations with RMSE < 2°C and Bias < 0.8°C (Stations 1, 2, 3, 4, and 5). However, three of the eight stations (Stations 6, 7, and 8) show Bias performance metrics higher than 1.5°C and RMSEs greater than 2°C. Again, this Bias is likely due to the presence of large water bodies upstream of these stations. Large water bodies increase the contact surface, which, during the summer period, has a warming effect on surface water outflow. The current version of CEQUEAU does not consider this lake effect, which may explain the underestimation of simulated water temperature. The impact of large lakes on the CEQUEAU temperature simulations is further discussed by Khorsandi *et al.*

(2022). The Vanderhoof station (Station 2) is of higher importance because it is the closest station with longest observation period to the location identified in the STMP (Finmoore), where the mean daily water temperature must remain ≤ 20°C during the Sockeye salmon migration.

The simulated results for the Vanderhoof station for the baseline period show RMSE = 1.27°C and Bias ≈ 0°C, which are acceptable results when using a hydrological model at the watershed scale. Although the simulated values are generally unbiased (Figure 7.4), there is a slight underestimation of the yearly number of days above 20°C (Figure 7.5).

The mean simulated daily water temperatures matched the observed values for the July-August period for each year (Figure 7.5a). These simulated values use ERA5 data (one dataset) for the 1980-2019 baseline period, and each boxplot shows the simulated water temperature range in a year. The number of days above 20°C for both observed and simulated data is not perfectly matched and the error (observed – simulated) in the number of days above 20°C for the simulated time series ranges from -5 to +9 days (Figure 7.5b). Hence, there are summers during which the model overestimates the number of days above 20°C; and there are summers during which the model underestimates the number of days above 20°C. These differences in the simulated versus the observed number of days should be considered when interpreting climate change simulations. The observed data show that the maximum number of days above 20°C that occurred in a year is nine days (during the July-August of 2019 in Figure 7.5b). Simulations indicate that the highest historical number of exceedances of 20°C is nine days (2016; Figure 7.5b). The STMP aims to regulate elevated water temperatures during Sockeye salmon migration in the Nechako River by manipulating the timing and volume of reservoir water releases from SLS (Bond, 2017). The management plan focuses on maintaining mean daily water temperatures at or below 20°C in the Nechako River at Finmoore (30 km downstream of Vanderhoof), upstream of the Stuart River confluence, during the critical migration period from July 20 to August 20. Cooler water temperatures are essential for the survival of Sockeye salmon during its migration. The program employs a comprehensive approach involving field data collection, weather forecasts, temperature predictions, and flow release decisions. By analyzing trends in observed and predicted water temperatures, Rio Tinto determines when to increase or decrease the release of water from SLS, aiming to achieve optimal conditions for salmon migration. This protocol enhances the resilience of the Nechako River ecosystem and contributes to the conservation of Sockeye salmon populations. The program also employs a decision protocol to adjust SLS releases between 14.2-453 m^3/s based on observed and forecasted water temperature trends, ensuring that the water temperature remains below the

critical threshold for salmon survival. Additionally, the release of water from SLS is carefully managed to maintain flow below Cheslatta Falls (Figure 7.1) between 170 m^3/s and 283 m^3/s.

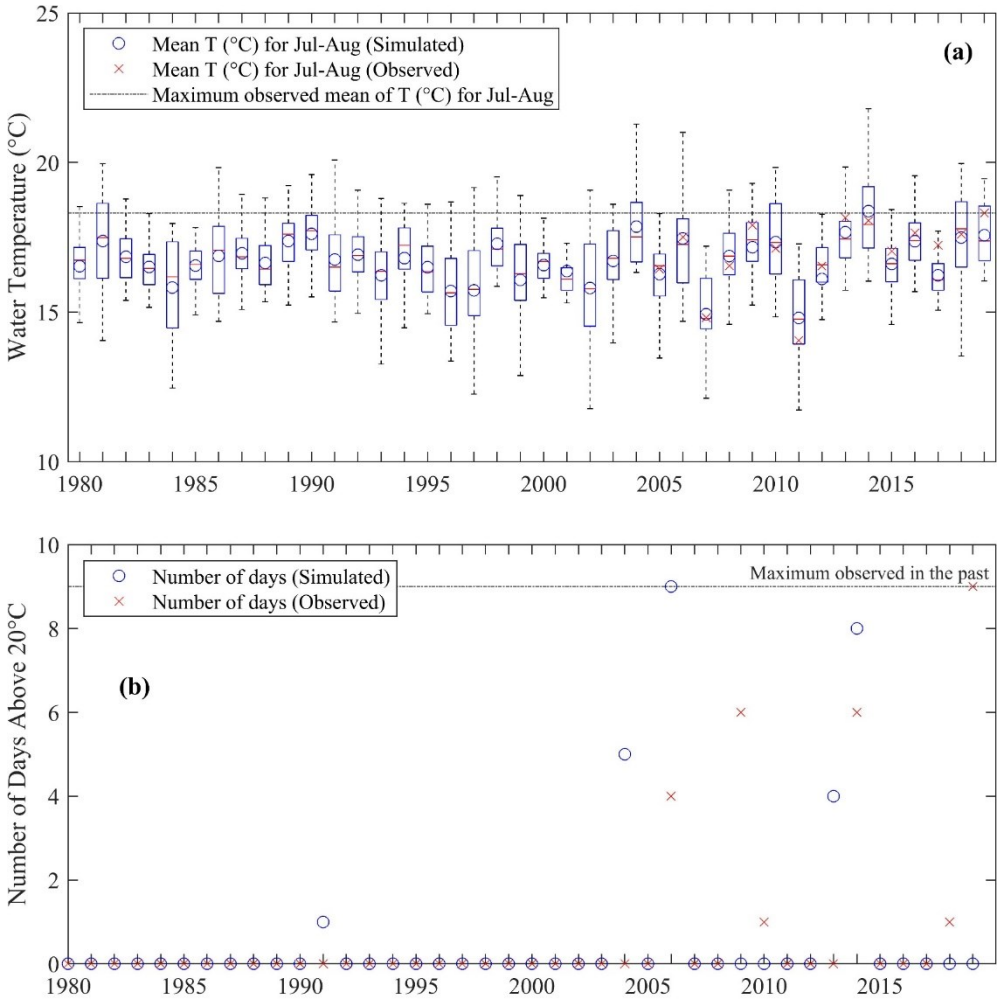

Figure 7.5 **(a) The yearly range [for July-August only] of simulated water temperature variability is shown using boxplots for the Vanderhoof station's baseline period 1980-2019. The red line represents the median value, and the blue circles show the mean simulated value. The upper and lower limits of the boxes show the 25th and 75th percentiles, respectively, and the whiskers show the most extreme data. The mean annual temperature of the observed data is shown with the red crosses (x) (2005-2019 only). (b) Red crosses and blue circles show the annual number of days on which the mean daily temperature exceeds the critical limit of 20°C for simulated and observed time series, respectively. The horizontal dotted line shows the maximum number of observed days above 20°C in a year, equal to 9 days**

The TDCs for observations and simulated values show the probability of exceedance for water temperatures higher than 5°C since the CEQUEAU model is not designed to model low water temperatures (Figure 7.6a). The timing of interannual mean values for each DOY from 1980 to 2019 for simulated match observed values during the warm months (from June to

September) (Figure 7.6b). The simulated values mostly follow the same timing pattern as the observed values (Figure 7.6b).

Although simulated temperatures lag a few days behind observed ones, the observed values fall within the lower and upper boundaries of simulated water temperatures. Figure 7.6b also shows that in long-term management through STMP implementation (adjusting SLS releases between 14.2-453 m^3/s at SLS from July 20 to August 20), STMP controls the water temperature at the Vanderhoof and the DOY mean during the warm season is below the 20°C threshold, as confirmed by observed water temperature values (Figure 7.6b).

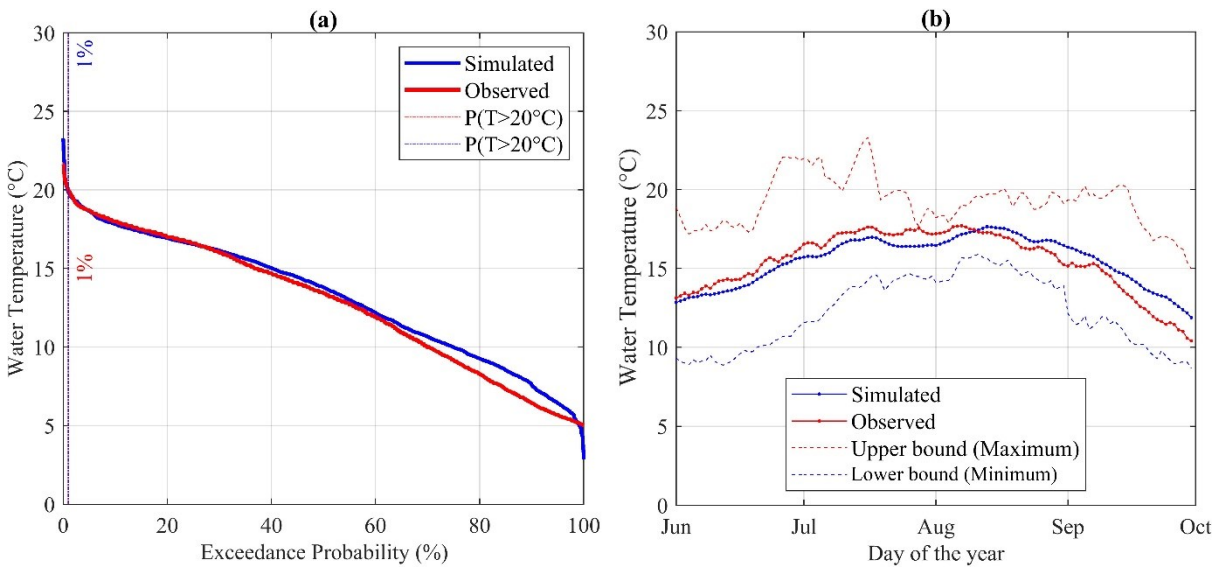

Figure 7.6 **(a) Duration curve of baseline (1980-2019) simulated data and observed water temperatures at the Vanderhoof station (b) Water temperature timing for the warm season (June-September) for observed and simulated time series together with the lower and upper boundary of simulated values**

For the Vanderhoof station on a short-term daily basis, the CEQUEAU model can accurately reproduce high-temperature days (Figure 7.4; for the Vanderhoof station in higher resolution, see Figure II-A1 in Appendix II-A). This simulation skill is visible over the long term through DOY mean values (Figure 7.6b). Therefore, the CEQUEAU simulated values for future climate change scenarios can provide reliable insight into the timing and extent of changes in high water temperatures.

7.3.3 Future projections

The selected climate scenarios data were used as inputs to the CEQUEAU model (Figure 7.2b). In addition, simulated water flow and water temperature at the SLS using CE-QUAL-

W2 were used as boundary conditions for the CEQUEAU model for future scenarios. Figures II-B1 to II-B5 in Appendix II-B show the simulated temperature results for the Vanderhoof station for the SSP2-4.5 2040-2069, SSP2-4.5 2070-2099, SSP5-8.5 2040-2069, and SSP5-8.5 2070-2099 horizons, respectively.

The Vanderhoof station shows an increase in summer water temperature compared to the baseline period (Figure 7.6a and Figure 7.7; Figure 7.6a shows TDC for the baseline, while Figure 7.7 provides TDC at the Vanderhoof station for all climate change scenarios). Comparing the water temperature increase from daily simulations for future horizons (Figures II-B1 to II-B5 in Appendix II-B) to the baseline (Figure 7.4 and Figure II-A1 in Appendix II-A) reveals the extent of change.

The TDCs for all eight climate models for all four climate change scenarios show a higher probability of exceedance for 20°C compared to the baseline period (Figure 7.7). The increasing probability of occurrence can be seen by moving from SSP4.5 to SSP8.5 and by going further in the time horizon. The exceedance probability starts from 3.8% for SSP2-4.5 (2040-2069 horizon) as the lowest value using MPI-ESM1-2-HR model data, and the highest exceedance value is 36% for SSP5-8.5 (2070-2099 horizon) using CMCC-ESM2 model data.

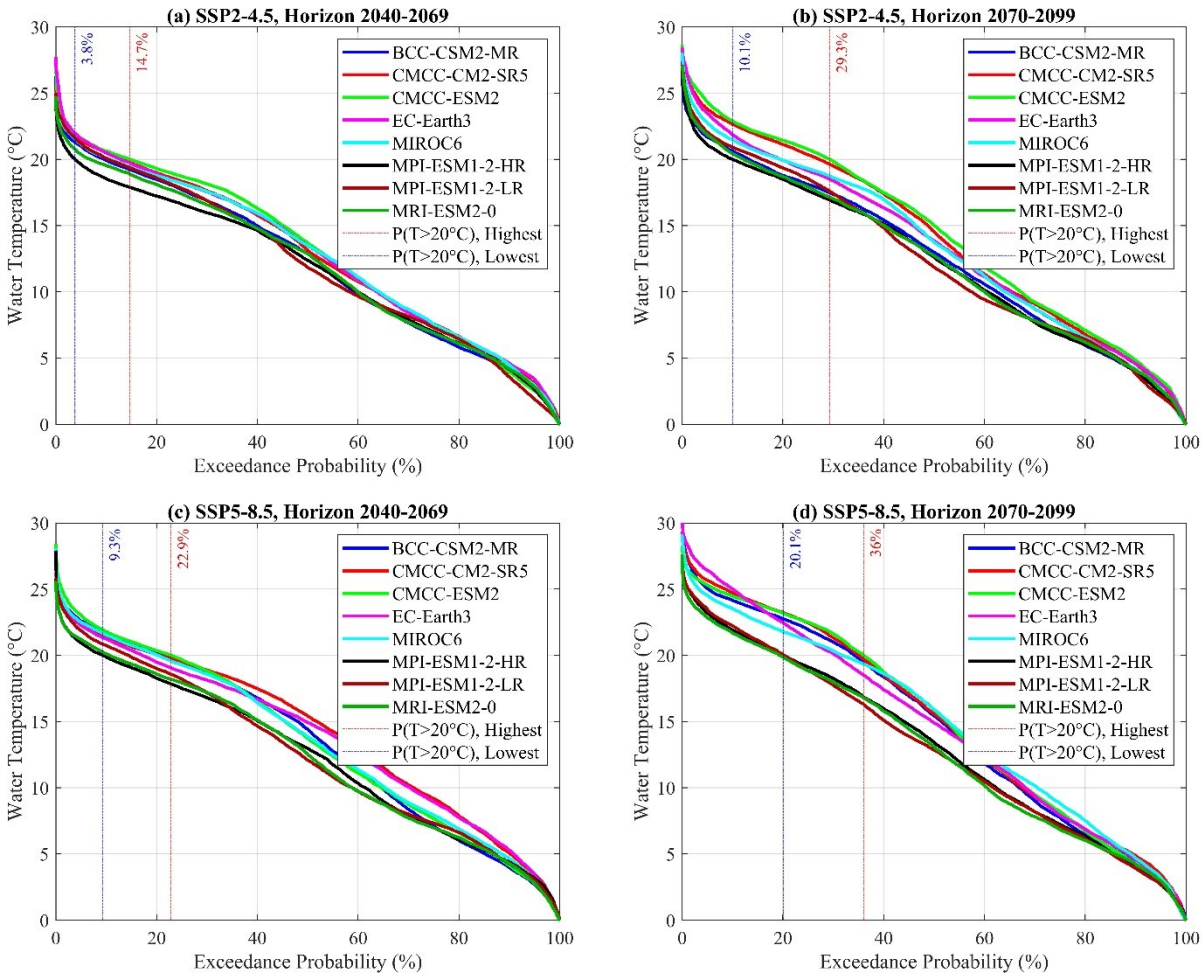

Figure 7.7 **Duration curve of simulated data using eight climate models for SSPs 4.5 and 8.5 W/m^2 and for the 2040-2069 and 2070-2099 time horizons at the Vanderhoof station. The models with the lowest and highest exceedance probabilities were used to calculate the minimum and maximum exceedance probabilities depicted on the graphs.**

The comparison of timing of high temperatures for future time horizons (Figure 7.8) compared to the past (Figure 7.6b) shows that August remains the warmest month, but September is warming more than July. The crucial time of year for warming is mid-July to mid-August, corresponding to the STMP period (Figure 7.6b). This crucial warming period will become longer and warmer. The future warm summers are expected to start in early July and last until early September. Considering the STMP fixed period (July 20 to August 20), we calculate the onset and end temperature for STMP using the interannual mean data for the baseline period. This onset-end threshold equals 17°C (by using Figure 7.6b for July 20 to August 20 period). Future scenarios show that June 8 to June 27 is the period for the future 17°C onset and September 12 to 25 is the end period for the 17°C threshold (Table 7.1), meaning more prolonged periods of warm days above these thresholds (Figure 7.8).

Table 7.1 **The onset and end dates for the 17°C thermal threshold (Figure 7.6b) based on the STMP period (July 20 to August 20)**

Scenario	Onset date (for 17°C)	End date (for 17°C)	Duration (days)
Baseline-ERA5, 1980-2019	July 20	August 20	32
SSP2-4.5, 2040-2069	June 27	September 12	78
SSP2-4.5, 2070-2099	June 16	September 14	91
SSP5-8.5, 2040-2069	June 16	September 16	93
SSP5-8.5, 2070-2099	June 8	September 25	110

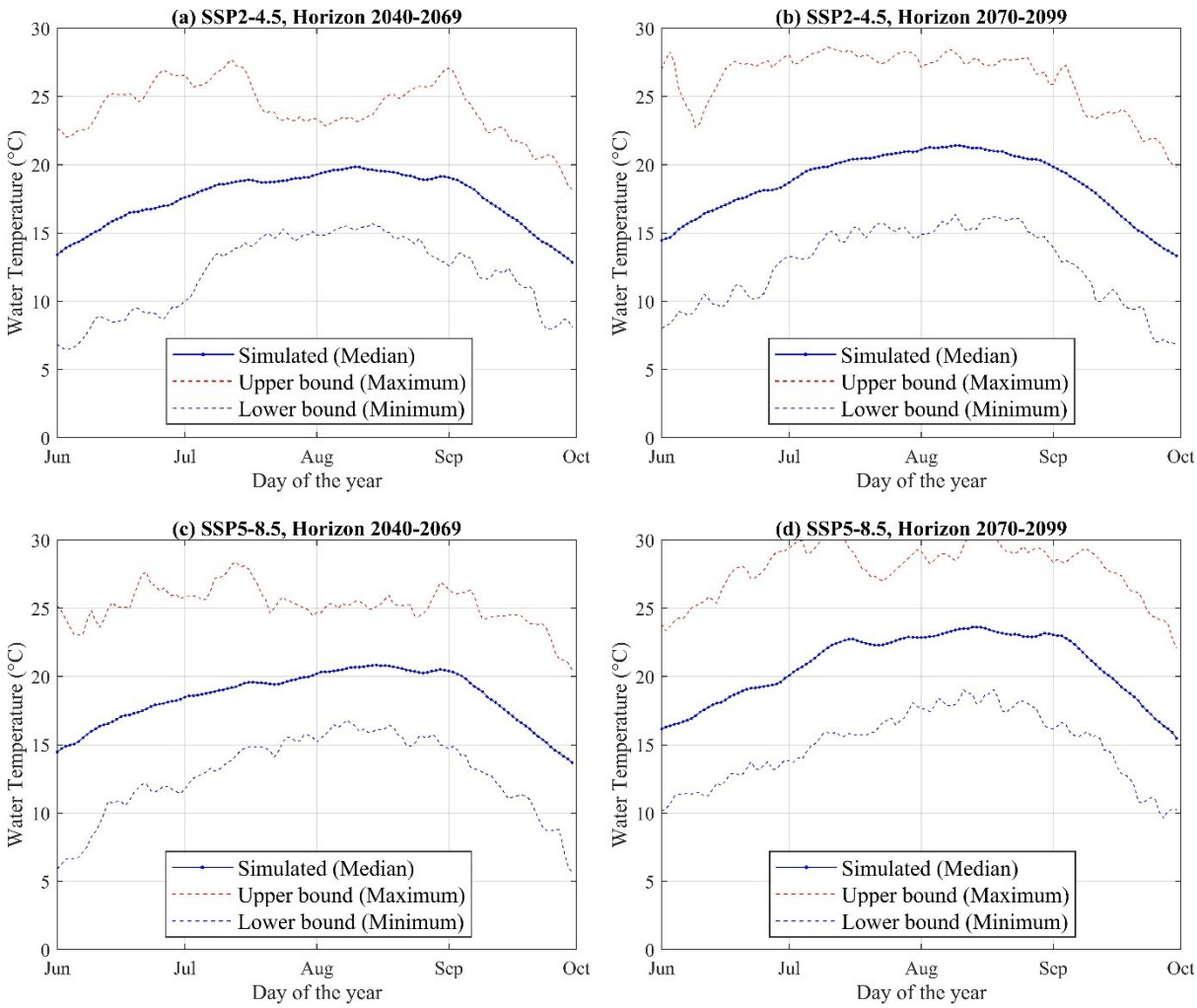

Figure 7.8 **High water temperatures timing for the June-September period for the simulated time series in the future, together with the lower and upper boundary of simulated values**

The mean and median temperature for the July-August period on a yearly basis, along with the number of days above 20°C per year, provide a better understanding of the potential frequency and duration of thermally stressful events for Sockeye salmon in future scenarios (Figure 7.9) compared to the past (shown in Figure 7.5).

The number of days above 20°C during summer each year differs between the four situations (Figures II-C1 to II-C4 in Appendix II-C). The four situations (SSP4.5 and SSP8.5 for the 2040-2069 and 2070-2099 periods) show an increasing maximum number of days above 20°C during summer, ranging from 38 to 62 days, which is above the model uncertainty of nine days for this criterion. The SSP8.5 for the 2070-2099 period shows that the daily mean water temperature for the STMP period is above 20°C. Future water temperature scenarios consistently show higher annual numbers of threshold exceedances than the maximum level that occurred during the baseline period, which was nine days.

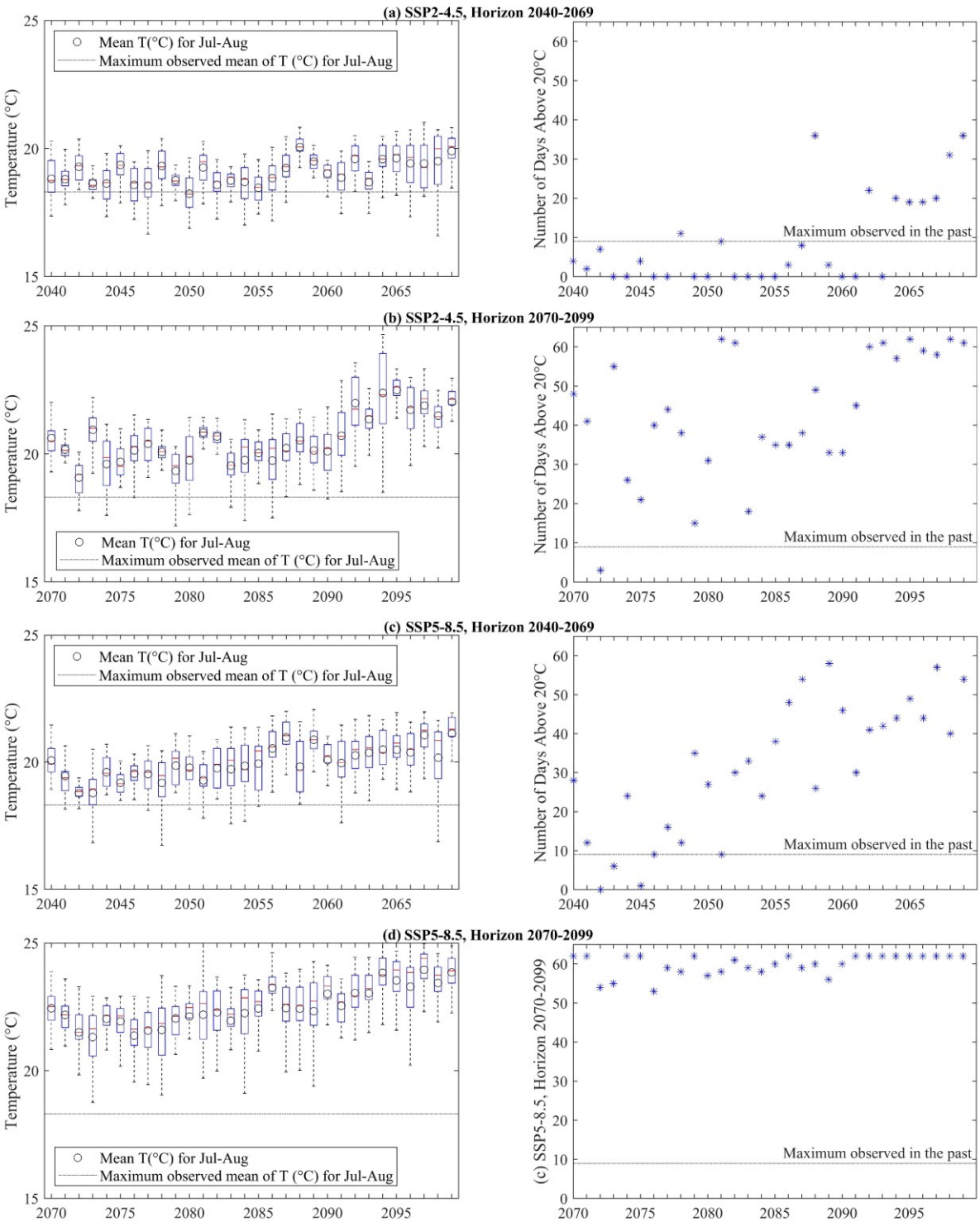

Figure 7.9 **The yearly range of water temperature variability is shown using boxplots for two time horizons and two SSPs. The red line represents the median multi-model value, and the black circles show the mean multi-model simulated value. The upper and lower limits of the boxes show 25th and 75th percentiles, respectively, and the whiskers show the most extreme data. The horizontal dotted line on the left panels shows the maximum observed mean of temperatures for July-August months in the past. Blue asterisks on the right-side plots show the annual number of days on which the temperature exceeds the critical limit of 20°C for simulated and observed time series using multi-model median values. The horizontal dotted line shows the maximum number of observed days above 20°C in a year, which occurred during the 2005-2019 period and is equal to 9 days.**

7.4 Discussion

7.4.1 Hydrothermal modeling

The CEQUEAU calibration performance metrics are satisfactory (KGE and NSE ≥ 0.6 for stream flows; RMSE ≤ 1.5°C for water temperature for Vanderhoof), comparable or better than those provided in previous studies on the Nechako (Table II-C1 in Appendix II-C) (Islam *et al.*, 2019; Kwak *et al.*, 2017b; Ouellet-Proulx, 2018; Ouellet-Proulx *et al.*, 2017b; Ouellet-Proulx *et al.*, 2019). However, the hydrological calibration for the Nautley station shows weaker performance metrics than at Vanderhoof. It is suspected that the weaker model performance at this station is related to the impact of relatively large lakes upstream of the station. The CEQUEAU model can conceptually account for the slower water routing through lakes, but it is somewhat limited since it considers flow routing in a simplified and conceptual form. Given that little is known about the water residence time in those lentic habitats, this could be revisited if additional information is gathered. Khorsandi *et al.* (2022) also showed that simulated water temperature is positively biased downstream of large water bodies in the Nechako River watershed (e.g., the CEQUEAU simulations are underestimating the observed temperature values downstream of the large water bodies with Bias > 0.5°C). Lake surfaces provide a large area exposed to solar radiation and warm air, leading to warming surface water. Stations 6-8 in Figure 7.4 show this Bias (underestimation in water temperature) for the stations upstream of the Nautley confluence. However, this Bias does not spread downstream, where the water temperature is also more influenced by local meteorological forcings (Gatien *et al.*, 2022; Khorsandi *et al.*, 2022). As a result, the simulations at the Vanderhoof station are unbiased. Therefore, we can say the future projections at the Vanderhoof station are unbiased, too.

For stations 6-8, part of this Bias may be due to water temperature boundary conditions at SLS. Synthetic time series were produced using DOY means for the reference period. Gatien *et al.* (2022), using the HEC-RAS hydraulic model for the reaches of the Nechako River between the SLS and Vanderhoof, showed that the impact of upstream boundary conditions on the simulated temperature at Vanderhoof is negligible. However, since CEQUEAU does not include a complex hydraulic water routing scheme like HEC-RAS, upstream boundary conditions may impact the stations upstream of the Nautley confluence (Stations 6-8), which may partly explain the presence of thermal simulation biases.

7.4.2 Impacts of future upstream boundary conditions on downstream water temperatures

This study used the CE-QUAL-W2 outputs for water temperature at the SLS based on using current reservoir operations (STMP) to constrain the upstream boundary condition. The current reservoir operation model defines powerhouse water intake and SLS outputs. This assumption is based on a business-as-usual operation based on STMP by Rio Tinto and considers the water temperature stratification inside the reservoir due to climate change impacts. Although CE-QUAL-W2 resolves water stratification inside the reservoir, water is only released at SLS from the surface layer of the reservoir, and reservoir management cannot provide cold water to the Nechako River unless the water is taken from deeper layers. However, both flow and temperature may experience regime shifts due to the impact of climate change on reservoir stratification and climate change mitigations by Rio Tinto (e.g., future updated STMP or installing a new water release facility). Currently, there is no facility to manage water temperature downstream at the Nechako Reservoir directly. Therefore, testing other reservoir management rules that explicitly consider water temperature downstream to reduce the exceedance of the 20°C threshold is required. Further studies can analyze the implications and efficiency of these changes on the upstream boundary conditions under climate change scenarios.

7.4.3 Future water temperature simulation under climate change impacts

The results emphasize a projected increase in water temperature at the Vanderhoof station during the 2040 to 2100 period. Past observed water temperature data and historical simulations at Vanderhoof show similar exceedances of the 20°C threshold (1% for observations and simulations; observations exist from 2005 to 2019). Previously modeled temperatures for the 1950-2015 historical period showed that the number of days above 20°C in the summertime doubled over 65 years for the whole basin, and the whole Fraser River basin's mean summer temperature rose by 1°C during this period (Islam *et al.*, 2019).

They reported that during the summer periods of 1960 to 2000 there was a three-day decrease in 20°C exceedances for the Vanderhoof station. This small reduction in 20°C exceedances for Vanderhoof station while reporting warmer water at the Fraser basin may be explained by the implementation of the STMP protocol by Rio Tinto, which focused on the Vanderhoof station to keep the mean temperature below the 20°C threshold.

Moreover, another study showed an increasing trend in water temperature for the Nechako watershed during summertime using climate change impact assessments (Picketts *et al.*, 2017). Warmer summers mean limiting conditions for Sockeye salmon, their migration, and spawning time upstream of Vanderhoof station (Macdonald, 2019). Our simulations for

exceedance of 20°C in near climate horizons (2040-2069) show 0-36 days for SSP 4.5 and 0-58 days for SSP 8.5, which is a high increase confirmed by previous studies with a 2°C water temperature increase for this horizon (Picketts *et al.*, 2017). Results for far future horizons (2070-2099) show even higher exceedance frequencies for the 20°C threshold. Our estimations indicate 3-62 days for SSP2-4.5. The SSP5-8.5 scenario simulations show 53-62 days beyond the 20°C threshold. These simulations resulted from considering natural or at least historic flow rules based on STMP. However, it may be possible to modify SLS water releases to target longer and more frequent heat waves to mitigate the increase in temperatures.

7.4.4 Ecological implications of Nechako River water temperature warming

Sockeye salmon is one of the most vulnerable species to rising water temperatures, with 100% mortality beyond 21°C in water temperature after 72 hours following exhaustive exercise (Middleton *et al.*, 2018; Robinson *et al.*, 2015). Scatter plots of the daily mean temperature from July-August for each year of the baseline period (Figure 7.5) indicate some years with nine days above 20°C, with a rising trend in the number of days.

All four projections (SSP2-4.5 and SSP5-8.5 for 2040-2069 and 2070-2099) show an increasing water temperature and an increased frequency of days with > 20°C water temperature. While for the past period, the maximum number of exceedance days is 9 for each summer, for these four combinations, the maximum number of 36, 62, 58, and 62 days is expected based on daily simulations using eight climate models. Considering the lethal threshold for Sockeye salmon (Table II-A1 in Appendix II-A), the number of days with a temperature of more than 20°C is a severe and alarming signal for this species' habitat (Deinet *et al.*, 2020).

7.5 Conclusion

This study provided water temperature simulations for the historical baseline period of 1980-2019, for the near future (2040-2069), and the far future (2070-2099) for the Nechako River in British Columbia, Canada. The study used CMIP6 climate models and SSP2-4.5 and SSP5-8.5 climate change scenarios. The CEQUEAU model was forced with ERA5 data for the past and eight CMIP6 climate models for the future. The study's main finding was the ability of the model to simulate water temperatures during the summer at all eight observation sites, acceptably. The CEQUEAU model provided reliable water flow and temperature values for the entire Nechako River watershed. The results of the study provide ensemble estimations of water temperature under climate change scenarios, which are necessary for decision-makers in the Nechako River watershed. Climate change scenarios indicate that there will be 3.8-36%

more days with water temperatures higher than 20°C, which may pose a severe threat to the Sockeye salmon population due to changes in the Nechako River's thermal regime. The frequency of days with a mean water temperature of more than 20°C is expected to be 0-62 days during July-August, compared to 0-9 days in the past. The water temperature simulation results indicated a relatively high probability of exceeding the 20°C thermal limit at Vanderhoof in the Nechako River. Scenarios indicate that potentially highly stressful conditions for cold-water species like Sockeye salmon during high-temperature events will likely occur more frequently. Future studies are needed to assess the possible impact of dam operation as an adaptation strategy to tackle this and to implement other solutions as mitigation for increasing water temperature.

Acknowledgement

This work was funded by the Canadian Natural Sciences and Engineering Research Council (NSERC) and Rio Tinto as part of a Collaborative Research and Development grant (Grant Number: CRDPJ 523640-18).

ARTICLE II BIBLIOGRAPHIE

Afshar A, Kazemi H, Saadatpour M (2011). Particle swarm optimization for automatic calibration of large scale water quality model (CE-QUAL-W2): Application to Karkheh Reservoir, Iran. *Water Resources Management* 25(10):2613-2632.

Ahmad SK, Hossain F, Holtgrieve GW, Pavelsky T, Galelli S (2021). Predicting the likely thermal impact of current and future dams around the world. *Earth's Future* 9(10):e2020EF001916.

Ahmadi-Nedushan B, St-Hilaire A, Ouarda TBMJ, Bilodeau L, Robichaud É, Thiémonge N, Bobée B (2007). Predicting river water temperatures using stochastic models: case study of the Moisie River (Québec, Canada). *Hydrological Processes* 21(1):21-34.

Algera DA, Kamal R, Ward TD, Pleizier NK, Brauner CJ, Crossman JA, Leake A, Zhu DZ, Power M, Cooke SJ (2022). Exposure risk of fish downstream of a hydropower facility to supersaturated total dissolved gas. *Water Resources Research* 58(6):e2021WR031887.

Arsenault R, Brissette F, Martel J-L (2018). The hazards of split-sample validation in hydrological model calibration. *Journal of hydrology* 566:346-362.

Arsenault R, Poulin A, Côté P, Brissette F (2014). Comparison of stochastic optimization algorithms in hydrological model calibration. *Journal of Hydrologic Engineering* 19(7):1374-1384.

Benyahya L, Caissie D, St-Hilaire A, Ouarda TB, Bobée B (2007). A review of statistical water temperature models. *Canadian Water Resources Journal* 32(3):179-192.

Bérubé S, Brissette F, Arsenault R (2022). Optimal Hydrological Model Calibration Strategy for Climate Change Impact Studies. *Journal of Hydrologic Engineering* 27(3):04021053.

Bond J (2017). *The 2017 Summer water temperature and flow management project.* Triton Environmental Consultants Ltd., BC, Canada. 40 p

Bosmans J, Wanders N, Bierkens MFP, Huijbregts MAJ, Schipper AM, Barbarossa V (2022). FutureStreams, a global dataset of future streamflow and water temperature. *Scientific Data* 9(1):307.

Cannon AJ (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics* 50(1):31-49.

Cole TM, Wells SA (2006). *CE-QUAL-W2: A two-dimensional, laterally averaged, hydrodynamic and water quality model, version 3.5.* US Army Engineering and Research Development Center. 679 p

Deinet S, Scott-Gatty K, Rotton H, Twardek WM, Marconi V, McRae L, Baumgartner LJ, Brink K, Claussen JE, Cooke SJ (2020). *The living planet index (LPI) for migratory freshwater fish: Technical report.* 30 p

Dugdale SJ, Allen Curry R, St-Hilaire A, Andrews SN (2018). Impact of Future Climate Change on Water Temperature and Thermal Habitat for Keystone Fishes in the Lower Saint John River, Canada. *Water Resources Management* 32(15):4853-4878.

Dugdale SJ, Hannah DM, Malcolm IA (2017a). River temperature modelling: A review of process-based approaches and future directions. *Earth-Science Reviews* 175:97-113.

Dugdale SJ, St-Hilaire A, Allen Curry R (2017b). Automating drainage direction and physiographic inputs to the CEQUEAU hydrological model: sensitivity testing on the lower Saint John River watershed, Canada. *Journal of Hydroinformatics* 19(3):469-492.

Elshall AS, Pham HV, Tsai FT-C, Yan L, Ye M (2015). Parallel inverse modeling and uncertainty quantification for computationally demanding groundwater-flow models using covariance matrix adaptation. *Journal of Hydrologic Engineering* 20(8):04014087.

Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D (2007). The shuttle radar topography mission. *Reviews of Geophysics* 45(2):33.

Ficklin DL, Luo Y, Stewart IT, Maurer EP (2012). Development and application of a hydroclimatological stream temperature model within the Soil and Water Assessment Tool. *Water Resources Research* 48(1):16.

Fniguire F, Laftouhi N-E, Al-Mahfadi AS, El Himer H, Khalil N, Saidi ME (2022). Hydrological modelling using the distributed hydrological model CEQUEAU in a semi-arid mountainous area: a case study of Ourika watershed, Marrakech Atlas, Morocco. *Euro-Mediterranean Journal for Environmental Integration* 7(1):89-102.

Fullerton AH, Sun N, Baerwalde MJ, Hawkins BL, Yan H (2022). Mechanistic simulations suggest riparian restoration can partly counteract climate impacts to juvenile salmon. *JAWRA Journal of the American Water Resources Association* 58(4):525-546.

Gatien P, Arsenault R, Martel J-L, St-Hilaire A (2022). Using the ERA5 and ERA5-Land reanalysis datasets for river water temperature modelling in a data-scarce region. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 10.1080/07011784.2022.2113917:1-18.

Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of hydrology* 377(1-2):80-91.

Hansen N (2023). The CMA evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772* https://doi.org/10.48550/arXiv.1604.00772.

Hansen N, Ostermeier A (1996). Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. *Proceedings of IEEE international conference on evolutionary computation.* (Nagoya, Japan, IEEE, p 312-317.

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J-N (2020a). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146(730):1999-2049.

Islam SU, Hay RW, Déry SJ, Booth BP (2019). Modelling the impacts of climate change on riverine thermal regimes in western Canada's largest Pacific watershed. *Scientific Reports* 9(1):11398.

Karakoyun Y, Yumurtacı Z, Dönmez AH (2018). Chapter 4.9 - Environmental fow assessment methods: A case study. *Exergetic, Energetic and Environmental Dimensions*, Dincer I, Colpan CO,Kizilkan O (Édit.) Academic Press, https://doi.org/10.1016/B978-0-12-813734-5.00060-3. p 1061-1074.

Karra K, Kontgis C, Statman-Weil Z, Mazzariello JC, Mathis M, Brumby SP (2021). Global land use/land cover with Sentinel 2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS.* (Brussels, Belgium, IEEE, p 4704-4707.

Khorsandi M, St-Hilaire A, Arsenault R (2022). Multisite calibration of a semi-distributed hydrologic and thermal model in a large Canadian watershed. *Hydrological Sciences Journal* 67(14):2147-2174.

Kim Y, Kim B (2006). Application of a 2-dimensional water quality model (CE-QUAL-W2) to the turbidity interflow in a deep reservoir (Lake Soyang, Korea). *Lake and Reservoir Management* 22(3):213-222.

Kwak J, St-Hilaire A, Chebana F (2017a). A comparative study for water temperature modelling in a small basin, the Fourchue River, Quebec, Canada. *Hydrological Sciences Journal* 62(1):64-75.

Kwak J, St-Hilaire A, Chebana F, Kim G (2017b). Summer season water temperature modeling under the climate change: Case study for fourchue river, Quebec, Canada. *Water (Switzerland)* 9(5).

Larabi S, Schnorbus MA, Zwiers F (2022). A coupled streamflow and water temperature (VIC-RBM-CE-QUAL-W2) model for the Nechako Reservoir. *Journal of Hydrology: Regional Studies* 44:101237.

Liang X, Wood EF, Lettenmaier DP (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global and Planetary Change* 13(1):195-206.

Macdonald J, Morrison J, Patterson D, Heinonen J, Foreman M (2007). *Examination of factors influencing Nechako River discharge, temperature, and aquatic habitats.* Fisheries and Oceans Canada, Vancouver, BC, (2773). 32 p

Macdonald JS (2019). *Saik'uz First Nation and Stellat'en First Nation v. Rio Tinto Alcan, et al. BCSC No.116524, Vancouver Registry, Sockeye Salmon and Water Temperature.* Department of Justice Canada, Canada. 1-48 p

Middleton CT, Hinch SG, Martins EG, Braun DC, Patterson DA, Burnett NJ, Minke-Martin V, Casselman MT (2018). Effects of natal water concentration and temperature on the behaviour of up-river migrating sockeye salmon. *Canadian Journal of Fisheries and Aquatic Sciences* 75(12):2375-2389.

Monteith J, Unsworth M (2013). *Principles of environmental physics: plants, animals, and the atmosphere.* Academic Press, USA. 403 p

Morin G, Couillard D (1990). Predicting river temperatures with a hydrological model. *Encyclopedia of fluid mechanic, surface and groundwater flow phenomena,* Volk Gulf Publishing Company, Houston, Texas Vol 10. p 171-209.

Murray FW (1967). On the computation of saturation vapor pressure. *Journal of Applied Meteorology and Climatology* 6(1):203-204.

Nash JE, Sutcliffe JV (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of hydrology* 10(3):282-290.

Nash KL, Cvitanovic C, Fulton EA, Halpern BS, Milner-Gulland EJ, Watson RA, Blanchard JL (2017). Planetary boundaries for a blue planet. *Nature Ecology & Evolution* 1(11):1625-1634.

Ouellet-Proulx S (2018). *Prévision thermique d'ensemble en rivière avec assimilation de données*. Doctoral (Université du Québec, Institut national de la recherche scientifique, Quebec, Canada). 222 p. https://espace.inrs.ca/7621/1/T750.pdf

Ouellet-Proulx S, Chiadjeu OC, Boucher M-A, St-Hilaire A (2017a). Assimilation of water temperature and discharge data for ensemble water temperature forecasting. *Journal of hydrology* 554:342-359.

Ouellet-Proulx S, St-Hilaire A, Boucher M-A (2017b). Water temperature ensemble forecasts: Implementation using the CEQUEAU model on two contrasted river systems. *Water* 9(7):457.

Ouellet-Proulx S, St-Hilaire A, Boucher MA (2019). Implication of evaporative loss estimation methods in discharge and water temperature modelling in cool temperate climates. *Hydrological Processes* 33(22):2867-2884.

Ouellet V, St-Hilaire A, Dugdale SJ, Hannah DM, Krause S, Proulx-Ouellet S (2020). River temperature research and practice: Recent challenges and emerging opportunities for managing thermal habitat conditions in stream ecosystems. *Science of The Total Environment* 736:139679.

Picketts IM, Parkes MW, Déry SJ (2017). Climate change and resource development impacts in watersheds: Insights from the Nechako River Basin, Canada. *The Canadian Geographer / Le Géographe canadien* 61(2):196-211.

Robinson KA, Hinch SG, Raby GD, Donaldson MR, Robichaud D, Patterson DA, Cooke SJ (2015). Influence of postcapture ventilation assistance on migration success of adult sockeye salmon following capture and release. *Transactions of the American Fisheries Society* 144(4):693-704.

Schnorbus M (2018). *VIC Glacier (VIC-GL) - Description of VIC model changes and upgrades, VIC Generation 2 Deployment Report.* Pacific Climate Impacts Consortium, Victoria, BC. 40 p

Sheedy B (2005). *Analysis of a cold water release facility in the Nechako Reservoir*. Master of Business Administration (Faculty of Business Administration-Simon Fraser University). 90 p. https://summit.sfu.ca/_flysystem/fedora/sfu_migrate/10292/etd2058.pdf

Shen H, Tolson BA, Mai J (2022). Time to update the split-sample approach in hydrological model calibration. *Water Resources Research* 58(3):e2021WR031523.

Smaoui H, Zouhri L, Kaidi S, Carlier E (2018). Combination of FEM and CMA-ES algorithm for transmissivity identification in aquifer systems. *Hydrological Processes* 32(2):264-277.

St-Hilaire A, Boucher M-A, Chebana F, Ouellet-Proulx S, Zhou QX, Larabi S, Dugdale S, Latraverse M (2015). Breathing a new life to an older model: The CEQUEAU tool for flow and water temperature simulations and forecasting. *Proceedings of the 22nd Canadian Hydrotechnical Conference.* (Montréal, Canada, April 29–May 2), p 10.

Stratton Garvin LE, Rounds SA, Buccola NL (2022). *Updates to models of streamflow and water temperature for 2011, 2015, and 2016 in rivers of the Willamette River Basin, Oregon.* U.S. Department of the Interior, U.S. Geological Survey, USA, (2022-1017). 88 p

Sullivan AB, Rounds SA (2021). *Modeling water temperature response to dam operations and water management in Green Peter and Foster Lakes and the South Santiam River, Oregon.* U.S. Department of the Interior, U.S. Geological Survey, USA, (2020-5145). 37 p

Wilson KL, Kay LM, Schmidt AL, Lotze HK (2015). Effects of increasing water temperatures on survival and growth of ecologically and economically important seaweeds in Atlantic Canada: implications for climate change. *Marine Biology* 162(12):2431-2444.

Xiong YJ, Yin J, Paw U KT, Zhao SH, Qiu GY, Liu Z (2020). How the three Gorges Dam affects the hydrological cycle in the mid-lower Yangtze River: a perspective based on decadal water temperature changes. *Environmental Research Letters* 15(1):014002.

Yu X, Bhatt G, Duffy C, Shi Y (2012). A Two-Scale Parameterization for Distributed Watershed Modeling Using National Data and Evolutionary Algorithm. *AGU Fall Meeting Abstracts.* p H31H-1231.

Zanaga D, Van De Kerchove R, De Keersmaecker W, Souverijns N, Brockmann C, Quast R, Wevers J, Grosu A, Paccini A, Vergnaud S, Cartus O, Santoro M, Fritz S, Georgieva I, Lesiv M, Carter S, Herold M, Li L, Tsendbazar N-E, Ramoino F, Arino O (2021). ESA WorldCover 10 m 2020 v100.).

Zhu S, Piotrowski AP (2020). River/stream water temperature forecasting using artificial intelligence models: a systematic review. *Acta Geophysica* 68(5):1433-1442.

Zouhri L, Kaidi S, Smaoui H (2021). Parameter Identification by High-Resolution Inverse Numerical Model Based on LBM/CMA-ES: Application to Chalk Aquifer (North of France). *Water* 13(11):1574.

8 ARTICLE III

Evaluation of downward longwave irradiance models with ERA5 input data in Canadian environments

Évaluation des modèles d'irradiance à ondes longues descendantes avec des données d'entrée ERA5 dans les environnements canadiens

Auteurs :

Mostafa Khorsandi[1], André St-Hilaire[1], Richard Arsenault[2]

[1]Canadian Rivers Institute and Institut National de la Recherche Scientifique, Centre Eau Terre Environnement (INRS-ÉTÉ), 490 De la Couronne St., Québec City, Quebec, G1K 9A9, Canada

[2]Hydrology, climate and climate change laboratory, École de technologie supérieure (ÉTS), Montréal, Quebec, Canada

L'article sera soumis dans la revue *Journal of Hydrometeorology*

Date de soumission : Janvier 2024

Contribution des auteurs:

Mostafa Khorsandi: Curation des données, méthodologie, conceptualisation, programmation, analyse formelle, enquête, validation, visualisation, rédaction - ébauche originale, rédaction - révision et édition.

André St-Hilaire: Acquisition de financement, administration de projet, ressources, supervision, validation, rédaction – révision & édition.

Richard Arsenault: Ressources, supervision, rédaction – révision & édition.

Résumé

Le rayonnement à ondes longues (LR) est l'une des composantes du bilan énergétique responsables du réchauffement et du refroidissement de l'eau. Tant le LR incident descendant, émis par l'atmosphère, que le LR sortant émis par la surface terrestre ne sont pas souvent mesurés. L'influence des nuages et de la canopée sur le bilan thermique du LR rend encore plus difficile l'établissement de formulations fiables pour les conditions variables du ciel. Cette étude utilise la température de l'air, la température du point de rosée, la couverture nuageuse et l'indice de surface foliaire (LAI) de la base de données de réanalyse ERA5 pour comparer neuf modèles de l'irradiance à ondes longues descendante (DLI) à la surface de la Terre avec le produit DLI d'ERA5. Notre travail utilise des données de DLI mesurées en continu sur une longue période à deux stations au Canada ainsi que la réanalyse ERA5, une source fiable pour les régions pauvres en données, telles que le centre de la Colombie-Britannique et le nord du Québec (Canada). Les résultats montrent la faisabilité de l'utilisation de différents modèles de DLI avec des données de réanalyse ERA5 pour des conditions de couverture nuageuse variables, avec des métriques RMSE variant de 30 à 38 W/m^2, ce qui est comparable à des études précédentes et peut facilement être appliquées à des échelles plus larges en l'intégrant dans des modèles hydrologiques. De plus, il est démontré que les données sur grille d'ERA5 pour le DLI présentent les meilleurs résultats avec un RMSE de 30 à 32 W/m^2 aux deux stations. Cette meilleure performance suggère l'utilisation directe des données ERA5 en tant que données d'entrée pour les modèles thermiques.

Mots-clés

Rayonnement à ondes longues, données ERA5, modèles empiriques.

Abstract

Longwave radiation (LR) is one of the energy balance components responsible for warming and cooling water. Both downward LR, emitted by the atmosphere, and outgoing LR emitted by the land surface are not widely measured. The influence of clouds and canopy on the LR heat budget makes it even harder to establish reliable formulations for sky conditions. This paper uses air temperature, dew point temperature, cloud cover, and leaf area index (LAI) from the ERA5 reanalysis database to compare nine models for downward longwave irradiance (DLI) at the Earth's surface and to compare them with ERA5's DLI product. Our work uses long-time continuous DLI measured data at two stations in Canada and the ERA5 reanalysis, a reliable source for data-scarce regions, such as central British Columbia and northern Quebec (Canada). The results show the feasibility of using several DLI models using ERA5 reanalysis data for all sky conditions with Root Mean Square Errors (RMSE) performance metrics ranging from 30 to 38 W/m^2, which is comparable with previous studies. Moreover, it is shown that ERA5 gridded data for DLI show the best results with RMSE ranging from 30 to 32 W/m^2 at the two stations. This higher performance suggests that ERA5 DLI can be used directly as input data for thermal models.

Key Words

Longwave radiation, ERA5 data, Empirical models.

8.1 Introduction

Longwave radiation (LR) is one of the main heat budget terms responsible for several environmental processes in land and aquatic environments (Bernhardt *et al.*, 2022; McFarlane & Clark, 2021; Ouellet *et al.*, 2014). LR exchange, which refers to wavelengths longer than 4 µm, is controlled by the surface's thermal emissivity, atmospheric objects, clouds, canopy, water vapor, carbon dioxide and other greenhouse gases (Herrero & Polo, 2012). There are two LR-related variables (downward and upward radiation); the downward longwave irradiance (DLI) acts as a signature of the atmosphere condition (Tang *et al.*, 2021). Therefore, it controls an important part of the surface heat budget (Zhou *et al.*, 2020). One of the critical roles of LR is the cooling/warming of water during summer, which highly impacts aquatic flora and fauna. A previous study has shown that this component of the heat budget balance is responsible for cooling the water temperature in the Nechako watershed in central British Columbia, Canada (Khorsandi et al., 2022). Therefore, using a more accurate longwave radiation balance in this region can improve the energy balance modeling. The more precise representation of DLI can improve the simulation results and analysis of hydrological and water quality models (e.g., water temperature) (Gatien *et al.*, 2022; Khorsandi *et al.*, 2022).

Over the past decades, several studies developed the scientific foundations for DLI (Zhou *et al.*, 2020). Although the physical foundation for DLI is well understood and represented by the Stephan-Boltzmann law, the knowledge related to the interaction of variables that affect atmospheric emissivity is still developing. As a result, several studies parameterized and studied DLI using several equations (Bárbaro *et al.*, 2009; Duarte *et al.*, 2006; Herrero & Polo, 2012; Huang *et al.*, 2007; Koll & Cronin, 2018; Sicart *et al.*, 2010). Idso and Jackson (1969) developed a general formula for the surface DLI estimation using a theoretical analysis of atmospheric thermal behavior. Using global-scale measurements, they found that DLI has the minimum effective emittance at 273 K for screen-level air temperature. Based on this finding, they presented a new ground surface DLI modeling equation.

Miller (1981) presented a knowledge synthesis and applications of DLI for ecosystem modeling. Afterward, several scholars developed formulations mainly for air emissivity. It is well known that air temperature and cloud cover are the primary controlling meteorological variables, impacting DLI through air emissivity (Li *et al.*, 2017; Zhou *et al.*, 2020).

By collecting more ground-based measurements, the impact of cloud cover has been studied extensively. Duarte *et al.* (2006) studied the clear-sky and all sky conditions for DLI estimation using well-known parameterizations in Southern Brazil; Sicart *et al.* (2010) did the same for Andean glaciers, Herrero and Polo (2012) for a mountainous site, and Voortman *et al.* (2015) for dune ecosystems. These studies have shown the efficiency of already-developed models in providing reliable estimates for clear-sky conditions.

More recent studies focused on all sky condition parameterizations (Li *et al.*, 2017; Long *et al.*, 2021). For example, Li *et al.* (2017) compared and recalibrated fifteen parametric models for DLI. They showed that previously proposed DLI models can be categorized into only a few families of models. They recommended the Brunt (Brunt, 1932) family of models for future use. Long *et al.* (2021) tested six models for clear-sky conditions and eight for cloudy-sky conditions over the Tibetan Plateau. Their results showed an mean RMSE of 26.4 W/m^2 for all sky conditions. Tang *et al.* (2021) showed that reanalysis products for DLI outperformed satellite retrievals over the land surface. Tang *et al.* (2021) and Niemelä *et al.* (2001) pointed out the extensive possibility of reanalysis data for DLI modeling.

Parallel to the cloud impact assessments and development on DLI, the impact of other meteorological variables has been studied for clear-sky conditions. Prata (1996) focused on clear-sky conditions and developed a new equation based on the impact of air temperature and water vapor on DLI. Several studies showed vapor pressure as the third (after air temperature and cloud cover) primary controlling variable on air emissivity (Abramowitz *et al.*,

2012; Dilley & O'brien, 1998; Huang *et al.*, 2007; Li *et al.*, 2017; Morcrette, 2002; Niemelä *et al.*, 2001; Zhou & Cess, 2000).

Using reanalysis data as a proxy for meteorological variables was recommended in the literature (Gatien *et al.*, 2022; Tarek *et al.*, 2020b). However, no previous research for data-scarce regions tested the benefit of directly using a reanalysis-based product for DLI or using other reanalysis variables as a proxy for meteorological measurements to calculate DLI.

This work presents a comparative analysis of several DLI formulations validated against field measurements taken from 2 December 2019 to 31 May 2022. The study area is the Nechako watershed (British Colombia, Canada), which has one DLI monitoring station [at Skins Lake spillway (SLS); see Figure 8.1]. Using this analysis, selected well-known parameterizations are calibrated for their parameters to calculate the air emissivity for all sky conditions (e.g., clear-sky and cloudy-sky conditions) based on hourly ERA5 data. The calibrations and validations are performed using measurements at SLS. A second calibration/validation is performed on an experimental drainage basin operated by the École de technologie supérieure (QC, Canada). This station is called the “Bassin Versant Expérimental de Sainte-Marthe” station (hereafter BVESM station). Two approaches were performed and validated at the study site for DLI modeling: (1) comparing the physically-based Stephen-Boltzmann law with several parameterizations for air-emissivity available in the literature and (2) directly using ERA5 gridded DLI data instead of calculating DLI. Finally, the possible application of the several parameterizations is addressed for future modeling at the Nechako watershed.

8.2 Methods

8.2.1 Study areas

The Nechako River watershed (downstream of Skins Lake Spillway to Finmoore) has a 12,805 km^2 drainage area with two main rivers, Nechako and Nautley. This watershed is located in central British Columbia, Canada (Figure 8.1). The present paper focuses on the aforementioned measuring site for radiation measurements, SLS, in addition to BVESM (Table 8.1). The SLS station is located on the northeast extremity of the Nechako reservoir, with a low canopy cover [Figures A1(a) and (b) in Appendix III-A]. The observed data were used without any pre-processing.

Table 8.1 Names and locations of longwave radiation measuring stations

Station	Longitude (West)	Latitude (North)	Elevation (masl)	Distance from closest ERA5 grid point (km)	Data availability	
					Start	End
Skins Lake Spillway	-125.997	53.77	865	2.5	2 Dec 2019	31 May 2022
BVESM	-74.28	45.43	144	12.2	1 Jan 2020	31 Dec 2020

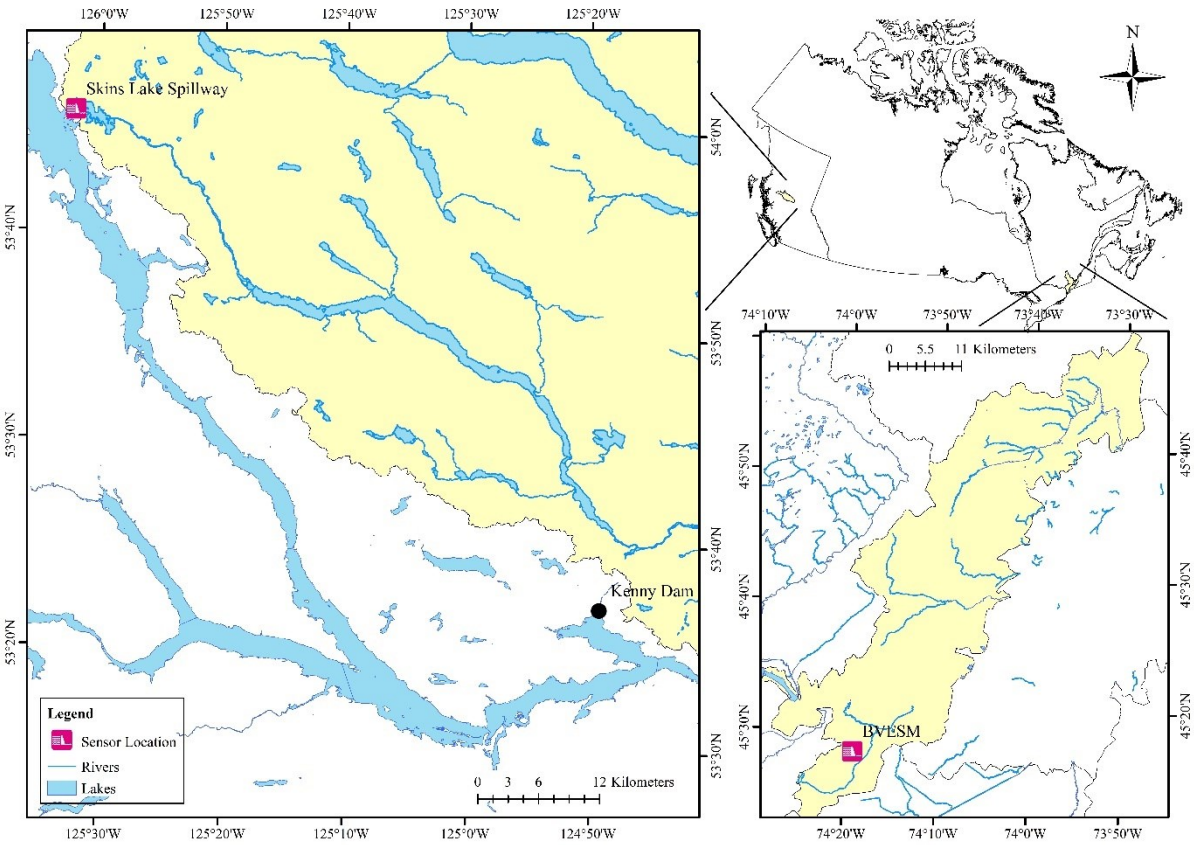

Figure 8.1 Nechako River and Bassin versant expérimental de Sainte-Marthe watersheds and measuring stations at Skins Lake Spillway and Bassin versant expérimental de Sainte-Marthe (BVESM)

The SLS station was used for calibration (60% of data), and validation was performed on the remaining 40% of data. The second station that was used for calibration and validation in the same way was the BVESM station (http://bvesm.etsmtl.ca/) (Figure 8.1 and Table 8.1). This station is located approximately 70 km west of Montreal in Quebec, Canada (45.43N, 74.28W). Among all the measured hydroclimatic variables, DLI was used. In contrast to the other station, the BVESM station is located in a small clearing within a dense canopy (Figures A1(c) and (d) in Appendix III-A).

8.2.2 ERA5 reanalysis data

The required input meteorological variables for the models were provided by the European Center for Medium-Range Weather Forecasts (ECMWF) through their European Reanalysis 5th generation (ERA5) (Hersbach *et al.*, 2020a). ERA5 reanalysis data include all required meteorological variables for this study. Gatien *et al.* (2022) and Khorsandi *et al.* (2022) showed the possibility of using ERA5 reanalysis data for water temperature modeling in calibration and simulation steps using the Nechako River watershed as their case study. In this study, five variables from the ERA5 database were used: (1) DLI (W/m^2), (2) air temperature at 2 m (K, converted to °C), (3) dew point temperature at 2 m (K, converted to °C), (4) total cloud cover (dimensionless), and (5) Leaf Area Index (LAI) for high vegetation (m^2/m^2). The 2 m data were used as the DLI sensors were installed at the same height. ERA5 hourly data available from the ECMWF Climate Data Store (CDS) disks have been pre-interpolated to regular latitude/longitude grid points (https://confluence.ecmwf.int/). The original computation cells for ECMWF's model have a resolution of 31 km (0.28125 degrees), which is interpolated to a reduced Gaussian grid called "linear grids" [mentioned as "native grid" in ERA5 documentation with the technical name "reduced Gaussian grid N320"] (https://confluence.ecmwf.int/; Figure III-A1 in Appendix III). The closest grid point to each station was used.

8.2.3 Longwave heat budget modeling

The most fundamental equation for longwave radiation is the Stefan-Boltzmann equation. The Stefan–Boltzmann law is a well-known equation to quantify the radiated flux from a black body in terms of its absolute temperature. It expresses that the total energy emitted from a black body object per unit area per time is proportional to the fourth power of black body temperature, in Kelvin:

$$J = \sigma T^4 \tag{8.1}$$

where J= radiant emittance (Wm^{-2}), σ= Stefan–Boltzmann constant ($5.670373 \times 10^{-8} Wm^{-2}K^{-4}$), and T= thermodynamic temperature (K). Not all radiation is emitted for natural objects, like a perfect black body, and the emissivity varies with wavelengths. However, the total emissivity for natural objects is usually estimated using a constant without considering wavelength dependency. The final equation is:

$$J = \beta \sigma T^4 \tag{8.2}$$

where β= emissivity of the object (dimensionless; $0 < \beta < 1$). This equation can be used as downward longwave radiation for clear-sky conditions. The role of water vapor, in absorbing longwave radiation is crucial due to its contribution to the greenhouse effect. Water vapor

absorbs infrared radiation, thereby trapping heat in the atmosphere (Satterlund, 1979). This effect is particularly pronounced in the lower parts of the atmosphere where water vapor concentration is highest. Consequently, equations derived for air emissivity often includes variables related to humidity (Satterlund, 1979). To estimate air emissivity (β), some equations are proposed for hydrological studies. For example, as a common equation for hydrological studies in Canada, Morin and Couillard (1990) presented the following equation:

$$\beta = 0{,}97(1 + 0{,}17B^2)(0{,}74 + 0{,}0065e_a) \tag{8.3}$$

where B= total cloud cover (dimensionless) and e_a= vapor pressure ($mm\,Hg$). e_a can be estimated using the Tetens equation and the dew point temperature as follows (Andrews, 2010; Murray, 1967; Upreti & Ojha, 2017):

$$e_a = e_s(T_{dew}) = 7{,}50062 \times \begin{cases} 0{,}61078 \times \exp\left(\frac{17{,}27T_{dew}}{T_{dew} + 237{,}3}\right) & T_{dew} \geq 0°C \\ 0{,}61078 \times \exp\left(\frac{21{,}875T_{dew}}{T_{dew} + 265{,}5}\right) & T_{dew} < 0°C \end{cases} \tag{8.4}$$

where e_s= saturated vapor pressure (mmHg), T_{dew}= dew point temperature ($°C$), and 7.50067 is conversion factor ($mmHg/kPa$).

Among all cloud-related variables, total cloud cover (B) is usually used for applied DLI studies (Morin & Couillard, 1990; Morin *et al.*, 1994; Ouellet *et al.*, 2014). However, different clouds with varying effects on DLI can have the same total cloud cover value (e.g., Cirrus clouds may have higher efficiency than low-level stratiform clouds in influencing DLI, while cirrus cloud and low-level stratiform clouds or their combination can give the same value of total cloud cover). The other reason for using the total cloud cover variable is the formulations of the cloud cover impact in the available DLI models. The variables used in these equations are close to the surface measurements and developed based on the total cloud cover variable. Therefore, the most similar variable from the ERA5 dataset to implement in this study is the total cloud cover (tcc). It has been shown that cloud cover and vapor pressure impact the incoming longwave radiation from the air (Hogan & Bozzo, 2016; Morcrette, 2002; Schafer, 2017). However, the relationship between effective emissivity ($\beta_{effective}$) for a cloudy condition related to cloud cover, air temperature, and vapor pressure is complex. One method is to use advanced global climate model output products, empowered with computational fluid dynamic methods, to tackle the complex relationship of thermodynamic variables (Tang *et al.*, 2021). However, simple mathematical formulations have historically been used to calculate DLI for local and terrestrial purposes. Sugita and Brutsaert (1993) studied the impact of cloud effect

on DLI. They categorized equations for two conditions: (1) clear-sky and (2) all sky, as described below.

8.2.3.1 Clear-sky longwave radiation

Sugita and Brutsaert (1993) showed that Equation (8.2) is accurate enough for a clear-sky. The only problem is estimating β for the atmospheric condition as a grey object. Duarte *et al.* (2006) showed DLI for clear-sky conditions is the net result of longwave emission and atmospheric absorption. However, almost all the equations for clear-sky conditions in the literature have the simplifying (but satisfactory) assumption that near-surface air temperature and vapor pressure control β. Morin and Couillard (1990), Sugita and Brutsaert (1993), and Duarte *et al.* (2006) reviewed the most common equations to estimate β.

Swinbank (1963) acknowledged the linear correlation between β and e_a. However, Swinbank (1963) used newly measured data from Australia, England, and France and proposed an alternative physically based equation for β in clear-sky conditions based on air temperature only:

$$\beta = 9.3645 \times 10^{-6} \times T_{air}^2 \tag{8.5}$$

[written as $J = 5.31 \times 10^{-14} T_{air}^6$ with J in mW/cm^2 in the original reference; by following the Stephan-Boltzmann equation, it can be written as $J = (9.3645 \times 10^{-6} T_{air}^2)\sigma T_{air}^4$ with J in W/m^2].

Maykut and Church (1973) documented that while Equation (8.5) demonstrates efficacy at lower latitudes, using long-term observations at Point Barrow in Alaska, its performance diminishes as one moves to higher latitudes.

Idso and Jackson (1969) followed-up the work of Swinbank (1963) and proposed a universally applicable equation based on measured air temperature. They showed that the effective emittance is minimum at 273 K (-0.15°C) with a symmetric increase for temperatures below and above this point. As an alternative universal model, they proposed:

$$\beta = 1 - 0{,}261 \times \exp\left[-7{,}77 \times 10^{-4}(273 - T_{air})^2\right] \tag{8.6}$$

Morin and Couillard (1990) used both variables in a linear relationship between β and e_a as:

$$\beta = 0.97 \times (0{,}74 + 0{,}0065 e_a) \tag{8.7}$$

where e_a is in mmHg. This relationship is empirical, and the values are calibrated based on the observations from previous studies.

In another study, Brutsaert (1975) developed a simplified equation by solving Schwarzschild's radiative transfer equation for a relatively standard atmosphere. The proposed function performed very well in humid conditions with relatively warm temperatures (Prata, 1996). Brutsaert (1975) assumed exponential temperature and water vapor density profiles with two decay constants. Brutsaert (1975) also introduced an adjustable parameter representing the ground level emissivity due to water vapor and CO_2 as a function of water vapor path length. Their developed equation is:

$$\beta = 1.24(\frac{e_a}{T_{air}})^{1/7} \ (e_a \text{ in hPa}) \tag{8.8}$$

Brutsaert (1975) equation results in $\beta \approx 0$ for higher latitudes [e.g. Alaska in the study by Maykut and Church (1973)], for conditions with e_a close to zero. These close to zero values violate the atmospheric background effects on emissivity due to greenhouse gases ($\beta > 0$ for $e_a \rightarrow 0$). Therefore Satterlund (1979) developed a new formulation to resolve the estimation of β especially for low temperatures. Satterlund (1979) mentioned that an equation for β for clear-sky conditions should meet three physical criteria: (1) It should include both temperature and vapor pressure (humidity), (2) It must not yield values that exceed the black body radiation at any temperature or vapor pressure extreme, and (3) it should not yield values lower than the β related to the CO_2 content of atmosphere. He argued that the correct form of the function to satisfy the three criteria has an exponential function of vapor pressure. Also, since the saturated vapor pressure (extreme condition for humidity) is a function of temperature, the exponent of vapour pressure should include the air temperature. The final proposed model was:

$$\beta = 1.08[1 - exp\,(-e_a{}^{\frac{T_{air}}{2016}})] \quad (e_a \text{ in hPa}) \tag{8.9}$$

This emissivity model performs better than the previously developed models while resolving the air emissivity for low temperatures.

Prata (1996) followed the work of Brutsaert (1975) and developed a new emissivity model to satisfy the criteria mentioned above for an emissivity model by using radiosonde profiles and a Lowtran-7 (Kneizys *et al.*, 1988) [atmospheric radiance/transmittance model which calculates the irradiance using radiosonde profiles] radiative transfer model as follows:

$$\beta = 1 - \left[1 + 46.5\left(\frac{e_a}{T_{air}}\right)\right] exp\,\langle -\left\{1.2 + 3\left[46.5\left(\frac{e_a}{T_{air}}\right)\right]\right\}^{1/2}\rangle \ (e_a \text{ in hPa}) \tag{8.10}$$

The term $46.5\left(\frac{e_a}{T_{air}}\right)$ (e_a in hPa and T_{air} in K) in their equation refers to the precipitable water content (Prata, 1996). Prata (1996) extensively tested the developed model and achieved good results across various latitudes, encompassing air temperatures ranging from -40°C to

+40°C. In their assessment, the newly developed model outperformed all the previously used models.

Niemelä *et al.* (2001) compared the effective emissivity versus T_{air} and e_a, for two winter and one summer period in Sodankylä, Finland. They observed the lowest effective emissivity for clear-sky conditions when the air temperature was about -10°C or $e_a \approx$ 2 hPa. By implementing the linear relationship between β and e_a, they developed the following equation:

$$\beta = \begin{cases} 0{,}72 + 0{,}009(e_a - 2). & e_a \geq 2 \\ 0{,}72 - 0{,}076(e_a - 2). & e_a < 2 \end{cases} \tag{8.11}$$

Niemelä *et al.* (2001) showed the significance of this equation compared to the previous equations, especially for cold and dry conditions. Since T_{air} and e_a are essential parameters to estimate β for clear-sky, Huang *et al.* (2007) studied clear-sky outgoing longwave radiation sensitivity to these variables. Their results showed that T_{air} has a dominant impact compared to e_a for clear-sky conditions in the low or middle troposphere, while in the upper troposphere, the impacts of each component are nearly equal.

8.2.3.2 All-sky longwave radiation

DLI equations should include cloud cover correction for practical all sky condition purposes (Duarte *et al.*, 2006). Maykut and Church (1973) studied the DLI for Point Barrow, Alaska. Their five-year data collection between 1962 and 1966 in Alaska's cold region provided information for a low vapor pressure (0.08-1.5 mmHg during much of the year) but mostly cloudy conditions. They found the empirical relationship between all-sky condition DLI and meteorological variables as follows (all-sky refers to both clear and cloudy sky conditions):

$$J_{All-sky} = (1 + 0.0004B^{2.75})\, \beta_{Clear-sky} \sigma T_{air}^4, \quad \beta_{Clear-sky} = 0.7855 \tag{8.12}$$

Their fitted empirical equation for $J_{All-sky}$ does not contain the e_a variable which means a constant correction factor ($\beta_{Clear-sky} = 0.7855$) for atmospheric effects, including the effects by e_a multiplied by the cloud cover correction factor. This equation confirms the Bolz (1949) or Budyko (1974) correction factors for cloudy conditions based on clear-sky conditions. Morin and Couillard (1990) showed a similar type of equation using different parameter values. Niemelä *et al.* (2001)'s study in the cold region of Finland showed a correction factor of $1.152 \leq \frac{J_{All-sky}}{J_{Clear-sky}} \leq 1.356$ for cloudy conditions. Duarte *et al.* (2006) concluded that a typical approach to adjust $J_{All-sky}$ based on $J_{Clear-sky}$ is to use correction factor equations [e.g., Bolz (1949) or Budyko (1974)] as follows:

$$J_{All-sky} = J_{Clear-sky}(1 + uB^{v}) \tag{8.13}$$

$$J_{All-sky} = J_{Clear-sky}(1 - B^{w}) + rB^{w}\sigma {T_{air}}^{4} \tag{8.14}$$

where B is the fractional cloud cover from human observers viewpoint, u, v, w, and r are dimensionless constants depending on cloud characteristics with different values in each reference. Duarte *et al.* (2006) showed better results using Equation (8.13) while fitting different DLI formulations using measurements. Therefore, this study uses the Equation (8.13) to convert clear-sky DLI to all sky conditions. Bolz (1949) used $v = 2$ and different values of u. Some examples of this form were presented by other scholars (Duarte *et al.*, 2006) for various case studies where $-1 \leq u \leq 0.5$ and $0 \leq v \leq 4$, such as that in Equation (8.3), where $u = 0.17$ and $v = 2$.

Although cloud cover and cloud parameters are measured by satellite imagery, ground-based observations are not always available. The all sky condition is easily observable, nevertheless hard to quantify. The complexity of measuring cloud cover is due to its relatively small scale size, the height of clouds, the cloud formation and types, and the expensive required measurement devices. Although cloud cover is monitored in crucial places like airports, their data are for local purposes and not transferable to larger scales. Therefore, when cloud cover measurements are unavailable, they can be estimated by the Crawford and Duchon equation (Crawford & Duchon, 1999):

$$B = 1 - \frac{measured\ solar\ irradiance}{clear - sky\ solar\ irradiance} \tag{8.15}$$

However, this equation can only be used for daytime when downward solar radiation measurements are available (Choi *et al.*, 2008; Long *et al.*, 2021).

There is an alternative method of DLI estimation for all sky conditions when implementing Equation (8.13). It is possible to use reanalysis products for cloud cover. The assimilated products of reanalysis climate models provide a reasonably acceptable estimation for cloud cover for large-scale studies (Tang *et al.*, 2021; Yao *et al.*, 2019; Yao *et al.*, 2020). Although different cloud types have a distinct influence on DLI, the total cloud cover is used. Cloud cover is one of the most uncertain variables among reanalysis products since the implemented Earth system model calculates it and is purely model-dependent (Buizza, 2023; Free *et al.*, 2016; Yao *et al.*, 2019; Yao *et al.*, 2020). The reanalysis of assimilated products is probably more reliable for the total cloud cover variable since it is the result of the Earth system model, assimilated by observed satellite data for column water vapor and in-situ measurements for humidity profiles (https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation#ERA5:datadocu

mentation-Observations). Also, total cloud cover variable is an aggregate indicator of all clouds in the sky. Yao *et al.* (2019) compared three reanalysis products and concluded that they could represent the cloud characteristics adequately. In their study, ERA5 reanalysis achieves the best performance metrics related to cloud cover (Yao *et al.*, 2019).

In this study T_{air} and e_a refer to 2-meter measurements. For near-surface DLI modeling, the canopy affects DLI irradiance (absorption and emission). Pomeroy *et al.* (2009) focused on DLI under the canopy and tested three methods for DLI modeling using (1) air temperature, (2) air temperature plus needle-branch and trunk temperatures, and (3) air temperature and shortwave extinction with an optimized efficiency function. They showed the superiority of method (3) and then method (2) to method (1), which only used sky factors. One needs to measure trunk temperatures to apply method (2) and derive the shortwave extinction function to use method (3) in their study. Given the relative paucity of such data, they proposed to assume that canopy temperature is equal to air temperature, with the following equation:

$$\beta_{air-canopy} = V_f \beta_{air} + (1 - V_f)\beta_{canopy} \tag{8.16}$$

where $\beta_{air-canopy}$ is the emissivity for mixed air-canopy situations and V_f is sky view factor (0-1). Oke (1987) reported that $\beta_{canopy} = 0{,}98$ and Verseghy *et al.* (1993) showed an empirical relationship to estimate V_f as follows for Canadian studies:

$$V_f = exp(-\alpha \times LAI) \tag{8.17}$$

where α is a constant for each canopy (with $\alpha = 0.5$ for their case study in Canada), and LAI is the leaf area index. Verseghy *et al.* (1993) proposed this function for the Canadian Land Surface Scheme (CLASS) for global circulation models originally developed to estimate DLI in coniferous forests.

We test this formulation together with Equation (8.3) to form the final model as follows:

$$\begin{aligned}\beta_{air-canopy.All-sky} &= (1 + uB^v)[exp(-\alpha \times LAI)]0{,}97 \times (0{,}74 + 0{,}0065e_a) \\ &+ 0{,}98 \times \langle 1 - exp(-\alpha \times LAI)\rangle\end{aligned} \tag{8.18}$$

where u, v, and α are parameters to be estimated from measurements for Equation (8.18) since it is different from Equation (8.3). The primary assumption of the final equation to calculate DLI $[(V_f\beta_{air} + (1 - V_f)\beta_{canopy})\sigma T^4]$ is the equality of canopy temperature with air temperature ($T_{canopy} = T_{air}$) (Pomeroy *et al.*, 2009).

A study conducted by Sicart *et al.* (2010) demonstrated that in settings characterized by low sun angles, overcast conditions, or dense forests where tree trunks receive limited direct sunlight, the potential for variation in canopy temperature compared to the surrounding air

temperature is small, therefore the $T_{canopy} = T_{air}$ assumption is valid. Conversely, in sunlit environments where tree trunks, branches, and needles are exposed to direct sunlight, the absorption of shortwave radiation can lead to an increased emission of longwave radiation from the canopy (Musselman & Pomeroy, 2017). Consequently, it is plausible that in certain instances, the fundamental premise underlying Equation (8.18) may not hold. However, the lack of canopy temperature measurements makes this assumption inevitable in the present case study.

8.2.4 ECMWF scheme for DLI data

Hogan and Bozzo (2016) presented the calculation scheme within the ECMWF model [Integrated Forecasting System (IFS)], which is the "Longwave layer-wise emission" scheme to calculate several atmospheric variables, including DLI. The formulation can be shown as the following equation:

$$\frac{dH_{long\downarrow}}{d\delta} = D(-H_{long\downarrow} + F_{i-\frac{1}{2}} + \delta F') \tag{8.19}$$

Where $H_{long\downarrow}$ = radiative flux (W/m^2), δ = optical depth measured downwards from zero at the top of the layer (m), $D = 1.66$ (m^{-1}) is the diffusivity factor accounting for longwave heat transport, $F_{i-\frac{1}{2}}$ is the Planck function above the ith atmospheric layer (W/m^2), and F' ($\frac{dF}{d\delta}$) is the gradient of the Planck function (W/m^2/m) (Hogan & Bozzo, 2016). This 3D (each calculation cell is based on longitude, latitude, and altitude differentiation) modeling of DLI by the ECMWF atmospheric model uses thermodynamic laws for the atmosphere system, which cannot be compared with 1D modeling approaches. However, the result of this model needs to be assessed for inland data-scarce regions for future implementations.

8.2.5 Calibration algorithm

The calibration algorithm is the Covariance Matrix Adaptation Evolution Strategy (CMA-ES), an evolutionary algorithm developed by Hansen and Ostermeier (1996). CMA-ES is a stochastic global optimization method able to solve high-dimensional problems. These benefits make it a good candidate as a general-purpose optimization method (Hansen & Ostermeier, 1996). This algorithm was assessed as a superior candidate for use in hydrology (Arsenault *et al.*, 2014). The only stopping criterion used is the number of evaluations equal to 10,000. The lower and upper boundaries for calibration variables were defined based on recommendations from previous studies that used the same model. Each calibration is repeated ten times to ensure the convergence of the algorithm. The final optimum is the mean of ten iterations.

8.2.6 Calibration objective function and performance metrics

This study used the RMSE as the objective function for all DLI model calibrations for each model's cloud cover factor coefficients. RMSE was used to calibrate the parameters since it is one of the general performance metrics that does not focus on a specific part of the data (Gupta *et al.*, 2009; Legates & McCabe Jr, 1999). At the same time, it is shown that none of the performance metrics perform ideally to show the goodness of fit (Krause *et al.*, 2005). Therefore, several performance metrics were calculated, as Legates and McCabe Jr (1999) and Gauch *et al.* (2023) strongly recommended.

Relative RMSE ($RRMSE$) was used to show each equation's skill dimensionless for calibration and validation steps. Nash–Sutcliffe efficiency (NSE) shows statistically meaningful efficiency of fit, as well as the performance of fit for high values. The correlation (r) and coefficient of determination (R^2) performance metrics were also calculated as recommended by Legates and McCabe Jr (1999). Since using more parameters and input variables can cause overfitting, there is a need to select the most parsimonious model among those of equal performance. Thus, the Akaike information criterion (AIC) (Akaike, 1974) and Bayesian information criterion (BIC) (Schwarz, 1978) were also calculated while considering the number of parameters. Finally, a simple bias performance metric was calculated. A summary of these performance metrics is presented as follows:

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{N}} \tag{8.20}$$

$$RRMSE_j = \frac{RMSE_j}{\bar{O}_j} \times 100 \tag{8.21}$$

$$r_j = \frac{\sum_{i=1}^{N}(S_{i,j} - \bar{S}_j)(O_{i,j} - \bar{O}_j)}{\sqrt{\sum_{i=1}^{N}(S_{i,j} - \bar{S}_j)^2 \sum_{i=1}^{N}(O_{i,j} - \bar{O}_j)^2}} \tag{8.22}$$

$$R_j^2 = \left[\frac{\sum_{i=1}^{N}(S_{i,j} - \bar{S}_j)(O_{i,j} - \bar{O}_j)}{\sqrt{\sum_{i=1}^{N}(S_{i,j} - \bar{S}_j)^2 \sum_{i=1}^{N}(O_{i,j} - \bar{O}_j)^2}}\right]^2 \tag{8.23}$$

$$NSE_j = 1 - \frac{\sum_{i=1}^{N}(S_{i,j} - O_{i,j})^2}{\sum_{i=1}^{N}(O_{i,j} - \bar{O}_j)^2} \tag{8.24}$$

$$Bias_j = \overline{O_j} - \overline{S_j} \tag{8.25}$$

$$AIC_j = N.Ln\left(\frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{N}\right) + 2k + \frac{2k(k+1)}{N-k-1} + N.Ln(2\pi) + N \tag{8.26}$$

$$BIC_j = N.Ln\left(\frac{\sum_{i=1}^{N}\left(S_{i,j} - O_{i,j}\right)^2}{N}\right) + k.Ln(N) + N.Ln(2\pi) + N \tag{8.27}$$

where $RMSE_j$, $RRMSE_j$, r_j, R_j^2, NSE_j, $Bias_j$, AIC_j and BIC_j are the performance metrics for the jth station; $O_{i,j}$ and $S_{i,j}$ respectively are observed and simulated values for the jth station at i^{th} time step; N is the number of time steps (measurements), and k is the number of parameters in the model; $\overline{O_j}$ is the mean of observed values. The AIC_j and BIC_j performance metrics not only use the variance between observed and simulated values but also the number of measurements (N) and the number of model parameters (k) to identify the most parsimonious model (lowest AIC and BIC values).

The reported calibration and validation performance metrics are based on the mean of ten iterations, which helps mitigate the impact of numerical precision and provides uncertainty bounds for parameters and their calculated performance metrics.

8.2.7 Calibration process

As presented in section 8.2.3, several formulations for longwave heat budget estimation exist. In this study, two methods for longwave calculation will be evaluated using ground-based measurements: (1) testing ERA5 DLI data (e.g., “Mean surface downward long-wave radiation flux” based on the ECMWF variable names) as a proxy instead of using DLI models and (2) estimating DLI using nine forms of models for all sky conditions (Table 8.2). Both methods are evaluated with ground-based observations for DLI using two installed sensors. The detailed parametrization is provided in Table 8.2.

These net radiometers used at both stations (Apogee) have four components: two pyranometers that measure incoming (upward-looking) and reflected (downward-looking) shortwave radiation and two pyrgeometers facing upward and downward for longwave radiations. All four radiation fluxes were recorded hourly (Table 8.1). The other measured variable at the stations is air temperature.

Table 8.2 Formulation of the nine DLI estimation models for this study

Model No.	Unit for J	Unit for e_a	The final form of the equation ($J_{All-sky} = (1 + uB^v) \times J_{Clear-sky}$)	Reference for $J_{Clear-sky}$	Reference for $(1 + uB^v)$
1	W/m^2	–	$J = a\sigma T_{air}^4$	Stephen-Boltzmann law for grey body	(Duarte et al., 2006)
2*	mW/cm^2	–	$J = (1 + uB^v) \times 5.31 \times 10^{-14} T_{air}^6$	(Swinbank, 1963; Equation 2)	(Duarte et al., 2006)
3	W/m^2	–	$J = (1 + uB^v)\langle 1 - 0{,}261 \times \exp[-7{,}77 \times 10^{-4}(273 - T_{air})^2]\rangle \sigma T_{air}^4$	(Idso et al., 1969)	(Duarte *et al.*, 2006)
4	W/m^2	$mmHg$	$J = [(1 + 0{,}17B^2)0{,}97 \times (0{,}74 + 0{,}0065e_a)]\sigma T_{air}^4$	(Morin et al., 1990; Equations 11 and 12)	(Morin et al., 1990; Equation 12)
5	W/m^2	hPa	$J = (1 + uB^v)[1{,}24(\frac{e_a}{T_{air}})^{1/7}]\sigma T_{air}^4$	(Brutsaert, 1975; Equation 11)	(Duarte et al., 2006)
6	W/m^2	hPa	$J = (1 + uB^v)\langle 1{,}08[1 - exp(-e_a^{\frac{T_{air}}{2016}})]\rangle \sigma T_{air}^4$	(Satterlund, 1979; Equation 5)	(Duarte et al., 2006)
7	W/m^2	hPa	$J = (1 + uB^v)(1 - \left[1 + 46{,}5\left(\frac{e_a}{T_{air}}\right)\right] exp\langle -\left\{1{,}2 + 3\left[46{,}5\left(\frac{e_a}{T_{air}}\right)\right]\right\}^{1/2}\rangle)\sigma {T_{air}}^4$	(Prata, 1996; Equation 20)	(Duarte et al., 2006)
8	W/m^2	hPa	$J = (1 + uB^v)\begin{cases}[0{,}72 + 0{,}009(e_a - 2)]\sigma {T_{air}}^4 & e_a \geq 2 \\ [0{,}72 - 0{,}076(e_a - 2)]\sigma {T_{air}}^4 & e_a < 2\end{cases}$	(Niemelä et al., 2001; Equation 18)	(Duarte et al., 2006)
9**	W/m^2	$mm\ Hg$	$J = \{(1 + uB^v)[exp(-\alpha \times LAI)]0{,}97 \times (0{,}74 + 0{,}0065e_a) + 0{,}98 \times \langle 1 - exp(-\alpha \times LAI)\rangle\}\sigma {T_{air}}^4$	(Pomeroy et al., 2009; Equation 3; Verseghy et al., 1993; Equation 7)	(Duarte et al., 2006)

* By conversion from mW/cm^2 to W/m^2 the constant is 5.31×10^{-13}. Also, following Stephen-Boltzmann law notation using σT_{air}^4 the term can be written as $(9.3645 \times 10^{-6} T_{air}^2)\sigma T_{air}^4$.

** The canopy transmissivity $exp(-\alpha \times LAI)$ is based on Beer's law of radiation (Verseghy *et al.*, 1993). Pomeroy *et al.* (2009) cited Oke (1987) for the $\beta_{canopy} = 0.98$.

The ERA5 variables were used without any correction as they were assumed as reliable proxy for data-scarce regions or ungauged watersheds over Canada by Tarek *et al.* (2020b) and Yao *et al.* (2020). Also, the efficiency of the direct use of ERA5 products over the Nechako watershed was confirmed by Gatien *et al.* (2022) and Khorsandi *et al.* (2022).

After calibration and validation of the DLI models at SLS, the nine models and ERA5 DLI product were also calibrated/validated separately on the BVESM station to test the methods' performance on a station with different hydroclimatological conditions. For Model 1, one parameter to calibrate is the emissivity as a constant for all sky conditions. For seven models (2,3, 5-9), the parameters for cloud cover were calibrated (u and v). Model 9 has one additional parameter (α). The main constraints for Models 3-9 are the upper and lower boundaries for u ($-1 \leq u \leq 0.5$), and v ($0 \leq v \leq 4$) which are based on the physically possible ranges and the

reported values in the literature. Moreover, based on the observed air temperature ($T_{air,observed}$) and observed DLI ($J_{Observed}$), the observed effective emissivity ($\beta_{Effective}$) range was set as $0.5 \leq \beta_{Simulated} = \frac{J_{Simulated}}{\sigma T_{air,ERA5}^4} \leq 1.5$ for all models. Based on Table 8.2, Model 4 has no parameters to calibrate since parameter values recommended for Canada were used. For Model 9, since there is no observed relationship between LAI and the sky view factor (V_f) at the two stations, lower boundary was considered as $0 \leq \alpha$ to keep the exponential relationship between the possible physical boundaries ($0 \leq V_f \leq 1$).

8.2.8 Experimental uncertainty analysis

In this research, an uncertainty analysis was conducted to evaluate the calibrated equations' reliability. This analysis used partial derivatives of $J = \beta\sigma T^4$, a fundamental tool in quantifying the propagation of uncertainties through implementing this equation (Ku, 1966). The uncertainty analysis enables a more comprehensive understanding of how input parameter variations influence the results' uncertainty. The estimate of the uncertainty for DLI (ΔJ) based on $J = \beta\sigma T^4$ and $\beta = f(T, e_a, B)$ can be written as (Ku, 1966):

$$\Delta J = \pm\sqrt{\left(\left[\frac{\partial J}{\partial \beta}\frac{\partial \beta}{\partial T} + \frac{\partial J}{\partial T}\right]\Delta T\right)^2 + \left(\left[\frac{\partial J}{\partial \beta}\frac{\partial \beta}{\partial e_a}\right]\Delta e_a\right)^2 + \left(\left[\frac{\partial J}{\partial \beta}\frac{\partial \beta}{\partial B}\right]\Delta B\right)^2} \tag{8.28}$$

where ΔJ, ΔT, Δe_a, and ΔB are estimation uncertainty (errors) in J, T, e_a, and B variables.

Using the measured DLI and air temperatures, effective emissivity ($\beta_{Effective}$) can be calculated as:

$$\beta_{Effective} = \frac{J_{obs}}{\sigma T_{obs}^4} \tag{8.29}$$

In a similar way, the uncertainty approximation for $\beta_{Effective}$ is:

$$\Delta\beta_{Effective} = \pm\sqrt{\left(\frac{\partial \beta_{Effective}}{\partial J_{obs}}\Delta J_{obs}\right)^2 + \left(\frac{\partial \Delta\beta_{Effective}}{\partial T_{obs}}\Delta T_{obs}\right)^2} \tag{8.30}$$

where ΔJ_{obs} and ΔT_{obs} are measurement errors for measured DLI (W/m^2) and measured air temperature (°C). The device's accuracy in measuring DLI (ΔJ_{obs}) is ±5% ($\pm 0.05 \times J_{obs}$ in W/m^2 measurment unit) based on the sensor specifications (www.apogeeinstruments.com).

8.3 Results

The presented results comprise a comparative analysis of DLI derived from the ERA5 reanalysis database and the nine models detailed in Table 8.2. Each model demonstrates the influence of air temperature (T_{air}), vapor pressure (e_a), and cloud cover (B) on the estimation of measured DLI, employing distinct formulations using one or more of these variables.

8.3.1 Model calibration

The RMSE for the calibration period for the two stations are presented in Table 8.3, along with other performance metrics. Table III-B1 in Appendix III-B shows the calibrated parameters for the nine models. For calibration convergence and to assess algorithmic uncertainty (including software, calibration algorithm, and numerical precision), CMA-ES calibration was iterated ten times per station and DLI model. Table III-B1 in Appendix III-B shows the mean value, standard deviation, and the resulting coefficient of variation (CV) for each parameter for each DLI model. Table III-B1(c) shows how much each calibrated parameter is sensitive to the algorithmic uncertainty.

Calibration performance metrics for each DLI model and both stations are determined from the ten-iteration means (Table 8.3) and visualized in scatter plots (Figure 8.2 for SLS and Figure 8.3 for BVESM). We calculated the mean value, standard deviation, and CV for each performance metric of each DLI model (see Table III-B3 and B4 in Appendix III-B for SLS and BVESM performance metrics).

Table 8.3 Calibration performance metrics for ERA5 data and nine DLI models compared with observed data at SLS and BVESM stations (10 Calibration Iterations). Best candidates are in bold, and alternatives are underlined and numbered.

Station	Model	RMSE (W/m^2)	RRMSE	Correlation	R^2	NSE	Bias (W/m^2)	AICc	BIC
SLS	**ERA5**	**32.02**	**11.81**	**0.78**	**0.60**	**0.55**	5.45	**101815.95**	**101815.95**
SLS	1	37.58	13.87	0.62	0.39	0.38	**0.48**	105155.79	105163.04
SLS	2	37.02	13.66	0.73	0.53	0.40	4.22	104844.14	104851.39
SLS	3	33.42	12.33	0.73	0.53	0.51	-0.66	102709.28	102716.53
SLS	4[1]	32.07	11.84	0.75	0.56	0.55	-0.82	101853.98	101853.98
SLS	5	35.92	13.25	0.77	0.59	0.44	5.12	104214.90	104222.15
SLS	6	33.57	12.39	0.75	0.56	0.51	-2.27	102804.42	102811.67
SLS	7	32.37	11.94	0.76	0.58	0.54	1.80	102045.58	102052.83
SLS	8	32.53	12.00	0.75	0.56	0.54	-1.15	102149.12	102156.37
SLS	9[2]	32.29	11.92	0.74	0.55	0.54	0.67	101998.36	102005.61
BVESM	**ERA5**	**30.37**	**9.89**	**0.88**	**0.78**	**0.76**	4.98	**49680.61**	**49680.61**
BVESM	1	36.02	11.73	0.81	0.66	0.66	**-0.22**	51436.68	51443.22
BVESM	2	38.36	12.49	0.85	0.72	0.61	6.69	52082.47	52089.02
BVESM	3	35.29	11.50	0.84	0.70	0.67	3.11	51226.52	51233.06
BVESM	4	32.02	10.43	0.87	0.75	0.73	6.55	50222.84	50222.84
BVESM	5	37.74	12.29	0.88	0.78	0.63	8.75	51916.14	51922.69
BVESM	6[2]	31.66	10.31	0.88	0.77	0.74	2.55	50111.04	50117.58
BVESM	7	32.80	10.69	0.88	0.77	0.72	4.70	50474.32	50480.86
BVESM	8	34.40	11.21	0.86	0.74	0.69	2.50	50963.36	50969.91
BVESM	9[1]	30.76	10.02	0.87	0.75	0.75	0.34	49812.18	49818.72

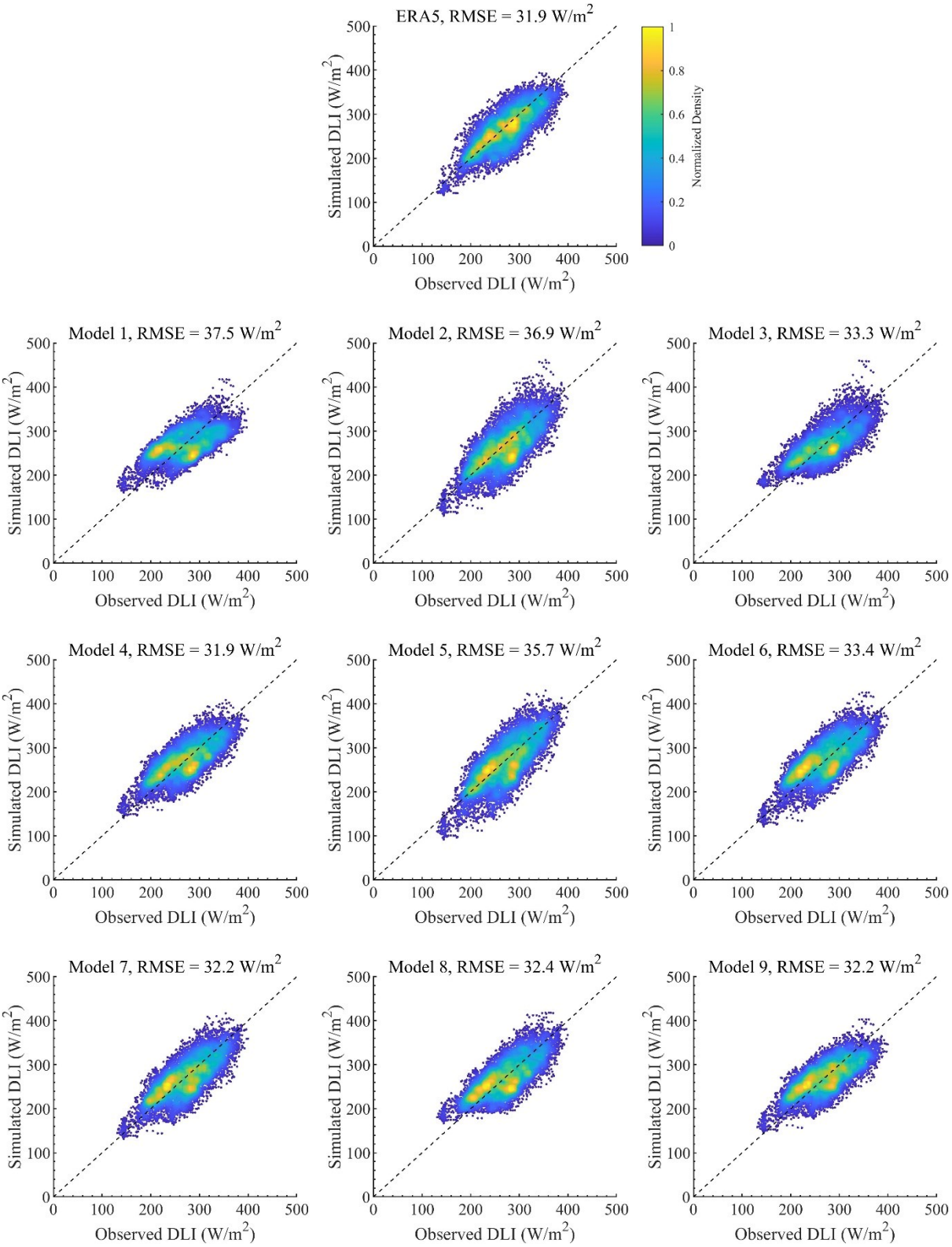

Figure 8.2 Scatter plots and RMSE for ERA5 DLI and nine DLI models in SLS station calibration

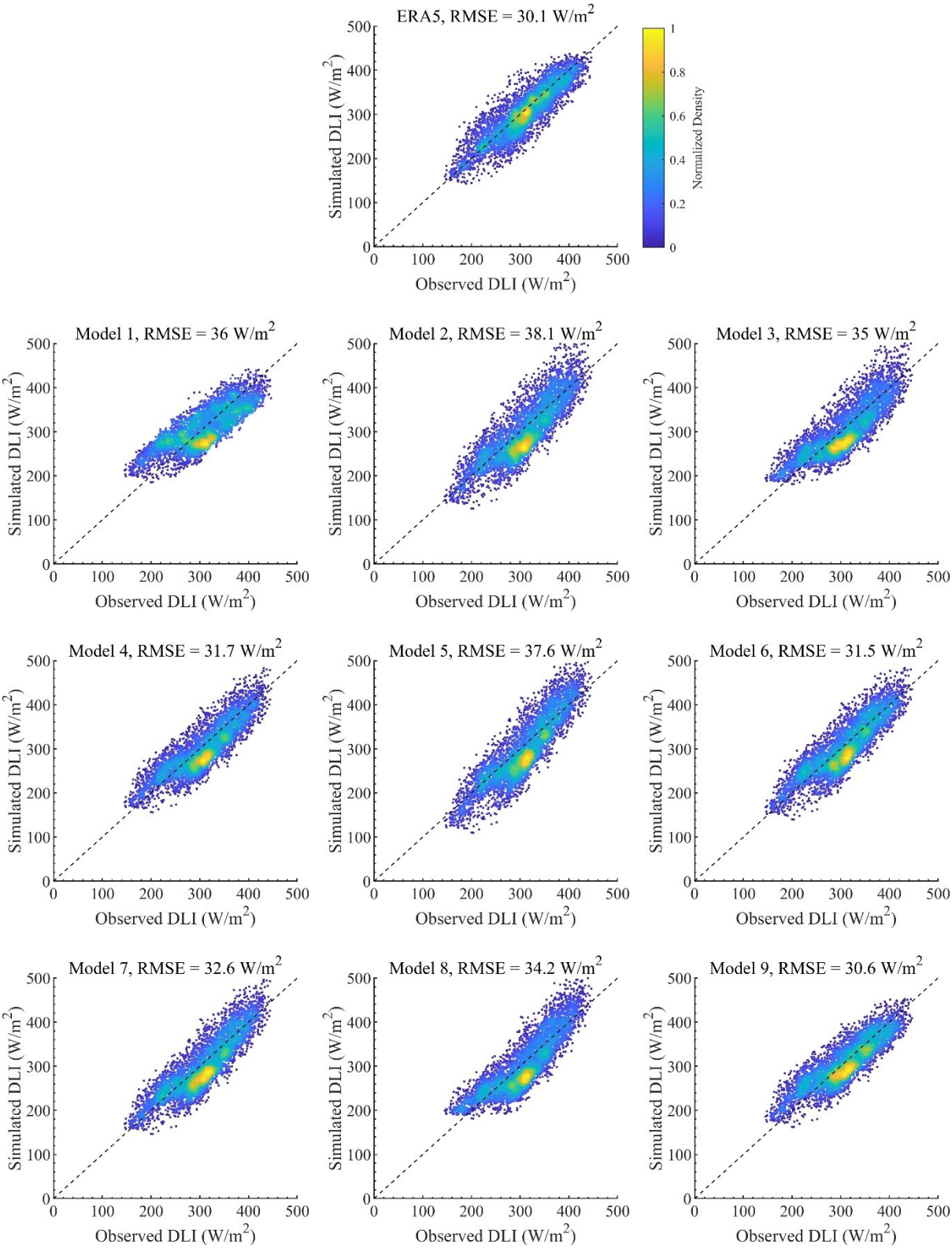

Figure 8.3 Scatter plots and RMSE for ERA5 DLI and nine DLI models in BVESM station calibration

Model 1, using a simple constant emissivity formulation for all sky conditions, serves as a benchmark for efficiency comparison across DLI models. For SLS station calibration, Model 2 exhibits a higher RMSE value (37 W/m²) and lower NSE (0.4) compared to other models,

ranking it similar to Model 1 (RMSE = 37.6 W/m²). Similarly, for BVESM, Model 2 shows RMSE = 38.4 W/m² (NSE = 0.61), performing less effectively than Model 1 (RMSE = 36.0 W/m², NSE = 0.66). The lower performance of Model 2 is attributed to its air emissivity formulation, discussed in the following section.

The remaining eight DLI estimation options at SLS exhibit small variability in RMSE values, ranging from a minimum of 32.0 W/m² (RRMSE = 11.8% for ERA5) to a maximum of 35.9 W/m² (RRMSE = 13.3% for Model 5). Notably, Model 4 and its updated version accounting for canopy effects based on LAI (Model 9) yield similar results, suggesting that canopy effects have limited impact at this station, resulting in comparable performance between these two models.

For the BVESM station calibration, the eight DLI estimation options (ERA5, models 3, 4, 5, 6, 7, 8, 9) exhibit similar RMSE results, ranging from a minimum of 30.4 W/m² (RRMSE = 9.9% for ERA5 data) to a maximum of 38.4 W/m² (RRMSE = 12.5% for Model 2). This similarity among the models (Table 8.3) is expected, given their foundation in physical radiation laws and reliance on variables influencing air emissivity. Notably, ERA5 outperforms all models at both SLS and BVESM, probably because of its comprehensive 3D radiation formulation. At the same time, it is calculating DLI on a 31 km resolution grid, so it cannot have local properties correctly modeled. The observed RMSE range (30.4-38.4 W/m²) aligns with findings by Long *et al.* (2021), who reported a mean RMSE of 26.4 W/m² (ranging from 23.2 to 42 W/m²) for all sky conditions in the Tibetan Plateau.

Furthermore, we computed additional performance metrics for each model, providing several criteria for comparison (Table 8.3). The RRMSE indicates that all DLI options perform well with values consistently below 15%, aligning with previous findings (Long *et al.*, 2021). Specifically, for the calibration step at the SLS station, direct use of ERA5 (without the need for calibration) exhibits an RRMSE of 11.8%, while at BVESM, it achieves 9.9%. NSE, which penalizes discrepancies in high values within a time series, reveals ERA5's superior performance (NSE = 0.55 for SLS and 0.76 for BVESM), followed by methods 4 and 9 (SLS; NSEs = 0.55 and 0.54 respectively) and, 9 and 6 (BVESM; NSEs = 0.75 and 0.74 respectively).

AIC and BIC values reveal ERA5 as the best option for both stations, with the minimum values. Following ERA5, Model 4 performs best for SLS, while Model 9 shows good performance at BVESM (Table 8.3). Notably, Model 9, incorporating an additional term related to LAI compared to Model 4, demonstrates superior performance metrics at BVESM, where the canopy's impact is more pronounced. For SLS, the performance metrics of Models 4 and 9 are nearly identical.

8.3.2 Model validations

With the calibrated parameters and DLI models from the calibration step on SLS and BVESM, the validation results are calculated on both stations and are presented in Table 8.4.

Table 8.4 Validation performance metrics for ERA5 data and nine DLI models compared with observed data at SLS and BVESM stations. Best candidates are in bold, and alternatives are underlined and numbered.

Station	Model	RMSE (W/m^2)	RRMSE	Correlation	R^2	NSE	Bias (W/m^2)	AICc	BIC
SLS	**ERA5**	**32.06**	**11.82**	**0.77**	**0.60**	**0.54**	5.46	**67483.29**	**67483.29**
SLS	1	37.60	13.86	0.62	0.38	0.37	0.65	69685.08	69691.92
SLS	2	37.18	13.71	0.72	0.52	0.39	4.32	69531.92	69545.60
SLS	3	33.53	12.36	0.72	0.52	0.50	**-0.42**	68106.42	68120.10
SLS	4[1]	32.23	11.88	0.74	0.55	**0.54**	-0.73	67560.95	67574.62
SLS	5	36.12	13.32	0.76	0.58	0.42	5.04	69132.48	69146.16
SLS	6	33.71	12.43	0.74	0.55	0.50	-2.21	68178.47	68192.15
SLS	7	32.56	12.00	0.76	0.57	0.53	1.90	67699.29	67712.97
SLS	8	32.68	12.05	0.74	0.55	0.53	-0.95	67750.93	67764.61
SLS	9[2]	32.39	11.94	0.73	0.54	0.53	0.82	67631.05	67651.57
BVESM	**ERA5**	**30.72**	**10.01**	**0.88**	**0.78**	**0.76**	5.64	**33499.14**	**33499.14**
BVESM	1	36.25	11.81	0.81	0.66	0.66	**-0.02**	34646.41	34652.55
BVESM	2	38.74	12.62	0.85	0.72	0.62	7.25	35108.49	35120.78
BVESM	3	35.71	11.63	0.84	0.70	0.67	3.54	34546.03	34558.32
BVESM	4	32.48	10.58	0.87	0.75	0.73	6.97	33889.36	33901.65
BVESM	5	38.04	12.39	**0.88**	**0.78**	0.63	9.21	34982.78	34995.07
BVESM	6[2]	31.95	10.41	**0.88**	0.77	0.74	3.00	33775.26	33787.56
BVESM	7	33.16	10.80	**0.88**	0.77	0.72	5.13	34032.47	34044.77
BVESM	8	34.77	11.33	0.86	0.74	0.69	2.83	34360.57	34372.87
BVESM	9[1]	31.09	10.13	0.87	0.75	0.75	0.92	33589.91	33608.35

Validation results for the SLS station yield performance metrics and conclusions mirroring the calibration phase, with the same models ranked in the same order. Thus, ERA5 and Model 4 provide the most reliable estimates for the SLS station in the Nechako watershed. Scatter plots illustrating the validation of the nine DLI models can be found in Figure 8.4.

Similarly, the validation results for the BVESM station closely resemble the calibration outcomes, with ERA5 and Model 9 remaining as the top choices. Scatter plots for model validation at the BVESM station are presented in Figure 8.5 for result interpretation.

Model 9, incorporating the LAI variable, performs notably better at the BVESM station, with relatively high canopy cover, whereas Model 9 produces results on par with Model 4 at the SLS station.

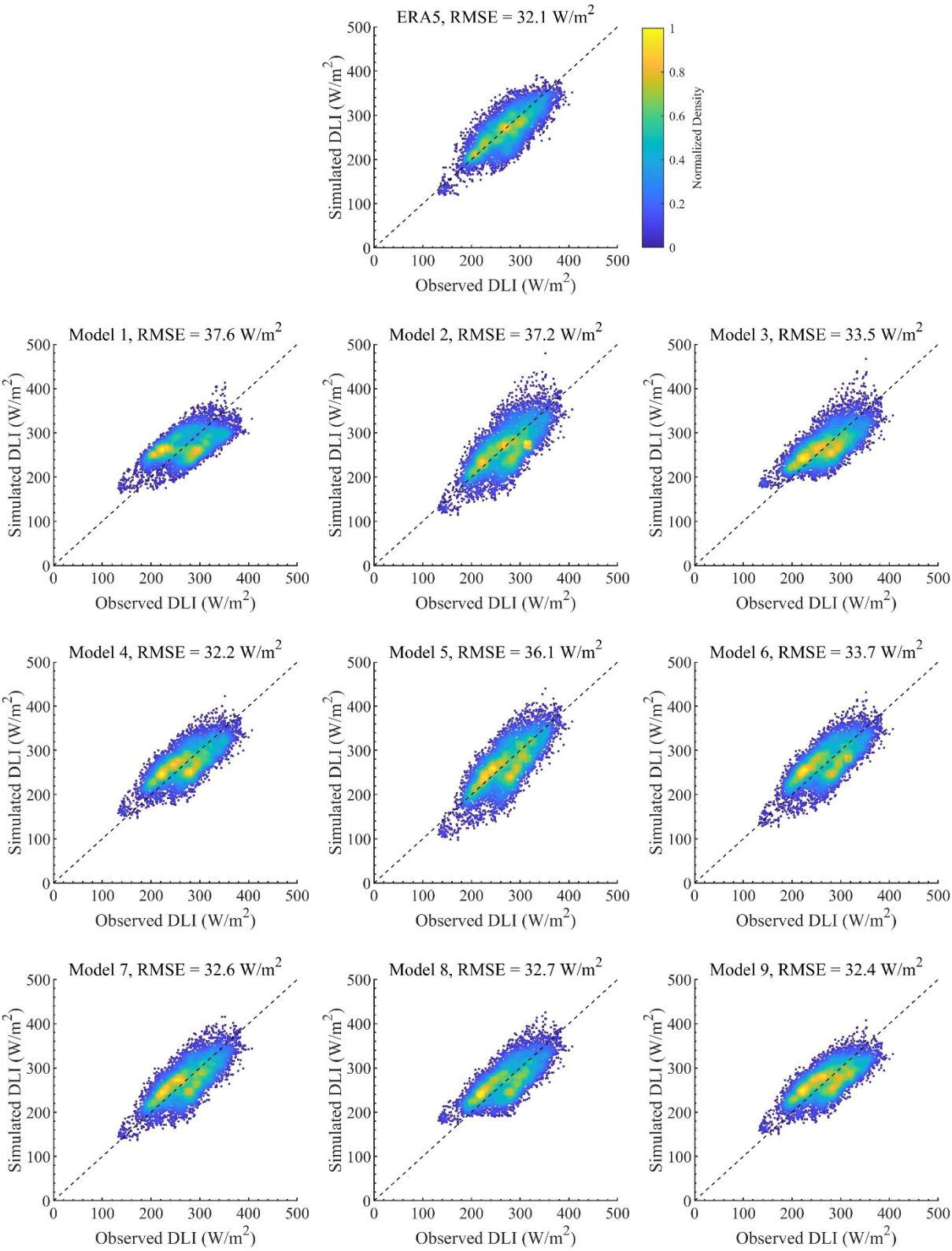

Figure 8.4 Scatter plots and RMSE for ERA5 DLI and nine DLI models in SLS station validation

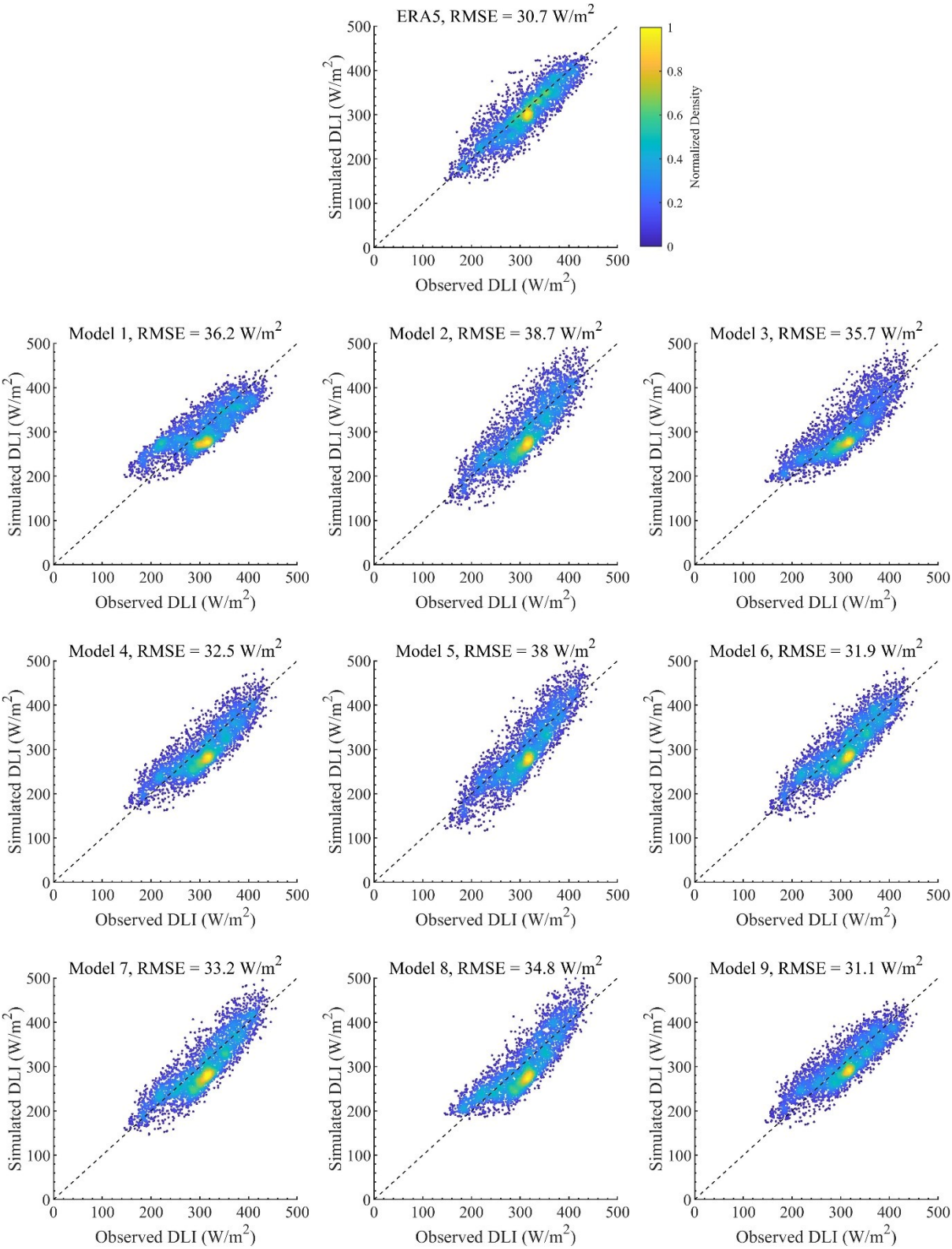

Figure 8.5 Scatter plots and RMSE for ERA5 DLI and nine DLI models in BVESM station validation

The results demonstrate that ERA5 data outperform other options at the BVESM station, even in the presence of a nearby dense canopy.

8.3.3 Data measurement error and propagation using the ERA5 dataset

Not all climate variables are measured at each site. At the SLS and BVESM stations, only air temperatures, aside from the DLI radiative variable, are measured. Consequently, beyond DLI, it is feasible to assess the reliability of using ERA5-based air temperatures instead of measured air temperatures.

The sensors employed for measuring air temperature in research settings, as in the cases of SLS and BVESM, provide an accuracy of ±0.5°C, which covers the natural temperature range of Canada (-40 to +30°C) (source: www.apogeeinstruments.com). In contrast, ERA5 data involves a reanalysis of observed data and global-scale land-atmosphere models for a ≈ 900 km^2 grid cell, whereas the station has a very local measurement depending on its exact environment. ERA5's near-surface temperatures and observed temperatures were compared (Figure 8.6). While the value of RMSE performance metrics are high (3°C and 3.7°C for SLS and BVE, respectively), they are comparable to error values reported in a similar study in the Tibetan Plateau (1.4 to 5.2°C for five stations; Long et al., 2021).

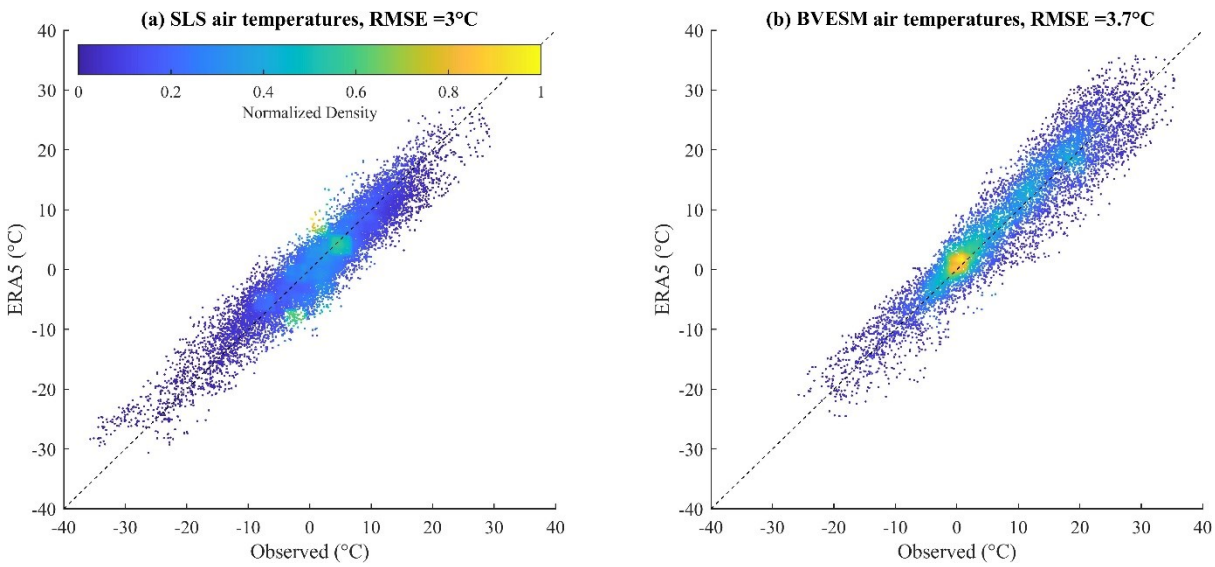

Figure 8.6 **Scatter plots comparing ERA5's 2 m air temperatures with hourly observed measurements at (a) SLS and (b) BVESM stations, along with their respective RMSE performance metrics.**

One of the critical meteorological variables for DLI modeling is relative humidity (or vapor pressure), which is notably absent from the ERA5 dataset for near-surface conditions (2 m or 10 m height). In the absence of direct vapor pressure data, an alternative approach involves utilizing dew point temperature to estimate vapor pressure for DLI modeling. It is worth noting that neither of the two stations has observed data for dew point temperature. However, it is

reasonable to anticipate a similar level of accuracy between measured and ERA5-based data for dew point air temperatures, with an expected accuracy of approximately 3-4°C.

The second measured variable in this study is DLI itself. As per sensor specifications, the measurements come with a ±5% error (equivalent to ±22 W/m² using the observed data) (source: www.apogeeinstruments.com). Using Equation (8.29), the $\beta_{Effective}$ varies from 0.59 to 1.13. When employing Equation (8.30) to account for measurement uncertainty in $\beta_{Effective}$, (considering ±5°C uncertainty in T_{obs}, can be written as $\pm\sqrt{\left(\frac{0.05}{\sigma T_{obs}^4}J_{obs}\right)^2+\left(\frac{2}{\sigma T_{obs}^5}J_{obs}\right)^2}$) uncertainty results in ±0.057 (±7%). However, for all the other climatic variables utilized in this study, there is no observed value available to estimate the accuracy of the calculations. The primary motivation for using ERA5 data rather than land-based measurements stems from this data deficiency.

In their study, Long *et al.* (2021) provided error metrics for air temperature (T) and vapor pressure (e_a) measurements at five stations in China, as well as modeling uncertainties for DLI models (ΔJ). Using RMSE as the error metric, they found that the errors were within the following ranges: $1.4℃ \leq |\Delta T| \leq 5.2℃$, $0.8\,hPa \leq |\Delta e_a| \leq 2.3\,hPa$, and $23.2\,W/m^2 \leq |\Delta J| \leq 42\,W/m^2$. They also calculated the attributed error in DLI estimation due to air temperature ($\left[\frac{\partial J}{\partial \beta}\frac{\partial \beta}{\partial T}+\frac{\partial J}{\partial T}\right]\Delta T$), which ranged from 3.1 to 12.5 W/m². Similarly, the attributed error in DLI estimation due to vapor pressure ($\left[\frac{\partial J}{\partial \beta}\frac{\partial \beta}{\partial e_a}\right]\Delta e_a$) varied from 4.8 to 13.8 W/m².

It is important to note that due to the absence of data measurements for cloud cover, they did not assess the attributed error in DLI estimation due to cloud cover estimation error ($\left[\frac{\partial J}{\partial \beta}\frac{\partial \beta}{\partial B}\right]\Delta B$). Instead, they used cloud correction factors (including the one in this study) to adjust clear-sky DLI estimations for all sky conditions. They estimated cloud cover using the Crawford and Duchon equation (Crawford & Duchon, 1999) and limited their analysis to daytime situations. With daytime data and cloud correction factors, the mean RMSE for all sky conditions was 26.4 W/m², varying from 23.2 to 42 W/m².

To quantify this overall error in our study, we can use the DLI equation as follows, taking Model 4 as an illustrative example, considering its wide use in Canada and its strong performance in our study for both stations:

$$\begin{aligned}\Delta J^2 = {} & ([4(1+0.17B^2)\times 0.97\times(0.74+0.0065e_a)\sigma T^3]\Delta T)^2 \\ & + ([0.97\times 0.0065(1+0.17B^2)\sigma T^4]\Delta e_a)^2 \\ & + ([0.34B\times 0.97\times(0.74+0.0065e_a)\sigma T^4]\Delta B)^2\end{aligned} \tag{8.31}$$

Considering $\Delta T = 3°C$ by using ERA5 data at SLS, the attributed error in J estimation due to air temperature errors ($[4(1 + 0.17B^2) \times 0.97 \times (0.74 + 0.0065e_a)\sigma T^3]\Delta T$) varies from 7 to 17 W/m^2, for Model 4 (values are 7.7 to 19.6 W/m^2 for BVESM). By assuming the same level of error in e_a as Long *et al.* (2021) ($|\Delta e_a| \leq 2.3\ hPa$ or $|\Delta e_a| \leq 1.73\ mmHg$), the maximum attributed error in J estimation due to e_a errors ($[0.0065 \times 0.97 \times (1 + 0.17B^2)\sigma T^4]\Delta e_a$) varies from 2.2 to 6.1 W/m^2 for Model 4 for SLS (2.4 to 6.4 W/m^2 for BVESM).

The attributed error stemming from cloud cover comprises several components that add to its complexity. Firstly, there is the cloud cover measurement error. Secondly, this study involves an uncertainty associated with cloud correction formulations. These formulations are based on empirical research from previous studies. Lastly, the error associated with cloud effects is particularly uncertain when compared to other variables. This uncertainty arises because various types of clouds can be simultaneously present, each exerting several influences on DLI. As a result, a single numerical value representing cloud cover does not adequately capture the aggregate effects on DLI.

The overall error propagation resulting from the above explanation is challenging to quantify and remains uncertain. To provide a simplified estimate, we can focus on the first element, which is the error in cloud cover measurements, while temporarily disregarding the other two error sources. The attributed error due to cloud cover for Model 4 ($[0.34B \times 0.97 \times (0.74 + 0.0065e_a)\sigma T^4]\Delta B$) can be assessed, assuming $|\Delta B| < 0.06$ for the most precise cloud cover sensors (source: https://www.atmos-meteo.com/meteorology/cloud-cover-sensors.html). This simplification results in an error ranging from 0 to 7.2 W/m^2 for SLS (0 to 8.7 W/m^2 for BVESM).

It is important to note that this estimated error related to cloud cover is highly simplified and underestimates the observed variation in DLI caused by cloud cover effects (see Figure 8.7 and Figure 8.8, which highlight the impact of clouds by comparing all sky and clear-sky conditions at SLS and BVESM). For instance, assuming a typical uncertainty of 20% for cloud cover (McMillan *et al.*, 2018), the uncertainty attributed to cloud cover would increase to 24 W/m^2 for SLS (28 W/m^2 for BVESM).

Considering all the assumptions outlined above, the overall DLI estimation for Model 4 [Equation (8.33)] ranges from 7.3 to 19.4 W/m^2 (calculated as $\sqrt{7^2 + 2.2^2 + 0^2} = 7.3$ to $\sqrt{17^2 + 6.1^2 + 7.2^2} = 19.4\ W/m^2$). Consequently, when using Model 4, ERA5 input data, and accounting for errors in air temperature (T) and vapor pressure (e_a) as discussed earlier, it becomes evident that T is the primary source of uncertainty in J estimation. For Model 4 as the DLI formulation, the errors stemming from e_a and B correction factors exhibit a similar magnitude, ignoring extra uncertainties in B measurements and impacts.

It is essential to acknowledge that the error analysis outlined above relies on several simplifying assumptions due to the lack of measured data, and thus, it likely underestimates the actual potential uncertainty. Nonetheless, it underscores the significance of DLI estimations, mainly when assessed using the RMSE performance metric for Model 4 under all sky conditions based on observed DLI values at SLS, which amount to 32.3 and 31.6 W/m^2 for calibration and validation, as detailed in Tables 8.3 and 8.4.

8.4 Discussion: Unveiling the complexities of DLI modeling

This section explores sources of error and their effects on DLI modeling. Furthermore, the physical principles underpinning the selected models are addressed.

8.4.1 Physical underpinnings of DLI formulations

The characterization of $\beta_{Effective}$ based on observed air temperature, DLI, and the Stefan-Boltzmann law, allow to investigate its variation as a function of air temperature, vapor pressure, and cloud cover from the ERA5 reanalysis dataset with relatively high accuracy (±0.057 absolute or ±7% relative).

By initially focusing on clear-sky conditions, the influence of T_{air} and e_a becomes evident (Figure 8.7a,b,c and Figure 8.8a,b,c). The observed patterns displayed in Figure 8.7a,b,c for SLS and Figure 8.8a,b,c for BVESM closely resemble findings by Niemelä *et al.* (2001) for Finland. As previously discussed, it is apparent that T_{air} exerts a more substantial impact on the $\beta_{Effective}$ value than e_a. However, Figure 8.7 reveals that under clear-sky conditions at SLS, $\beta_{Effective}$ fluctuates within a narrow range, spanning from 0.6 to 0.8 (mean = 0.69, variance = 0.0046) and a limited spread across e_a values. Notably, most clear-sky conditions at SLS coincide with low vapor pressure (typically less than 5 mmHg, Figure 8.7b) and encompass temperatures ranging from -10°C to 10°C, indicating that more humid periods tend to align with cloudy sky conditions at SLS. Consequently, for the SLS station, a straightforward estimation of $\beta_{Effective}$, represented by a nearly constant value of approximately 0.69, is sufficiently indicative of the $\beta_{Effective}$ behavior under clear-sky conditions.

Figure 8.8 presents a larger range of $\beta_{Effective}$ values under clear-sky conditions at BVESM when compared to SLS, particularly in relation to e_a values. Here, $\beta_{Effective}$ fluctuates between 0.6 and 0.9, accompanied by a higher degree of variability than observations at SLS (mean = 0.75, variance = 0.0057). Notably, clear-sky conditions at BVESM encompass variable vapor pressures, ranging from low to very high, and encompass a wide range of temperature variations, spanning from -20°C to 30°C. BVESM station experiences comparatively warmer conditions than the SLS station due to its lower latitude and lower mean

elevation. More humid conditions can also be linked to the microclimate effects of surrounding canopies. Consequently, BVESM exhibits a broader thermal range compared to SLS, as depicted in Figure 8.6, and it demonstrates higher observed DLI values, as depicted in Figure 8.2.

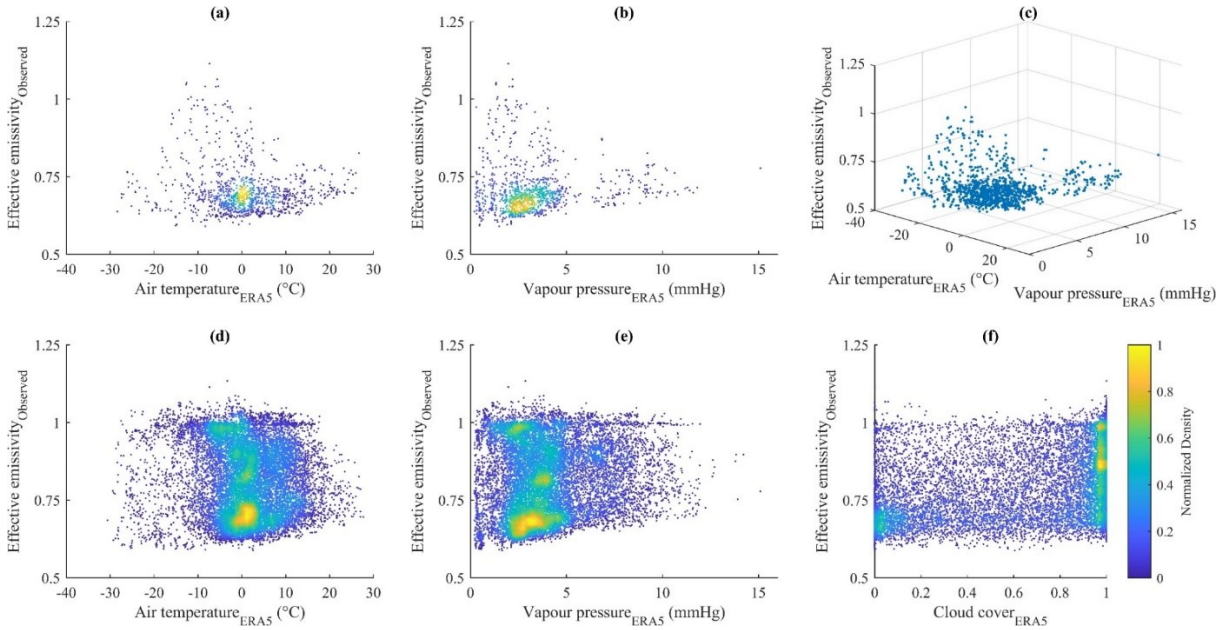

Figure 8.7 **Scatter plots of $\beta_{Effective}$ vs. climatic variables (T, e_a, and B) from ERA5 dataset for (a, b, and c) clear-sky conditions and (d, e, and f) all sky conditions for the SLS station**

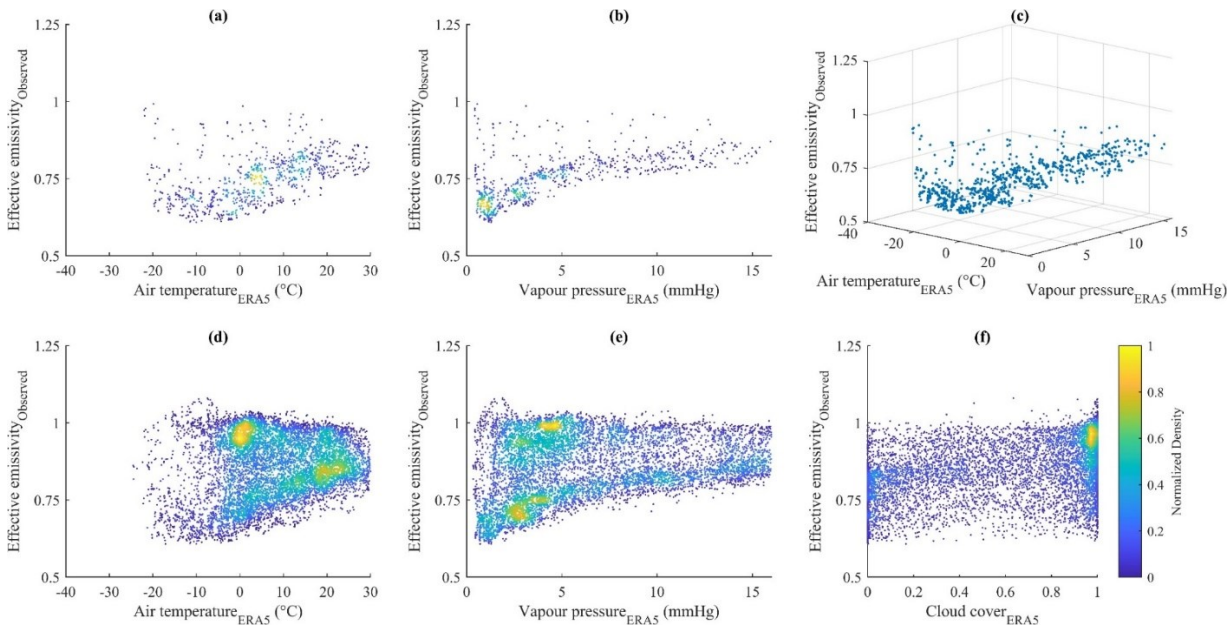

Figure 8.8 **Scatter plots of $\beta_{Effective}$ vs. climatic variables (T, e_a, and B) from ERA5 dataset for (a, b, and c) clear-sky conditions and (d, e, and f) all sky conditions for the BVESM station**

When examining all sky conditions at SLS (Figure 8.7f), the observed patterns for $\beta_{Effective}$ during cloudy-skies align with findings by Niemelä *et al.* (2001) in Finland and partially resemble the mathematical patterns proposed by Maykut and Church (1973) based on observations in Alaska under similar conditions.

It remains evident that most temperatures range from -10°C to 10°C, and vapor pressure typically registers below 5 mmHg. However, under all sky conditions, $\beta_{Effective}$ exhibits high variations, ranging from 0.6 to 1.1 (mean = 0.82, variance = 0.0137). Despite the high variability in $\beta_{Effective}$ attributable to T_{air}, e_a, and B, observations primarily cluster around 0°C, accompanied by vapor pressure levels of 2-4 mmHg and a mean $\beta_{Effective}$ value of 0.82. This mean value closely aligns with the calibration result for Model 1, which utilizes a simple constant value of $\beta = 0.82$. The similarity between this constant and the mean $\beta_{Effective}$ explains why Model 1 can estimate observed DLI with an RMSE of approximately 38 W/m^2. This accuracy level is consistent with previous studies' findings (de Kok *et al.*, 2020; Long *et al.*, 2021; Tang *et al.*, 2021).

Examining all sky conditions at the BVESM station, the pattern closely resembles the findings of de Kok *et al.* (2020) and Sicart *et al.* (2010), aligning with the parameterization in Model 5 (Brutsaert, 1975). However, it must be noted that Model 5 exhibits low performance, particularly for scenarios with low vapor pressures. This deficiency is effectively addressed by the refined formulations presented in Models 6 and 7.

Among all the models considered, Models 2 and 3 estimate β solely based on a single variable, e.g., air temperature. Model 3 employs a Gaussian relationship between T_{air} and β, centered around $T_{air} = 273K\ (\approx\ 0°C)$. In contrast, Model 2 derives β from T_{air} using a parabolic relationship without assuming symmetry. The implementation of Model 3 results in the anticipation of a minimum value for β around 0°C under clear-sky conditions. This behavior aligns with the observed patterns in Figure 8.7a and Figure 8.8a. Model 2 predicts a positive correlation between T_{air} and β across all temperature ranges, a trend not reflected in the observations presented in Figure 8.7a and Figure 8.8a. When considering the use of T_{air} as the sole determinant for β, Model 3 outperformed Model 2 (refer to Table 8.3 and Table 8.4).

Figure 8.7b substantiates the findings of Niemelä *et al.* (2001), revealing a clear-sky condition where $\beta_{Effective}$ assumes its minimum value for $e_a = 2\ hPa$ $(1.5\ mmHg; T_{air} = -10°\mathrm{C})$. Given the estimation accuracy for both variables based on ERA5 data, they introduced Equation (8.11), which establishes a linear relationship between e_a and β, subsequently incorporated into Model 8. Our study's results corroborate their interpretation of $\beta_{Effective}$ under clear-sky conditions. However, when considering the influence of cloud cover effects on DLI, this

previously observed relationship becomes less discernible (as illustrated in Figure 8.7e and Figure 8.8), as the dominant factor becomes the impact of cloud cover on $\beta_{Effective}$.

Morin et al. (1990) established a linear relationship between β and e_a across all temperatures for clear-sky conditions. This linear relationship, represented by Equation (8.7) and incorporated into Model 4, simplifies the interplay between e_a and β but does not account for the negative correlation observed when $e_a < 2\ hPa$ $(1.5\ mmHg)$. Observations at BVESM suggest that considering a linear relationship between $\beta_{Effective}$ and e_a is associated with negligible errors. However, when these models are applied to cloudy conditions, the effects of cloud cover take precedence. This outcome is reflected in the RMSE values for the SLS station (RMSE = 32.1 W/m^2 for Model 4 and RMSE = 32.5 W/m^2 for Model 8). Conversely, results for the BVESM station indicate slightly better performance for Model 4 compared to Model 8 (RMSE = 32.0 W/m^2 for Model 4 and RMSE = 34.4 W/m^2 for Model 8). A clear demonstration of Model 4's superiority over Model 8 under all sky conditions can be observed through a comparison of NSE or AIC between the two models (as seen in Table 8.3 and Table 8.4). The superior performance of Model 4 suggests that a linear relationship between e_a and β, corrected using cloud cover data, leads to enhanced model accuracy.

Models 5, 6, and 7 incorporate non-linear relationships between e_a, T_{air} and β, wherein β increases with rising e_a. These models are designed to adhere to the three physical constraints outlined by Satterlund (1979). Models 5 and 7 employ the term e_a/T_{air} and an exponential relationship in their formulations to satisfy these constraints. By plotting $\beta_{Effective}$ against e_a/T_{air} values, we obtain Figure 8.9 for both clear-sky and all sky conditions at SLS, and Figure 8.10 for BVESM.

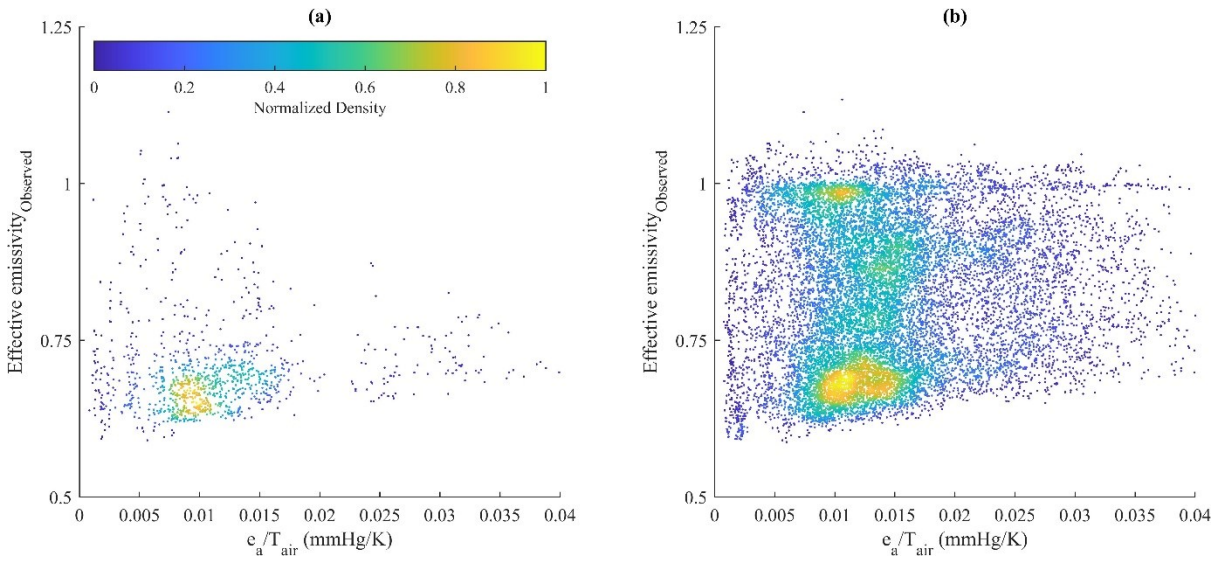

Figure 8.9 **Scatter plots between the measured effective emissivity ($\beta_{Effective}$) and (e_a/T_{air}) from the ERA5 dataset for (a) clear-sky conditions and (b) all sky conditions at the SLS station**

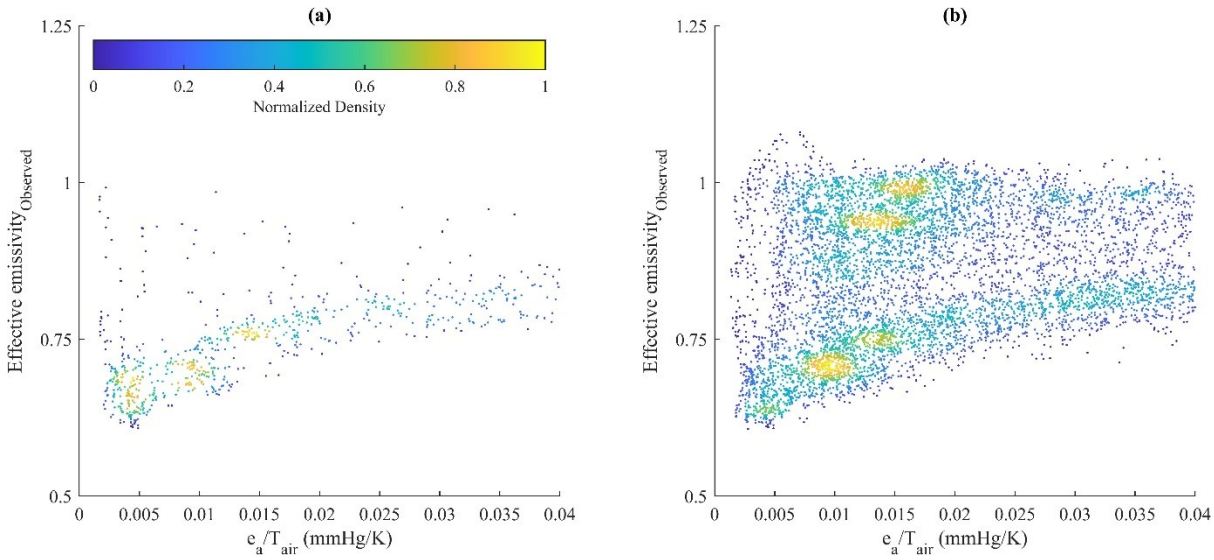

Figure 8.10 **Scatter plots between the measured effective emissivity ($\beta_{Effective}$) and (e_a/T_{air}) from the ERA5 dataset for (a) clear-sky conditions and (b) all sky conditions at the BVESM station**

Figure 8.9a for SLS exhibits limited similarity with the observations made by de Kok *et al.* (2020) for clear skies. This similarity primarily holds for conditions with $e_a/T_{air} > 0.005\, mmHg/K$. However, it becomes challenging to validate the findings of de Kok *et al.* (2020) for clear skies at SLS since vapor pressures exceeding 5 mmHg are almost always associated with cloudy-skies. In contrast, BVESM, with a greater number of clear-sky observations, supports the observed pattern between $\beta_{Effective}$ and e_a/T_{air} proposed by de Kok *et al.* (2020).

Our observations for $\beta_{Effective}$ at SLS ($e_a/T_{air} > 0.005\ mmHg/K$) and BVESM, in conjunction with the findings of de Kok *et al.* (2020), align with the Brutsaert (1975) relationship. This relationship is employed in Model 5. However, it did not yield satisfactory results compared to the other available options (Table 8.3 and Table 8.4).

Simultaneously, Figure 8.9a and Figure 8.10 reveal that the observed $\beta_{Effective}$ values at SLS and BVESM, corresponding to $e_a/T_{air} < 0.005\ mmHg/K$, do not conform to the power function proposed by Brutsaert (1975). This equation generates β values close to zero for low vapor pressures ($e_a/T_{air} < 0.005\ mmHg/K$), whereas the observed $\beta_{Effective}$ values indicate effective emissivity levels exceeding 0.6. This discrepancy extensively affects the model performance (Table 8.3 and Table 8.4) for both stations, making them less efficient when compared to the simplest model (Model 1) or the straightforward linear relationship between β and e_a in Model 4. Prata (1996) effectively tackled this issue in Model 5 and introduced Equation (8.10) (implemented in Model 7) to enhance the model's accuracy by considering the e_a/T_{air} term. The results for Model 7, under all sky conditions for both SLS and BVESM, clearly demonstrate this enhancement across all the model performance metrics, making Model 7 one of the top alternatives for both SLS and BVESM stations.

Regarding canopy impact modeling for DLI, obtaining precise data on air temperature within the canopy and branch/trunk temperatures is often challenging. For this reason, it is typically assumed that air and canopy temperatures are equivalent. In this study, we employed the relationship between the Leaf Area Index (LAI) and sky view (V_f) due to the unavailability of direct temperature measurements.

The ERA5 dataset provides LAI data, but it comes with unique characteristics. ERA5's LAI values are derived from monthly climatology studies, lacking inter-annual variations (https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation). These values are computed for specific grid cells area and not site-specific. For SLS, where the station is notably distant from any canopies (over 100 meters; see Figure III-A1b in Appendix III-A), ERA5's LAI values tend to overestimate the actual conditions. In contrast, BVESM, situated in an area surrounded by diverse canopies (see Figure III-A1d in Appendix III-A), aligns better with the general LAI values from ERA5.

The LAI values in ERA5 reflect a broader grid area (31 km × 31 km), making them more suitable for regional studies rather than specific site analysis. The scatter plots depicting the relationship between $\beta_{Effective}$ and LAI under clear-sky and all sky conditions are shown in Figure 8.11 for SLS and Figure 8.12 for BVESM.

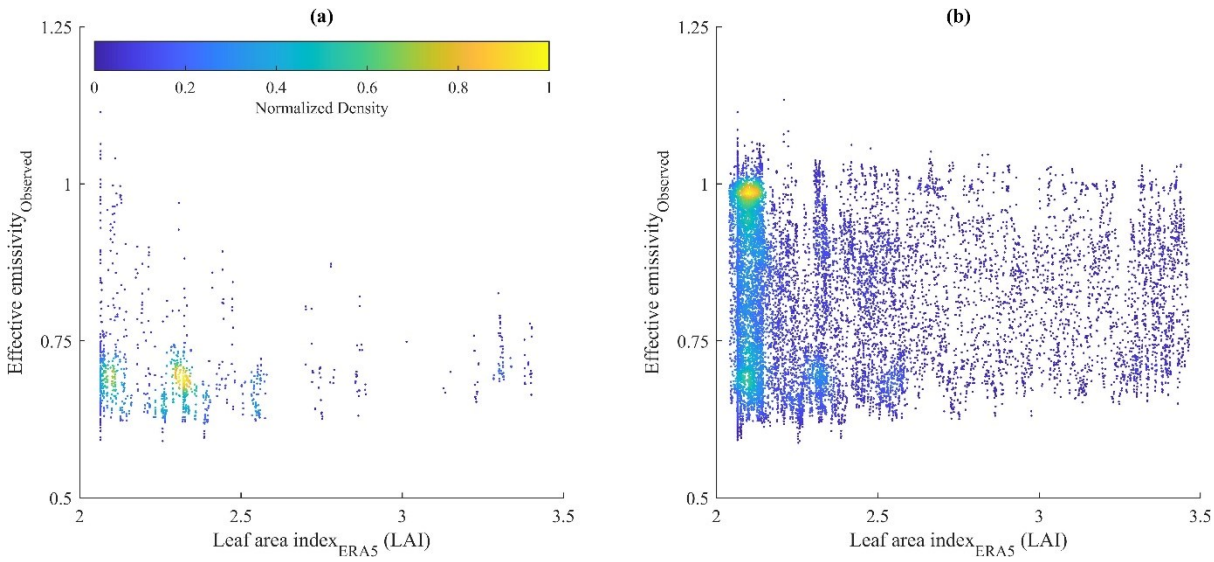

Figure 8.11 **Scatter plots between the measured effective emissivity ($\beta_{Effective}$) and Leaf Area Index (LAI) from the ERA5 dataset for (a) clear-sky condition and (b) all sky conditions for the SLS station**

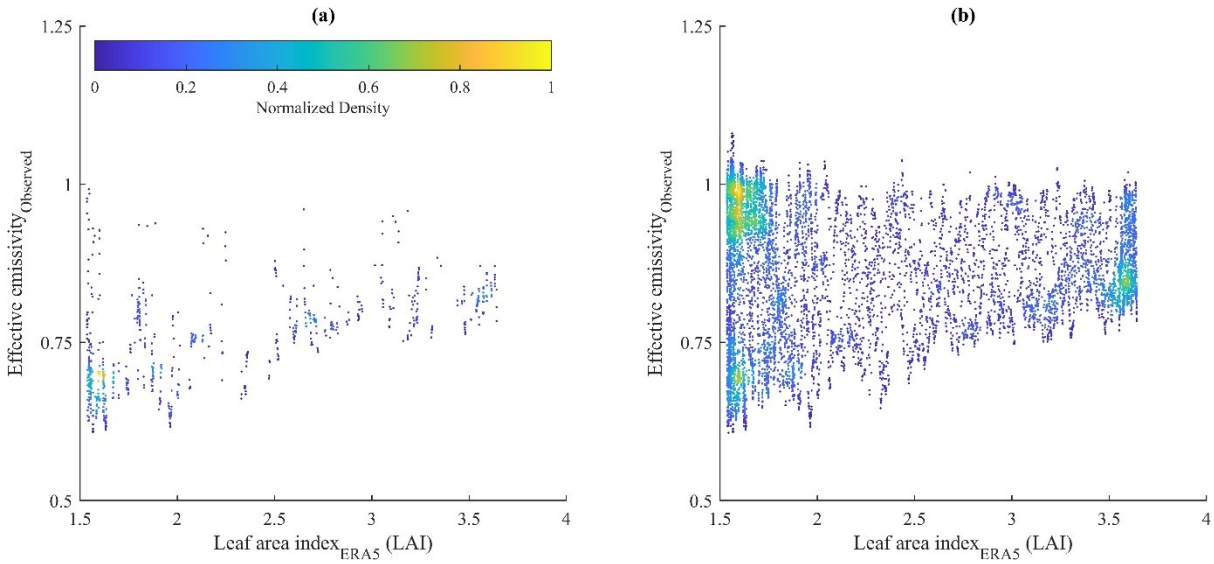

Figure 8.12 **Scatter plots between the measured effective emissivity ($\beta_{Effective}$) and Leaf Area Index (LAI) from the ERA5 dataset for (a) clear-sky condition and (b) all sky conditions for the BVESM station**

In Figure 8.11a and Figure 8.12a, we can observe seasonal variations in LAI, with low LAI values corresponding to winter and high values to summer. Most observations fall in between, representing spring and fall seasons. There is a positive correlation between LAI and $\beta_{Effective}$, particularly under clear-sky conditions.

However, in the all sky condition (Figure 8.11b and Figure 8.12b), the influence of LAI on $\beta_{Effective}$ is overshadowed by the combined effects of e_a and B. For Model 9 ($\alpha = 0.24$),

calibrated for SLS and BVESM, the LAI's impact on $\beta_{Effective}$ is high, however cloud cover impact is dominant. There is a more visible positive correlation between LAI and $\beta_{Effective}$ for BVESM, likely due to vapor pressure's interaction with the canopy cover. These intricate interactions make it challenging to determine the individual contributions of e_a, B, and LAI to $\beta_{Effective}$ using ERA5 data. Further measurements, including canopy temperature, cloud type, sky view, and shortwave extinction function for the canopy, would be required to unravel their specific effects.

Model 9, an extension of Model 4, introduces the LAI-based correction factor to distinguish open area DLI from canopy DLI. After calibration, the additional parameter in this model, denoted as α, is determined to be 0.24, signifying the influence of the canopy at both stations. This outcome aligns with expectations, considering the substantial impact of canopy cover on DLI. As demonstrated in Figure 8.11 and Figure 8.12, incorporating LAI as a variable affecting DLI enhances model accuracy, leveraging the existing ERA5 LAI data. While the results for SLS remained similar to those of Model 4, including LAI notably improved DLI estimates for the BVESM station.

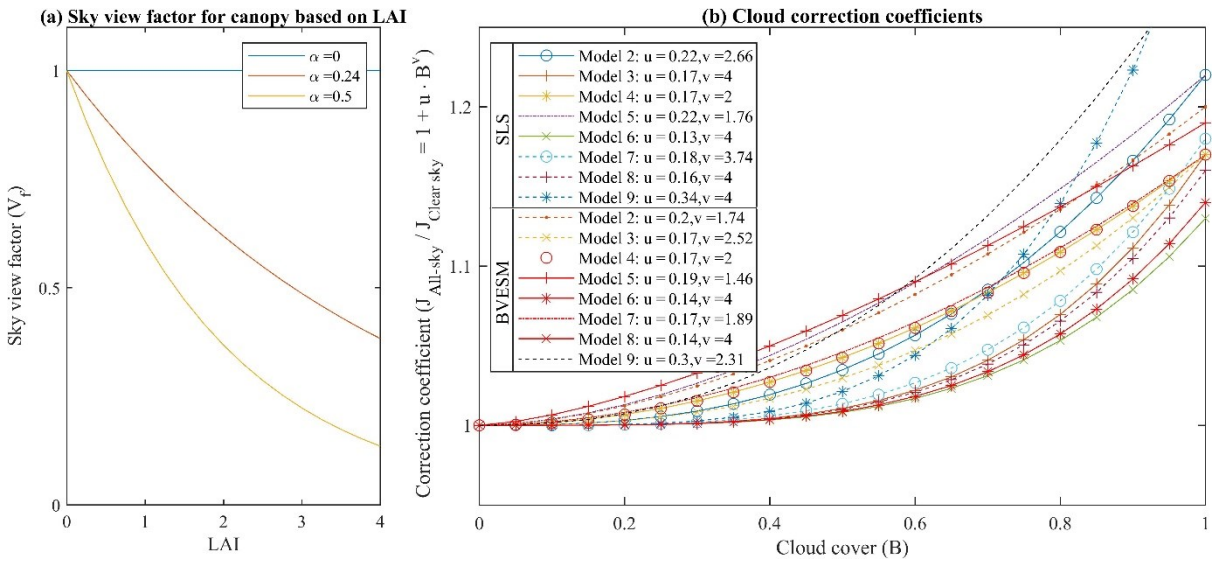

Figure 8.13 **(a) Relationship between the sky view variable (V_f) and the Leaf Area Index (LAI) using Equation (8.17) for three α values. (b) Computed cloud correction factors determined from Equation (8.13) and calibrated u and v values in this study for both SLS (first 8 rows numbered from Model 2 to 9) and BVESM (second 8 rows numbered from Model 2 to 9) stations (Table III-B1 in Appendix III-B).**

In summary, the comprehensive assessment of performance metric values in Table 8.3 and Table 8.4, along with the calibration results presented in Figure 8.2 and Figure 8.3, indicate that the straightforward utilization of the DLI data from the ERA5 dataset outperforms all the

DLI models for both the SLS and BVESM stations, emphasizing the ERA5 dataset's clear superiority in this context.

This superiority has two important messages: (1) ERA5-based DLI data are possibly reliable for all sky conditions (clear and cloudy-skies), and (2) they are reliable data for both bare land and canopy-dominated stations by precision comparable to the uncertainty of using DLI models and best instruments, at least for the conditions tested in this study.

Out of the several DLI models, three options (models 4, 6, and 9) show the most accurate estimations, which are our recommended options for future studies at the Nechako River watershed. This result agrees with Sugita and Brutsaert (1993), who mentioned that the Stephan-Boltzmann formulation is accurate enough. The only problem is finding the correct and precise formulation for β for all sky conditions. Although using models 4, 6, or 9 with two or three calibrated parameters improved RRMSE compared to Model 1 (from RRMSE = 13.9% to 11.8-11.9% for calibration at SLS and from 11.7% to 10-10.3% at BVESM), it is at the price of adding parameters; which means air temperature is the primary factor in estimating DLI and air emissivity, followed by cloud and finally by vapor pressure (controlled by added parameters in different models) (Long *et al.*, 2021).

The dominance of cloud cover in estimating β compared to vapor pressure for the Nechako agrees with findings for daytime DLI by Abramowitz *et al.* (2012). The enhanced performance seen in the BVESM station when incorporating LAI data into the models emphasizes the importance of considering local canopy characteristics in DLI estimations for areas with varying vegetation cover.

8.5 Conclusion

This study's main objective was to compare nine models for DLI using ERA5 reanalysis data for the Nechako watershed in British Columbia and BVESM in Quebec, Canada. More precise models for DLI for this watershed can help the assessment of the effect of heat budget for river temperature, a key variable for aquatic life. Calibration was station-based for each DLI model at SLS and BVESM, followed by validating those models on each station. Our results showed that the Stephan-Boltzmann equation could provide reasonable estimates for DLI using the proper formulation to estimate β. Three formulations (models 4, 6, and 9), our proposed options for hydrological model developments at the Nechako watershed, showed the best performance for both SLS and BVESM. Their superiority was shown using several performance metrics. Besides, the DLI downloaded data from the ERA5 reanalysis database provides the best results in terms of all performance metrics considered in this study, which suggests the direct use of ERA5 data instead of its calculation using a model. Our proposed

options can be incorporated into hydrological-thermal models like CEQUEAU (Khorsandi *et al.*, 2022), which is of interest to decision-makers in the Nechako watershed in British Columbia. Since meteorological data have a high local impact in this watershed and longwave radiation is one of the primary heat budget terms for water cooling during the summer, using more optimal options for DLI formulation in hydrothermal models can increase the model's prediction skill for water temperature modeling. This improved skill can also provide a better tool for climate change studies. Future studies are recommended to apply proposed models at the watershed scale using available physically-based hydrothermal models.

Acknowledgment

We thank Richard Loubier (Rio Tinto) and Michel Baraër (École de technologie supérieure) for providing longwave radiation data. This work was funded by the Canadian Natural Sciences and Engineering Research Council (NSERC) and Rio Tinto as part of a Collaborative Research and Development grant (Grant Number: CRDPJ 523640-18).

ARTICLE III BIBLIOGRAPHIE

Abramowitz G, Pouyanné L, Ajami H (2012). On the information content of surface meteorology for downward atmospheric long-wave radiation synthesis. *Geophysical Research Letters* 39(4).

Akaike H (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6):716-723.

Andrews DG (2010). *An introduction to atmospheric physics.* Cambridge University Press, Cambridge. 249 p

Arsenault R, Poulin A, Côté P, Brissette F (2014). Comparison of stochastic optimization algorithms in hydrological model calibration. *Journal of Hydrologic Engineering* 19(7):1374-1384.

Bárbaro EW, Oliveira AP, Soares J, Ferreira MJ, Božnar MZ, Mlakar P, Escobedo JF (2009). Downward Atmospheric longwave radiation in the city of São Paulo. *AIP Conference Proceedings.* AIP, p 447-450.

Bernhardt ES, Savoy P, Vlah MJ, Appling AP, Koenig LE, Hall RO, Arroita M, Blaszczak JR, Carter AM, Cohen M, Harvey JW, Heffernan JB, Helton AM, Hosen JD, Kirk L, McDowell WH, Stanley EH, Yackulic CB, Grimm NB (2022). Light and flow regimes regulate the metabolism of rivers. *Proceedings of the National Academy of Sciences* 119(8):e2121976119.

Bolz H (1949). Die Abhängigkeit der infraroten Gegenstrahlung von der Bewölkung. *Z. Meteorol* 3(Pt. 7).

Brunt D (1932). Notes on radiation in the atmosphere. I. *Quarterly Journal of the Royal Meteorological Society* 58(247):389-420.

Brutsaert W (1975). On a derivable formula for long-wave radiation from clear skies. *Water Resources Research* 11(5):742-744.

Budyko MI (1974). *Climate and life.* Academic Press, Inc., New York, USA. 508 p

Buizza R (2023). *Weather prediction: What everyone needs to know.* Oxford University Press,

Choi M, Jacobs JM, Kustas WP (2008). Assessment of clear and cloudy sky parameterizations for daily downwelling longwave radiation over different land surfaces in Florida, USA. *Geophysical Research Letters* 35(20).

Crawford TM, Duchon CE (1999). An improved parameterization for estimating effective atmospheric emissivity for use in calculating daytime downwelling longwave radiation. *Journal of Applied Meteorology and Climatology* 38(4):474-480.

de Kok RJ, Steiner JF, Litt M, Wagnon P, Koch I, Azam MF, Immerzeel WW (2020). Measurements, models and drivers of incoming longwave radiation in the Himalaya. *International Journal of Climatology* 40(2):942-956.

Dilley A, O'brien D (1998). Estimating downward clear sky long-wave irradiance at the surface from screen temperature and precipitable water. *Quarterly Journal of the Royal Meteorological Society* 124(549):1391-1401.

Duarte HF, Dias NL, Maggiotto SR (2006). Assessing daytime downward longwave radiation estimates for clear and cloudy skies in Southern Brazil. *Agricultural and Forest Meteorology* 139(3):171-181.

Free M, Sun B, Yoo HL (2016). Comparison between Total Cloud Cover in Four Reanalysis Products and Cloud Measured by Visual Observations at U.S. Weather Stations. *Journal of Climate* 29(6):2015-2021.

Gatien P, Arsenault R, Martel J-L, St-Hilaire A (2022). Using the ERA5 and ERA5-Land reanalysis datasets for river water temperature modelling in a data-scarce region. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 10.1080/07011784.2022.2113917:1-18.

Gauch M, Kratzert F, Gilon O, Gupta H, Mai J, Nearing G, Tolson B, Hochreiter S, Klotz D (2023). In Defense of Metrics: Metrics Sufficiently Encode Typical Human Preferences Regarding Hydrological Model Performance. *Water Resources Research* 59(6):e2022WR033918.

Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of hydrology* 377(1-2):80-91.

Hansen N, Ostermeier A (1996). Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. *Proceedings of IEEE international conference on evolutionary computation.* (Nagoya, Japan, IEEE, p 312-317.

Herrero J, Polo M (2012). Parameterization of atmospheric longwave emissivity in a mountainous site for all sky conditions. *Hydrology and Earth System Sciences* 16(9):3139.

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J-N (2020a). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146(730):1999-2049.

Hogan RJ, Bozzo A (2016). *ECRAD: A new radiation scheme for the IFS.* European Centre for Medium-Range Weather Forecasts, (787). 33 p

Huang Y, Ramaswamy V, Soden B (2007). An investigation of the sensitivity of the clear-sky outgoing longwave radiation to atmospheric temperature and water vapor. *Journal of Geophysical Research: Atmospheres* 112(D5).

Idso S, Jackson R (1969). Thermal radiation from the atmosphere. *Journal of Geophysical Research* 74(23):5397-5403.

Khorsandi M, St-Hilaire A, Arsenault R (2022). Multisite calibration of a semi-distributed hydrologic and thermal model in a large Canadian watershed. *Hydrological Sciences Journal* 67(14):2147-2174.

Kneizys F, Shettle E, Abreu L, ANDERSON G, CHETWYND J, GALLERY W, SELBY J, CLOUGH S (1988). Users guide to LOWTRAN 7, Optical/Infrared Technology Division. *Technical Report AFGL-TR-88-0177,* US Air Force Geophysics Laboratory, Hanscom Air Force Base Massachusetts.

Koll DD, Cronin TW (2018). Earth's outgoing longwave radiation linear due to H2O greenhouse effect. *Proceedings of the National Academy of Sciences* 115(41):10293-10298.

Krause P, Boyle D, Bäse F (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in geosciences* 5:89-97.

Ku HH (1966). Notes on the use of propagation of error formulas. *Journal of Research of the National Bureau of Standards* 70(4).

Legates DR, McCabe Jr GJ (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research* 35(1):233-241.

Li M, Jiang Y, Coimbra CFM (2017). On the determination of atmospheric longwave irradiance under all-sky conditions. *Solar Energy* 144:40-48.

Long Z, Ding L, Zhou J, Zhou T (2021). Toward the Estimation of All-Weather Daytime Downward Longwave Radiation over the Tibetan Plateau. *Atmosphere* 12(12):1692.

Maykut GA, Church PE (1973). Radiation Climate of Barrow Alaska, 1962–66. *Journal of Applied Meteorology and Climatology* 12(4):620-628.

McFarlane V, Clark SP (2021). A detailed energy budget analysis of river supercooling and the importance of accurately quantifying net radiation to predict ice formation. *Hydrological Processes* 35(3):e14056.

McMillan HK, Westerberg IK, Krueger T (2018). Hydrological data uncertainty and its implications. *WIREs Water* 5(6):e1319.

Miller DH (1981). *Energy at the Surface of the Earth: An Introduction to the Energetics of Ecosystems.* Elsevier, New York, USA

Morcrette J-J (2002). The surface downward longwave radiation in the ECMWF forecast system. *Journal of Climate* 15(14):1875-1892.

Morin G, Couillard D (1990). Predicting river temperatures with a hydrological model. *Encyclopedia of fluid mechanic, surface and groundwater flow phenomena,* Volk Gulf Publishing Company, Houston, Texas Vol 10. p 171-209.

Morin G, Nzakimuena TJ, Sochanski W (1994). Prevision des temperatures de l'eau en rivieres a l'aide d'un modele conceptuel: le cas de la riviere Moisie. *Canadian Journal of Civil Engineering* 21(1):63-75.

Murray FW (1967). On the computation of saturation vapor pressure. *Journal of Applied Meteorology and Climatology* 6(1):203-204.

Musselman KN, Pomeroy JW (2017). Estimation of Needleleaf Canopy and Trunk Temperatures and Longwave Contribution to Melting Snow. *Journal of Hydrometeorology* 18(2):555-572.

Niemelä S, Räisänen P, Savijärvi H (2001). Comparison of surface radiative flux parameterizations: Part I: Longwave radiation. *Atmospheric research* 58(1):1-18.

Oke TR (1987). *Boundary layer climates.* Routledge,

Ouellet V, Secretan Y, St-Hilaire A, Morin J (2014). Water temperature modelling in a controlled environment: comparative study of heat budget equations. *Hydrological Processes* 28(2):279-292.

Pomeroy JW, Marks D, Link T, Ellis C, Hardy J, Rowlands A, Granger R (2009). The impact of coniferous forest temperature on incoming longwave radiation to melting snow. *Hydrological Processes* 23(17):2513-2525.

Prata A (1996). A new long-wave formula for estimating downward clear-sky radiation at the surface. *Quarterly Journal of the Royal Meteorological Society* 122(533):1127-1151.

Satterlund DR (1979). An improved equation for estimating long-wave radiation from the atmosphere. *Water Resources Research* 15(6):1649-1650.

Schafer SA (2017). *What is the global impact of 3D cloud-radiation interactions?* Doctor of Philosophy (University of Reading, UK). 148 p. https://centaur.reading.ac.uk/72752/1/20800931_Schafer_thesis.pdf

Schwarz G (1978). Estimating the dimension of a model. *The Annals of Statistics* 6(2):461-464.

Sicart JE, Hock R, Ribstein P, Chazarin JP (2010). Sky longwave radiation on tropical Andean glaciers: parameterization and sensitivity to atmospheric variables. *Journal of glaciology* 56(199):854-860.

Sugita M, Brutsaert W (1993). Cloud effect in the estimation of instantaneous downward longwave radiation. *Water Resources Research* 29(3):599-605.

Swinbank WC (1963). Long-wave radiation from clear skies. *Quarterly Journal of the Royal Meteorological Society* 89(381):339-348.

Tang W, Qin J, Yang K, Zhu F, Zhou X (2021). Does ERA5 outperform satellite products in estimating atmospheric downward longwave radiation at the surface? *Atmospheric research* 252:105453.

Tarek M, Brissette FP, Arsenault R (2020b). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.* 24(5):2527-2544.

Upreti H, Ojha CSP (2017). Estimation of Relative Humidity and Dew Point Temperature Using Limited Meteorological Data. *Journal of irrigation and drainage engineering* 143(9):05017005.

Verseghy DL, McFarlane NA, Lazare M (1993). Class—A Canadian land surface scheme for GCMS, II. Vegetation model and coupled runs. *International Journal of Climatology* 13(4):347-370.

Voortman B, Bartholomeus R, Van der Zee S, Bierkens MF, Witte JPM (2015). Quantifying energy and water fluxes in dry dune ecosystems of the Netherlands. *Hydrology and Earth System Sciences* 19(9):3787-3805.

Yao B, Liu C, Yin Y, Liu Z, Shi C, Iwabuchi H, Weng F (2019). Assessment of cloud properties from the reanalysis with satellite observations over East Asia. *Atmos. Meas. Tech. Discuss.* https://doi.org/10.5194/amt-2019-223:1-36.

Yao B, Liu C, Yin Y, Liu Z, Shi C, Iwabuchi H, Weng F (2020). Evaluation of cloud properties from reanalyses over East Asia with a radiance-based approach. *Atmos. Meas. Tech.* 13(3):1033-1049.

Zhou YP, Cess RD (2000). Validation of longwave atmospheric radiation models using Atmospheric Radiation Measurement data. *Journal of Geophysical Research: Atmospheres* 105(D24):29703-29716.

Zhou Z, Lin A, Wang L, Qin W, Zhong Y, He L (2020). Trends in downward surface shortwave radiation from multi-source data over China during 1984–2015. *International Journal of Climatology* 40(7):3467-3485.

9 ARTICLE IV

Modeling the impact of downward longwave irradiance on water temperature modeling at the watershed scale using ERA5 input data in western Canada

Modélisation de l'impact de l'irradiance descendante à ondes longues sur la modélisation de la température de l'eau à l'échelle du bassin versant à l'aide des données d'entrée ERA5 dans l'ouest du Canada

Auteurs :

Mostafa Khorsandi[1], André St-Hilaire[1], Richard Arsenault[2], Suraj Patel[1]

[1]Canadian Rivers Institute and Institut National de la Recherche Scientifique, Centre Eau Terre Environnement (INRS-ÉTÉ), 490 De la Couronne St., Québec City, Quebec, G1K 9A9, Canada

[2]Hydrology, climate and climate change laboratory, École de technologie supérieure (ÉTS), Montréal, Quebec, Canada

L'article sera soumis dans la revue *Hydrological Processes*

Date de soumission: Janvier 2024

Contribution des auteurs :

Mostafa Khorsandi: Curation des données, méthodologie, conceptualisation, logiciel, programmation, analyse formelle, enquête, validation, visualisation, rédaction - ébauche originale, rédaction - révision et édition.

André St-Hilaire: Curation des données, acquisition de financement, administration de projet, ressources, supervision, validation, rédaction – révision & édition.

Richard Arsenault: Curation des données, acquisition de financement, ressources, supervision, validation, rédaction – révision & édition.

Suraj Patel : Programmation, visualisation, rédaction – révision & édition.

Résumé

Le rayonnement net à ondes longues (LR) joue un rôle crucial dans le refroidissement de la température de l'eau des rivières pendant les étés chauds, mais sa quantification à l'échelle des rivières reste limitée. Cette étude quantifie l'irradiance descendante à ondes longues (DLI) à l'aide de neuf modèles sélectionnés. Les données d'entrée du modèle hydrologique-thermique CEQUEAU comprennent les précipitations, la température de l'air, la température du point de rosée, le rayonnement net à ondes courtes, la vitesse du vent, la couverture nuageuse et l'indice de surface foliaire (LAI) issues du produit de réanalyse ERA5. Le bassin versant de la rivière Nechako dans le centre de la Colombie-Britannique, Canada, sert d'étude de cas. Les résultats indiquent que le modèle qui prend explicitement en compte les effets de la canopée grâce à l'indice de surface foliaire (LAI), offre les meilleurs résultats optimaux. Ces résultats fiables sont utilisés pour la modélisation de la température de l'eau par le biais d'un calage multisite à l'échelle du bassin versant. Diverses formulations de l'équation de Stephan-Boltzmann se révèlent efficaces pour quantifier l'impact de la DLI sur la température de l'eau à cette échelle. De plus, les données DLI de la grille ERA5 démontrent des performances presque optimales. Les performances améliorées du meilleur modèle mettent en évidence le potentiel de partitionner la DLI en zones ouvertes et avec couverture de canopée, en utilisant des formulations mathématiques et des données ERA5 en tant qu'entrées pour les modèles hydrologiques et écologiques. La signification de ces résultats est discutée, complétée par une analyse approfondie de l'incertitude, contribuant à notre compréhension de la modélisation de la température de l'eau et des études hydrologiques avec le modèle CEQUEAU.

Mots-clés

Irradiance longue onde descendante, Données ERA5, Température de l'eau, Modèle CEQUEAU, Échelle du bassin versant.

Abstract

Net Longwave radiation (LR) plays a crucial role in cooling river water temperature during hot summers, yet its quantification at the watershed scale remains limited. This study quantifies downward longwave irradiance (DLI) using nine selected models. Precipitation, air temperature, dew point temperature, net shortwave radiation, wind speed, cloud cover, and leaf area index (LAI) from the ERA5 reanalysis product are utilized as inputs for the CEQUEAU hydrological-thermal model. The Nechako watershed in central British Columbia, Canada, serves as the case study. Results indicate that the only tested model which explicitly considers canopy effects through LAI, yields the best outcomes. These reliable results are employed for water temperature modeling after multisite calibration at the watershed scale. Various formulations of the Stephan-Boltzmann equation prove effective for quantifying DLI's impact on water temperature at this scale. Furthermore, ERA5 gridded data for DLI demonstrate near-optimal performance. The enhanced performance of the best model highlights the potential for partitioning DLI into open areas and canopies, utilizing mathematical formulations and ERA5 data as input for hydrological and ecological models. The significance of these findings is discussed, advancing our understanding of water temperature modeling and hydrological studies with the CEQUEAU model.

Key Words

Downward Longwave Irradiance, ERA5 Data, Water Temperature, CEQUEAU Model, Watershed Scale.

9.1 Introduction

Longwave radiation (LR) plays a crucial role in the thermal dynamics of water bodies, impacting both cooling and warming processes. Enhancing the mathematical representation of LR holds the potential to refine hydrological estimations related to heat flux partitioning in water temperature modeling. LR is emitted by all objects due to their temperatures being above absolute zero. The Stefan-Boltzmann law allows for the precise micro-scale modeling of LR heat fluxes. Within this physical framework, two key variables emerge as the primary controlling variables: the object's absolute temperature (T) and its emissivity (β). Notably, emissivity values vary between air and water. Research has demonstrated that the emissivity coefficient for water (β_{water}) approaches unity (Morin & Couillard, 1990), implying that water closely resembles a black body in the context of LR radiation. In contrast, air's emissivity (β_{air}) is lower than water's, particularly in clear-sky conditions, rendering air closer to a grey body (Nussbaumer & Pinker, 2012).

The emissivity of air with regard to LR is closely linked to air temperature, water vapor pressure, and the extent of cloud cover within the atmospheric column (Niemelä *et al.*, 2001). Irrespective of the specific study environment, the net longwave radiation exerts influence on the hydrological system through two distinct mechanisms: (a) by contributing to the heat budget when $T_{air} - \sqrt[4]{\frac{\beta_{water}}{\beta_{air}}} T_{water} > 0$, or (b) by functioning as a cooling mechanism when this difference is less than zero (T_{air} and T_{water} are air and water temperatures in Kelvin; see Appendix IV-A for further details). Consequently, the estimation of β_{air} as a dynamic variable reflecting atmospheric conditions becomes crucial for an improved DLI heat budget assessment.

Estimation models for DLI are commonly categorized based on their utilization of specific β_{air} formulations (Choi *et al.*, 2008; Long *et al.*, 2021). Furthermore, DLI estimation models can be classified based on the nature of their input data. These models encompass a spectrum from physically based to regionally empirical forms (Bárbaro *et al.*, 2009; Herrero & Polo, 2012; Ouellet *et al.*, 2014; Prata, 1996; Sardar *et al.*, 2017). On the one hand, the physically-based models are all derived from the Stefan-Boltzmann equation, complemented by parameters accounting for cloud cover and atmospheric humidity corrections. Notably, research indicates that employing the Stefan-Boltzmann equation is adequate for achieving highly accurate DLI estimations (Choi *et al.*, 2008). However, the primary challenge lies in precisely estimating β_{air} at the site or regional scales.

Numerous studies have conducted site-scale comparisons of DLI models, aiming to identify the optimal formulation for β_{air} tailored to specific locations (de Kok *et al.*, 2020; Du *et al.*, 2023; Li *et al.*, 2017; Long *et al.*, 2021; Tang *et al.*, 2021). Additionally, the assessment of DLI models for hydrological-thermal modeling at a regional scale is made feasible by juxtaposing estimated DLI fluxes with field measurements (Ouellet *et al.*, 2014). However, due to challenges stemming from economic constraints, limited access to data, and technical barriers, only a handful of studies have focused on validating DLI models on a larger scale using field measurements (Brown *et al.*, 2021; Tang *et al.*, 2021). Consequently, an alternative approach indirectly validates DLI models as sub-modules integrated within hydrological and thermal models (Ouellet-Proulx et al., 2019). Despite the existence of studies that have compared various algorithms for other components of the heat budget, such as evaporation and evapotranspiration (Ouellet-Proulx *et al.*, 2019), the assessment of DLI estimation's impact at the watershed scale within the context of hydrological and thermal modeling remains unexplored. To the best of our knowledge, the work by Ouellet *et al.* (2014) stands as a key precedent, albeit conducted within a single controlled experimental environment.

Numerous climate products, including the fifth-generation ECMWF atmospheric reanalysis product (ERA5), offer valuable data on near-surface heat fluxes (Hersbach *et al.*, 2020a). An innovative alternative to traditional DLI modeling via the Stefan-Boltzmann equation involves directly incorporating these climate products into water temperature models, much like other meteorological input variables. Within Section 8 of this thesis (ARTICLE III), the superior performance of the ERA5 in modeling DLI flux is demonstrated, with a comparison against measured data and nine widely recognized modeling formulations based on the Stefan-Boltzmann law. Furthermore, Tang et al. (2021) presented an important finding, i.e., the ERA5 product for DLI surpasses satellite retrievals for land areas, marking a notable advancement in DLI estimation. While several climate variables have been extensively examined at the station or watershed scales using the ERA5 database, including precipitation, temperature (Taghizadeh *et al.*, 2021; Tarek *et al.*, 2020b), cloud cover (Yao *et al.*, 2020), water temperature (Gatien *et al.*, 2022; Khorsandi *et al.*, 2022), and irradiance variables (Jiang *et al.*, 2020), the direct utilization of ERA5's DLI product for water temperature modeling is a novel and promising approach.

The majority of hydrological models rely on a single equation for their DLI calculations (Dugdale *et al.*, 2017a; Ficklin *et al.*, 2012). Among these models, those purposefully tailored for the study of water temperature prove to be more suitable candidates for evaluating the effects of model structure on water temperature simulations, as demonstrated in Table IV-B1 (see Appendix IV-B) (DeChant & Moradkhani, 2014).

In light of the critical role of longwave radiation in water temperature modeling and its dependence on various factors, this study aims to investigate the impact of different DLI estimation formulations on water temperature simulations. Additionally, it explores the potential advantages of using ERA5's DLI data as a substitute for the internal formulations within the CEQUEAU model. To address these objectives the following research questions are considered:

(1) To what extent do different DLI estimation formulations impact water temperature simulations when incorporated into a semi-distributed model?
(2) Can the use of ERA5 DLI data, as an alternative to built-in formulations within the CEQUEAU model, yield potential benefits in water temperature modeling?

9.2 Methods

9.2.1 Study area

The primary focus of this research is the Nechako River sub-watershed, extending from the Skins Lake Spillway (SLS) to a point 16 km upstream of its confluence with the Stuart River.

Situated in British Columbia, Canada (Figure 9.1), this region has garnered considerable attention in prior studies due to its critical role in both private and public sectors (James *et al.*, 1991; Larabi *et al.*, 2022; Parkes, 2022; Picketts *et al.*, 2017; Sheedy, 2005). Notably, Rio Tinto, a key stakeholder, operates the SLS to meet industrial requirements while adhering to stringent environmental guidelines.

The Nechako River watershed experiences a wide annual temperature range, from -44°C to 35°C, reflecting its climatic diversity. The elevation ranges from 628 to 2240 meters above sea level (masl). The mean annual precipitation levels vary from 460 mm/year to 2177 mm/year, depending on the location within the watershed (mountainous or interior) and based on meteorological data from 1980 to 2019.

Within this watershed, the SLS plays a vital role in regulating the flow of the Nechako River between the spillway and its confluence with the Nautley River. Subsequently, the combined flow of the Nautley and Nechako Rivers traverses the town of Vanderhoof before converging with the Fraser River.

Hydrometric data are collected at three stations made available through collaboration with Rio Tinto and Environment and Climate Change Canada (data accessed in March 2022). These stations, progressing from upstream to downstream, include (1) the SLS station, (2) the Nautley River station, located just upstream of its confluence with the Nechako River, and (3) the Vanderhoof station (see Figure 9.1). Notably, the SLS and Nautley stations serve as boundary conditions, furnishing essential flow and temperature data for our model's hydrological and thermal modules.

Seven stations within the Nechako watershed measure water temperature hourly, numbered from 1 to 7 in a downstream-to-upstream sequence (see Figure 9.1). This study uses data from stations 1 to 4 due to their small influence from large upstream lakes (Khorsandi *et al.*, 2022). Additionally, the substantial distance between the upstream boundaries and the measurement point allows for the manifestation of local meteorological variables' influence on water temperature (Gatien *et al.*, 2022).

The drainage area from the SLS to Station 1 spans approximately 5,433 km^2. The flow at the SLS is managed by Rio Tinto, serving both industrial needs and environmental considerations, including flood mitigation. Notably, Rio Tinto adheres to stringent environmental regulations during the Sockeye salmon (*Oncorhynchus nerka*) migration season, from mid-July to mid-August. To maintain mean daily water temperatures below 20°C at the Finmoore station, located roughly 30 km downstream from the Vanderhoof station, Rio Tinto operates the Summer Temperature Management Program (STMP) water release protocol. This protocol

was developed in collaboration with Fisheries and Oceans Canada between 1980 and 1983 and remains in effect (Macdonald, 2019).

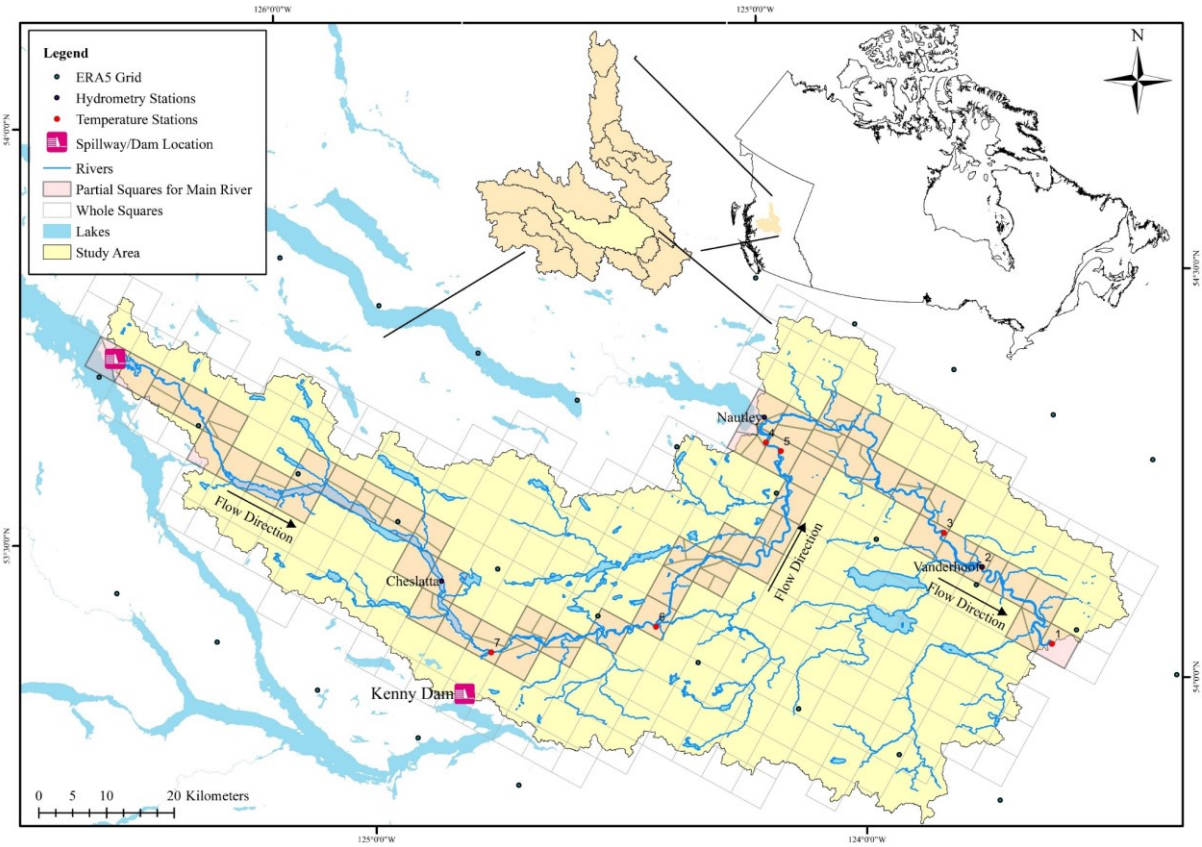

Figure 9.1 **Nechako River watershed study site, presenting important locations together with labeled monitoring stations. The model and ERA5 grids are also shown. Among all 615 partial squares (CPs), only those related to the main river are colored.**

9.2.2 ERA5 reanalysis dataset

ERA5, the fifth generation of the European reanalysis dataset, is made available by the European Center for Medium-Range Weather Forecasting (ECMWF). It offers gridded meteorological data for the atmosphere at a spatial resolution of 31 km. In this study, we used eight key daily climate variables from the ERA5 database. These variables include precipitation, air temperature (°C), dew point temperature (°C), wind speed (km/h; converted from m/s, the unit used in ERA5 dataset), total cloud cover (ranging from 0 to 1), net solar shortwave radiation (W/m^2), downward longwave irradiance (DLI, W/m^2), and leaf area index (LAI, m^2/m^2).

Initially, ERA5 data were downloaded in its original gridded format and recorded at hourly intervals. Subsequently, daily means were computed based on the hourly data, with daily

precipitation summed. The heat flux variables are converted from W/m^2 to MJ/m^2/day by summation of sub-daily values. Finally, the daily gridded data were interpolated onto the CEQUEAU model grid, as depicted in Figure 9.1. We employed the built-in CEQUEAU interpolator to facilitate this interpolation, employing a nearest neighbors approach (Ouellet-Proulx *et al.*, 2019). This process ensures the ERA5 data are seamlessly integrated into our modeling framework for comprehensive analysis and simulations.

9.2.3 CEQUEAU model

CEQUEAU, originally designed as a modular, semi-distributed hydrological model (as depicted in Figure IV-C1 in Appendix IV-C), was initially developed to model water flow and quality (Morin & Couillard, 1990). It has evolved and currently can simulate/forecast flow and water temperature. This versatile model supports sub-daily time steps for water flow simulations and daily time steps for water temperature predictions. Numerous prior studies have successfully applied the CEQUEAU model for various hydrological purposes, including flow simulation (Fniguire *et al.*, 2022), flow forecasting (Mai *et al.*, 2020), water temperature simulation (Ouellet-Proulx *et al.*, 2019), and water temperature forecasting (Ouellet-Proulx *et al.*, 2017a; Ouellet-Proulx *et al.*, 2017b).

One of CEQUEAU's strengths lies in its general-purpose design, allowing it to perform effectively under diverse hydrological conditions (Mai *et al.*, 2020). Notably, our choice of the CEQUEAU model for this comparative analysis is further justified by its recent reprogramming, which enables the selection of several algorithms for key heat budget components (St-Hilaire et al., 2015). This adaptability and its track record of successful applications make it a fitting candidate for our study.

The CEQUEAU model employs a simple discretization scheme involving a two-step spatialization process for watershed representation. In the first step, the watershed is subdivided into equal square-shaped areas, which serve as gridded units for hydrometeorological variables. These square units are referred to as "Whole Squares," denoted as "CE" following the original French notation (Morin & Couillard, 1990). Each CE requires elevation and land cover data for modeling. Critical CE-based variables include the proportion of land covered by three distinct land cover classes: forests, lakes and wetlands, and bare soil.

In the second step, for conceptual hydrological modeling, smaller hydrological response units are established, known as "Partial Squares" or "CP" (based on the original French notation). CPs are created by overlapping each CE with sub-basin boundaries, with each CE potentially containing up to four CPs (Morin & Couillard, 1990). Each CP operates as a hydrological

response unit, utilizing CEQUEAU's production function. This production function is grounded in a conceptual vertical water routing scheme, primarily relying on three interconnected reservoirs.

The first reservoir represents storage in lakes and wetlands, the second corresponds to the first layer of soil, and the third signifies the saturated soil layer (Khorsandi *et al.*, 2022). Subsequently, water is routed downstream for each CP at every time step, in proportion to the CP's area.

Within the CEQUEAU model structure, the hydrological module is responsible for computing the water balance, which serves as an essential input for the thermal module. The heat budget calculations within the thermal module are performed using Equation (9.1) (St-Hilaire *et al.*, 2015). This equation provides the foundation for quantifying the various heat terms and their contributions to the overall thermal dynamics within the studied watershed:

$$\Delta H = H_{short} + H_{long} + H_{ev} + H_{sens} + H_{ups} + H_{local} + H_{ground} \tag{9.1}$$

where ΔH is the overall change of heat budget for a CP (MJ); H_{short} is net shortwave radiation (MJ); H_{long} is net heat gain or loss by longwave radiation (MJ); H_{ev} is the latent heat loss (MJ) (evaporation); H_{sens} is the net sensible heat (MJ); H_{ups} is the heat gain from upstream (MJ); H_{local} is the heat coming from local runoff and interflow (MJ); and H_{ground} is the heat term related to groundwater (MJ). The equations to calculate each of the above-mentioned heat terms are as follows:

$$H_{short} = CRAYSO.A.R_{short} \tag{9.2}$$

$$H_{long} = CRAYIN.A.\sigma(\beta_{air}T_{air}^4 - \beta_{water}T_{water}^4) \tag{9.3}$$

$$H_{ev} = -CEVAPO.L_{ev}.A.E \tag{9.4}$$

$$H_{sens} = CCONVE.A[0.2U(T_{air} - T_{water})] \tag{9.5}$$

$$H_{ups} = V_{ups}.\theta.T_{water,ups} \tag{9.6}$$

$$H_{local} = V.\theta.T_{air} \tag{9.7}$$

$$H_{ground} = V_{ground}.\theta.TNAP \tag{9.8}$$

where A is the water surface area in the CP (m^2); $CRAYSO$, $CRAYIN$, $CEVAPO$, and $CCONVE$ are empirical dimensionless coefficients for thermal module calibration (presented in Table IV-D6 in Appendix IV-D); R_{short} is net solar radiation (MJ/m^2); σ is Stefan-Boltzmann constant (5.6704×10^{-8} W/m^2/K^4 or 4.8992×10^{-9} MJ/m^2/K^4/day); β_{air} and β_{water} are the sky and water emissivities (dimensionless, $\beta_{water} = 0.97$ in CEQUEAU); T_{air} and T_{water} are air and water temperatures, respectively (K) for the current CP; L_{ev} is the depth of evaporated water from river reach, which is calculated by the hydrological module using the Thornthwaite equation (m); E is the latent heat of evaporation (2480 MJ/m^3, considering constant water density), and U is wind speed (km/h); $T_{water,ups}$ is upstream CP's water temperature, $TNAP$ is CP's groundwater temperature; V , and V_{ups} are the estimated water volume by the hydrological module for the current, and upstream CPs, respectively (m^3), and θ is the heat capacity of water (4.186 MJ/kg/°C) (Morin & Paquet, 2007). After calculating all the heat terms, the thermal module calculates water temperature change using the following equation:

$$\Delta T = \frac{\Delta H}{V.\theta} \tag{9.9}$$

where ΔT is the change in the water temperature (°C).

This study's current model structure for the Nechako watershed has 271 CEs and 615 CPs. The model's modular architecture provides easy addition and modification of routines for several heat budget terms (St-Hilaire *et al.*, 2015) (Figure IV-C1 in Appendix IV-C). The current structure of CEQUEAU has 34 parameters (Tables IV-D1 to D3 in Appendix IV-D). They include 24 parameters for the hydrological module calibration, two parameters for evaporation calculation using the Thornthwaite equation which are hydrological-thermal calibration parameters, and eight parameters for the thermal module calibration.

In our model, the term "parameter" indicates that its value remains constant throughout the simulation, while the term "global" implies that these parameters maintain a uniform value across all CPs. As a recent addition to the meteorological input data, ERA5's DLI data and LAI data are introduced. These newly incorporated variables undergo interpolation and integration into the model structure, following the same procedure as previous meteorological inputs. ERA5's DLI data serves as an alternative to traditional DLI heat flux models. Of all the heat terms under consideration, this study places a particular emphasis on $H_{long\downarrow}$ [based on Equation (9.1) and Equation IV-A-1 in Appendix IV-A]. The literature provides several modeling options for this term, detailed in the following section.

9.2.3.1 Longwave radiation formulation

The original formulation for the CEQUEAU model for net longwave radiation is Equation (9.3), in which the air emissivity is:

$$\beta_{air} = [0.97(1 + 0.17B^2)(0.74 + 0.0065e_a)] \tag{9.10}$$

where β_{air} is the air emissivity (dimensionless); e_a= water vapor pressure ($mmHg$) and B is total cloud cover (0-1, dimensionless). Since vapor pressure is not one of the ERA5 variables for near-surface, it is calculated using the Tetens equation with dew point temperature at 2 m height as follows (Ouellet-Proulx, 2018):

$$e_a = e_s(T_{dew}) = 7{,}50062 \times \begin{cases} 0{,}61078 \times \exp\left(\frac{17{,}27T_{dew}}{T_{dew} + 237{,}3}\right) & T_{dew} \geq 0\ °C \\ 0{,}61078 \times \exp\left(\frac{21{,}875T_{dew}}{T_{dew} + 265{,}5}\right) & T_{dew} < 0\ °C \end{cases} \tag{9.11}$$

where e_s= saturated vapor pressure (mmHg), T_{dew}= Dew point temperature ($°C$), and 7.50067 is the conversion factor (from kPa to $mmHg$). Based on this formulation, DLI (coming from the near-ground atmosphere or $H_{long\downarrow}$) is:

$$H_{long\downarrow} = A.[0.97(1 + 0.17B^2)(0.74 + 0.0065e_a)]\sigma T_{air}^4 \tag{9.12}$$

which is weighted by $CRAYIN$ during model calibration.

ARTICLE III compared nine different formulations for DLI (more specifically, different formulations for β_{air}) and recommended these alternative formulations for further CEQUEAU development, as shown in Table 9.1.

Table 9.1 Formulation of the nine DLI estimation models for this study

Model No.	Unit for J	Unit for e_a	The final form of the equation ($J_{All-sky} = (1 + uB^v) \times J_{Clear-sky}$)	Reference for $J_{Clear-sky}$	Reference for $(1 + uB^v)$
1	W/m^2	–	$J = a\sigma T_{air}^4$	Stephen-Boltzmann law for grey body	(Duarte et al., 2006)
2*	mW/cm^2	–	$J = (1 + uB^v) \times 5.31 \times 10^{-14} T_{air}^6$	(Swinbank, 1963; Equation 2)	(Duarte et al., 2006)
3	W/m^2	–	$J = (1 + uB^v)\langle 1 - 0{,}261 \times \exp[-7{,}77 \times 10^{-4}(273 - T_{air})^2]\rangle \sigma T_{air}^4$	(Idso et al., 1969)	(Duarte *et al.*, 2006)
4	W/m^2	$mmHg$	$J = [(1 + 0{,}17B^2)0{,}97 \times (0{,}74 + 0{,}0065e_a)]\sigma T_{air}^4$	(Morin et al., 1990; Equations 11 and 12)	(Morin et al., 1990; Equation 12)
5	W/m^2	hPa	$J = (1 + uB^v)[1{,}24(\frac{e_a}{T_{air}})^{1/7}]\sigma T_{air}^4$	(Brutsaert, 1975; Equation 11)	(Duarte et al., 2006)
6	W/m^2	hPa	$J = (1 + uB^v)\langle 1{,}08[1 - exp(-e_a^{\frac{T_{air}}{2016}})]\rangle \sigma T_{air}^4$	(Satterlund, 1979; Equation 5)	(Duarte et al., 2006)
7	W/m^2	hPa	$J = (1 + uB^v)(1 - \left[1 + 46{,}5\left(\frac{e_a}{T_{air}}\right)\right] exp\langle -\left\{1{,}2 + 3\left[46{,}5\left(\frac{e_a}{T_{air}}\right)\right]\right\}^{1/2}\rangle)\sigma {T_{air}}^4$	(Prata, 1996; Equation 20)	(Duarte et al., 2006)
8	W/m^2	hPa	$J = (1 + uB^v)\begin{cases}[0{,}72 + 0{,}009(e_a - 2)]\sigma {T_{air}}^4 & e_a \geq 2 \\ [0{,}72 - 0{,}076(e_a - 2)]\sigma {T_{air}}^4 & e_a < 2\end{cases}$	(Niemelä et al., 2001; Equation 18)	(Duarte et al., 2006)
9**	W/m^2	$mm\ Hg$	$J = \{(1 + uB^v)[exp(-\alpha \times LAI)]0{,}97 \times (0{,}74 + 0{,}0065e_a) + 0{,}98 \times \langle 1 - exp(-\alpha \times LAI)\rangle\}\sigma {T_{air}}^4$	(Pomeroy et al., 2009; Equation 3; Verseghy et al., 1993; Equation 7)	(Duarte et al., 2006)

* By conversion from mW/cm^2 to W/m^2 the constant is 5.31×10^{-13}. Also, following Stephen-Boltzmann law notation using σT_{air}^4 the term can be written as $(9.3645 \times 10^{-6} T_{air}^2)\sigma T_{air}^4$.

** The canopy transmissivity $exp(-\alpha \times LAI)$ is based on Beer's law of radiation (Verseghy *et al.*, 1993). Pomeroy *et al.* (2009) cited Oke (1987) for the $\beta_{canopy} = 0.98$.

where u and v are dimensionless calibration parameters. The e_a unit is mmHg for Model 4 and Model 9 while the e_a unit is hPa for other models. The other alternative recommended by Khorsandi *et al.* (2023; ARTICLE III) is using DLI gridded values from the ERA5 database named "mean surface downward longwave radiation flux" or "msdwlwrf" (W/m^2) (Hersbach *et al.*, 2020a).

9.2.4 Model calibration

In this section, the calibration of various DLI alternative models is detailed. First, the CEQUEAU model is calibrated hydrologically using the water flow observations at Vanderhoof. The resulting parameters are presented in Tables IV-D1 in Appendix IV-D. This

process is iterated eight times to ensure the convergence of hydrological calibration, and the mean of the iterated results is used for thermal calibrations to ensure robustness and reliability.

The thermal calibrations 1-9 involve using alternative models for DLI, as outlined in Table 9.1. Model 4 represents CEQUEAU's original DLI formulation. Additionally, the tenth proposed model directly leverages DLI data from the ERA5 database, integrating it into the CEQUEAU model without utilizing a built-in DLI formulation.

The calibration and simulation period is from 2016 to 2019, during which the first three months of 2016 were considered the spin-up period (for hydrological calibrations, the spin-up period was two years, covering 2014 to 2015). Recommended by Arsenault *et al.* (2014) for hydrological studies, the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) optimization algorithm (Hansen, 2016) was implemented for parameter estimation using a maximum of 5000 model evaluations as the stopping criterion for all calibration scenarios.

The process involves multisite calibration, and following this calibration, performance metrics are computed for each site. In this context, multisite calibration harnesses observed data from all measuring sites to optimize the model as a single-objective optimization problem as follows:

Minimize: $$Objective\ Function = RMSE_{multisite}$$

$$RMSE_{multisite} = \sqrt{\frac{\sum_{j=1}^{M}\sum_{t=1}^{N_j}\left(S_{t,j} - O_{t,j}\right)^2}{\sum_{j=1}^{M} N_j}} \tag{9.13}$$

Subject to: $$Lower\ boundary \leq X \leq Upper\ boundary$$

where $RMSE_{multisite}$, is the multisite Root Mean Squared Error (RMSE) performance metric for all M stations; $O_{t,j}$ and $S_{t,j}$ respectively are observed and simulated values for the jth station at t^{th} time step; N_j is the number of observed time steps for j^{th} station (measurements), X is calibration decision variable as a vector, and the two $Lower$ and $Upper$ boundaries are the optimization constraints for the decision variable as two vectors (Table IV-D3 in Appendix IV-D). After finishing the calibration process, the single-site $RMSE_j$ performance metric is re-calculated at each site individually for validation using the multisite optimum parameter set.

9.3 Results

9.3.1 Model calibration

Figure 9.2 and Table 9.2 show the RMSE performance metric for each site and the $RMSE_{multisite}$ for all stations for each calibration scenario. Figure 9.2 and Table 9.2 show that

Model 1 has the poorest metrics than the other models, with higher RMSE values for all sites using the single-site metrics and higher $RMSE_{multisite}$ for multisite calibration. Station 1, the downstream station, shows the lowest RMSE values compared to the other sites, probably due to lower variance due to a shorter measurement period (234 days). Out of all the installed sensors in the watershed, stations 1-4 were selected due to their locations far from the upstream lakes. Therefore, all the stations are not dominantly impacted by upstream large lakes, which is the weakness of the current version of CEQUEAU in modeling the reservoir effect on water temperature (Auffray *et al.*, 2023; Khorsandi *et al.*, 2022). Also, near-zero water temperatures should be ignored since CEQUEAU does not have a module for ice effects on water temperature (the freezing period is December to March, which is not considered for calibrations and simulation performance metrics).

On the other hand, station 2, with the most prolonged measurement period (1461 days), shows higher RMSE values than stations 1 and 3. Since station 2 is between stations 1 and 3 with higher RMSE compared to them, the RMSE difference is partially due to extra variance due to more extended measurements at station 2. Also, there is a large river (Sinkut River) between stations 1 and 2 and no tributaries between stations 2 and 3, which can be part of the reason for the variance change. Station 4 shows higher RMSE values compared to stations 1-3. Station 4 is the only station upstream of the Nechako-Nautley confluence. This station is probably impacted by upstream large water bodies and the tributaries joining the Nechako River. Also, it has a more extended measurement period compared to stations 1 and 3 (623 days).

Calculating the $RMSE_{multisite,}$ the aggregate RMSE value from simulated water temperatures at the watershed scale, helps find the best DLI candidate for all stations without focusing on a specific station. Using $RMSE_{multisite}$, Model 1 has the highest RMSE value (1.31°C; lowest performance), and Model 9 resulted in the lowest value (1.14°C; highest performance) (Table 9.2). The remaining DLI models (2, 3, 4, 5, 6, 7, and 8) have similar performances (Figure 9.2). It can be seen that the ranking of most and least optimal DLI models at the site scale and the watershed scale are the same; however, the rankings for other models are slightly different (Table 9.2 and Table 8.4 in ARTICLE III).

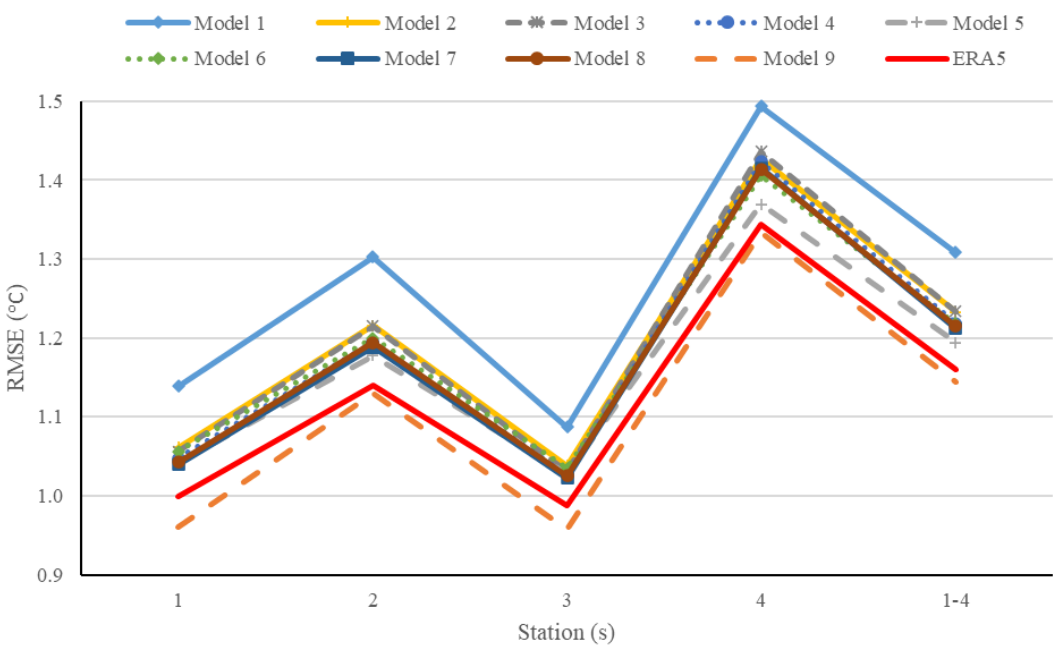

Figure 9.2 **Performance metric (RMSE) resulted from calibrations for (a) each site using derived multisite calibration (numbered 1, 2, 3, and 4 for according stations), and (b) $RMSE_{multisite}$ using all stations (1-4).**

Table 9.2 **Performance metrics by calibrating the thermal module parameters of CEQUEAU using temperature measurements via multisite calibration using stations 1-4. The best DLI formulation method with the best RMSE is bold. The model marked by a star is the CEQUEAU's default model.**

Scenario number	DLI model	Target station	Number of observation days	RMSE (°C) at each station using derived parameters from multisite calibration and $RMSE_{multisite}$ for all stations (°C) in each scenario (Bias values are calculated in parentheses with °C unit)				
				1	2	3	4	1-4
1	Model 1	St 1-4	2561	1.14 (0.23)	1.30 (-0.14)	1.09 (0.3)	1.49 (0.31)	1.31 (0.07)
2	Model 2	St 1-4	2561	1.06 (0.09)	1.22 (-0.08)	1.04 (0.18)	1.43 (0.43)	1.23 (0.10)
3	Model 3	St 1-4	2561	1.06 (0.12)	1.21 (-0.08)	1.03 (0.21)	1.44 (0.42)	1.23 (0.1)
4	Model 4*	St 1-4	2561	1.05 (0.15)	1.19 (-0.08)	1.03 (0.22)	1.42 (0.41)	1.22 (0.11)
5	Model 5	St 1-4	2561	1.04 (0.05)	1.18 (-0.05)	1.03 (0.12)	1.37 (0.43)	1.19 (0.1)
6	Model 6	St 1-4	2561	1.06 (0.08)	1.20 (-0.06)	1.03 (0.17)	1.41 (0.43)	1.22 (0.1)
7	Model 7	St 1-4	2561	1.04 (0.11)	1.19 (-0.05)	1.02 (0.18)	1.42 (0.44)	1.21 (0.12)
8	Model 8	St 1-4	2561	1.04 (0.08)	1.19 (-0.04)	1.03 (0.16)	1.41 (0.45)	1.21 (0.12)
9	**Model 9**	St 1-4	2561	**0.96** (0.09)	**1.13** (-0.11)	**0.96** (0.14)	**1.33** (0.46)	**1.14** (0.08)
10	ERA5	St 1-4	2561	1.00 (0.08)	1.14 (-0.05)	0.99 (0.18)	1.34 (0.42)	1.16 (0.11)

By comparing the $RMSE_{multisite}$ in Table 9.2 depicted in Figure 9.2, it is evident that using Model 9 and ERA5 data for DLI presents the best result for every station and the whole watershed. The calibrated thermal parameters for all ten calibrations are shown in Table 9.3.

Table 9.3 Thermal parameters values for different multisite calibration scenarios

	DLI model name									
P. Name	1	2	3	4	5	6	7	8	9	10 (ERA5)
CRAYSO	0.180	0.183	0.196	0.213	0.199	0.198	0.206	0.203	0.206	0.234
CRAYIN	0.255	0.279	0.298	0.366	0.295	0.344	0.321	0.330	0.675	0.368
CEVAPO	0.050	0.050	0.050	0.050	0.163	0.051	0.054	0.055	0.189	0.050
CCONVE	0.213	0.199	0.182	0.162	0.204	0.176	0.185	0.180	0.109	0.157
CRIGEL	974.208	0.748	0.754	0.748	31.874	0.810	2.175	5.395	29.618	31.868
TNAP	9.094	5.195	8.364	6.851	9.308	6.933	8.137	7.652	8.482	6.691
BASSOL	6.405	17.242	19.886	14.691	6.538	10.158	19.467	18.968	7.413	19.686
CORSOL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
XAA	0.000	0.000	0.000	0.000	0.000	0.021	0.009	0.009	0.008	0.000
XIT	38.291	12.462	8.080	39.065	13.540	36.076	6.294	38.821	36.452	11.140

The simulated water temperatures for four stations and ten DLI options are shown in Appendix IV-E1 to E10 in Appendix IV-E.

The DLI models have been validated in the previous study (ARTICLE III) using observed values at the SLS (In Nechako) station and Bassin Versant Expérimental de Sainte-Marthe (BVESM) station (70 km west of Montreal, QC) using hourly temporal resolution (ARTICLE III). The results showed high performance using DLI data from the ERA5 dataset for low canopy (SLS) and high canopy (BVESM) stations. The calibration and validations were based only on two sites; however, since the equations for DLI are validated for several case studies in the literature, they are considered as reliable alternatives for DLI in CEQUEAU, as shown in Figure 9.2 and Table 9.2. Model 9 provided the best performance metrics at the watershed scale, followed by using ERA5 DLI directly.

9.3.2 Multi-model ensemble simulations

In addition to evaluating the individual performance of each of the ten models, we present ensemble simulations that offer a better perspective on the range of potential outcomes (Figure 9.3). These ensemble multi-model simulations incorporate all ten models, providing lower and upper boundaries for the simulated water temperature. The ensembles allow us to explore the inherent uncertainty associated with hydrological-thermal modeling using DLI alternatives, as they encompass a spectrum of heat budget term variations and corresponding impacts on water temperature predictions.

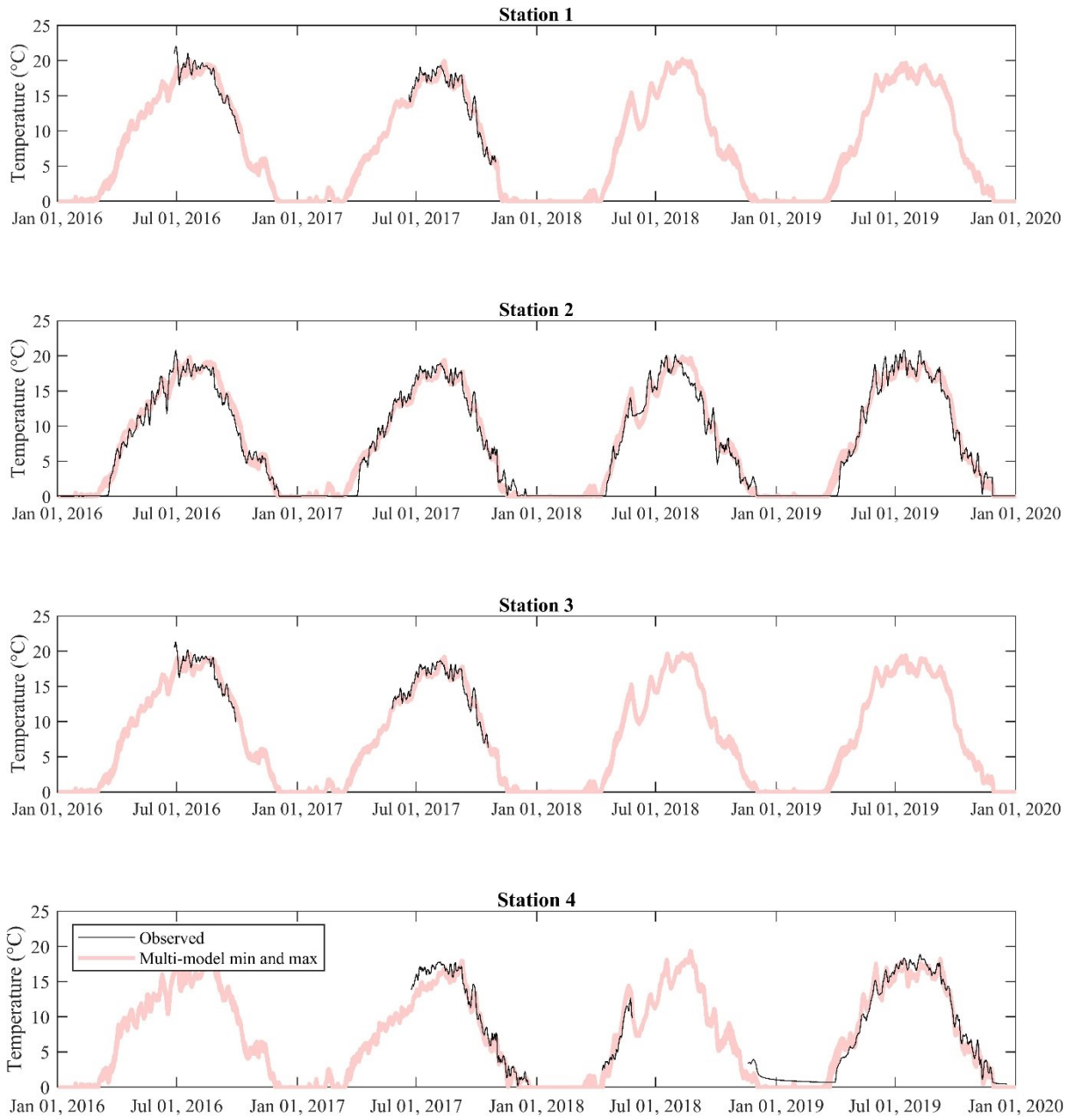

Figure 9.3 **Simulated water temperature time series for the four stations using calibrated CEQUEAU equipped with ten DLI modeling alternatives. The shaded area shows lower and upper estimates for each time step, and the black line is the observed value.**

Figure 9.3 shows the uncertainty range for the water temperature simulation by changing DLI models. The uncertainty range ($|Simulated_{max} - Simulated_{min}|$) is from 0 to 2°C, with lower uncertainty in the summers. The narrow range of uncertainty for April to November is not surprising since all the DLI models are close estimations for H_{long} by implementing models 1 to 10.

9.3.3 Uncertainty quantification

9.3.3.1 Observation uncertainty

This study uses three sets of observations: (1) upstream boundary conditions (at two stations for water flow and temperature), (2) observed water flows at one station used for hydrological calibration, and (3) water temperatures at four stations for thermal calibrations. The simulation time steps are daily, which means the diurnal variation of water flow and temperature is ignored for all the stations since the mean daily water temperature is used for operational management of the spillway for environmental flows. This diurnal variation range is available for water temperature based on hourly measured data for stations 1 to 4, Nautley, and SLS (Table 9.4).

Table 9.4 **Minimum and maximum diurnal variation ($|Observed_{max} - Observed_{min}|$) for temperature measurements at six stations.**

Station		1	2	3	4	5 (Nautley)	6 (SLS)
*Diurnal thermal variation (°C)	Min	0.1	0	0.4	0	0.2	0
	Max	2	5.1	3.7	3.6	7.3	6

*Rio Tinto measured the acquired data at hourly time steps

Not only are the lateral, longitudinal, and vertical temperature variations not measured in the Nechako River, but they are also neglected during simulations. The reason is that station measurements represent the entire CP for CEQUEAU, and the temperature in the prism volume of a CP is uniform. The current CEQUEAU model uses a two-step discretization scheme with 5 × 5 km grid cells for CEs. Although there are three-dimensional temperature variations within each CP, these are overlooked due to CEQUEAU's simplification as a hydrological model. There is no estimation for these variations at the Nechako watershed; however, some studies showed this variation could be greater than 2°C for a wide, slow flowing river reach and even more in case of cold/warm flows from tributaries (Hani *et al.*, 2023). Lastly, it is crucial to consider the instrument's precision for measuring water temperature. The temperature sensors installed in the Nechako River exhibit a precision error ranging from 0.2°C to 0.5°C (https://www.onsetcomp.com/products/data-loggers).

9.3.3.2 Input data uncertainty

The ERA5 data are not purely observed values but a reanalysis of observations with simulations by computational fluid dynamic models for atmosphere, ocean, and land interactions. Therefore, by using ERA5 data, part of the total $RMSE_{multisite}$ (residual uncertainty) is due to this modeling element related to the input data. Among all the meteorological variables used in this study for the 2016-2019 period, there are observed values for December 2019 to May 2022 (with hourly time steps) for four variables: (1) T_{air}, (2) downward solar radiation (DSR), (3) DLI, and (4) net solar radiation (NSR) for the SLS station. The comparison of daily mean values of these data provides an estimate of the data uncertainty for modeling in this study (Figure 9.4).

Comparing the ERA5 data with observations based on daily time steps shows relatively high accuracy for air temperature and DLI (Figure 9.4a,c), acceptable accuracy for DSR, and acceptable but biased values for NSR (Figure 9.4b,d). The NSR values are overestimated by the ERA5 dataset (Bias = NSR_{ERA5}-$NSR_{Observed}$ ≈ 2 MJ/m^2/day). Considering the RMSE as a metric for uncertainty quantifications, the air temperature uncertainty is 1.5°C, and three heat budget terms with uncertainty less than 10% (using RRMSE) show comparably high-quality input data for hydrological modeling (McMillan *et al.*, 2018).

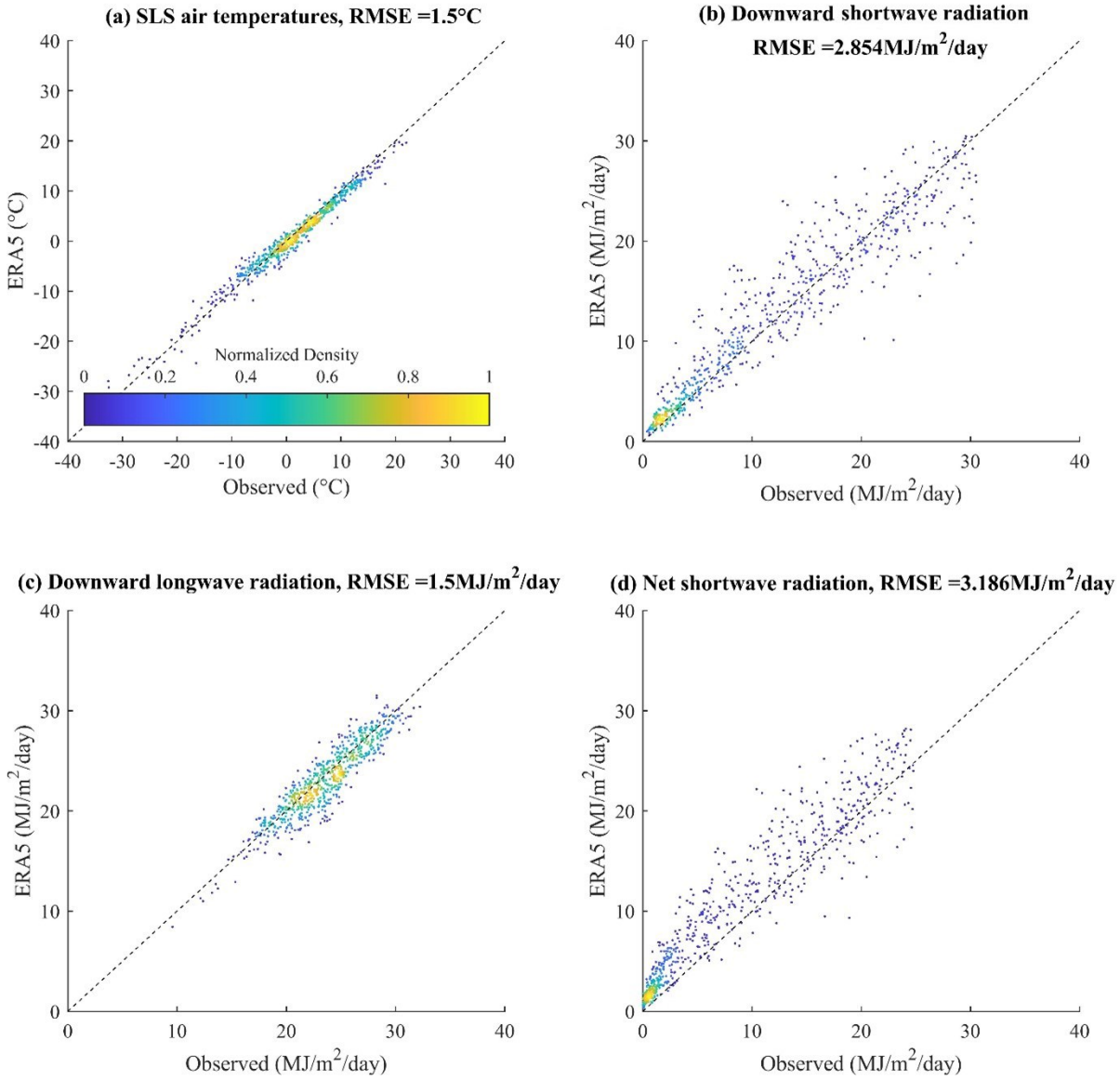

Figure 9.4 **Scatter plots for the comparison of ERA5's meteorological data at 2 m height with observed measurements at the SLS station on a daily time basis (the air temperature data are averaged, and radiations are summed using hourly data)**

Shortwave heat gain or loss is absent at night, and aggregating the energy budget on daily time steps gives a better comparison between the magnitude of DLI and DSR for daytime and night combined. Figure 9.4 shows that on a daily time step, the DLI magnitude equals the incoming and outgoing shortwave radiation (10-30 MJ/m^2/day for DLI compared to 0-30 MJ/m^2/day for DSR and NSR). The outgoing longwave radiation (and consequently net longwave radiation) is not compared since it is controlled by the land surface modeling in the model used by ECMWF, and the 31 km × 31 km spatial resolution of this global model does not provide reliable estimates for land surfaces with finer scale [e.g., Outgoing longwave radiation (OLR)]. Also, in the CEQUEAU model, which calculates the water heat budget, we

incorporated just DLI and not OLR since the OLR is controlled by simulated water temperature by the CEQUEAU itself and not the forced data.

9.3.3.3 Stream flow impact on water temperature simulation

Another source of uncertainty is the hydrologic calibration—the more uncertain the water quantity, the more uncertain the water temperature simulation. The measurements at the first two stations are boundary conditions for the model, and the model is calibrated just using one station (Vanderhoof). Therefore, residual uncertainty in hydrological calibration is inevitable, impacting thermal simulation uncertainty. The calibration results for Vanderhoof can provide an estimate for this flow uncertainty (Figure 9.5).

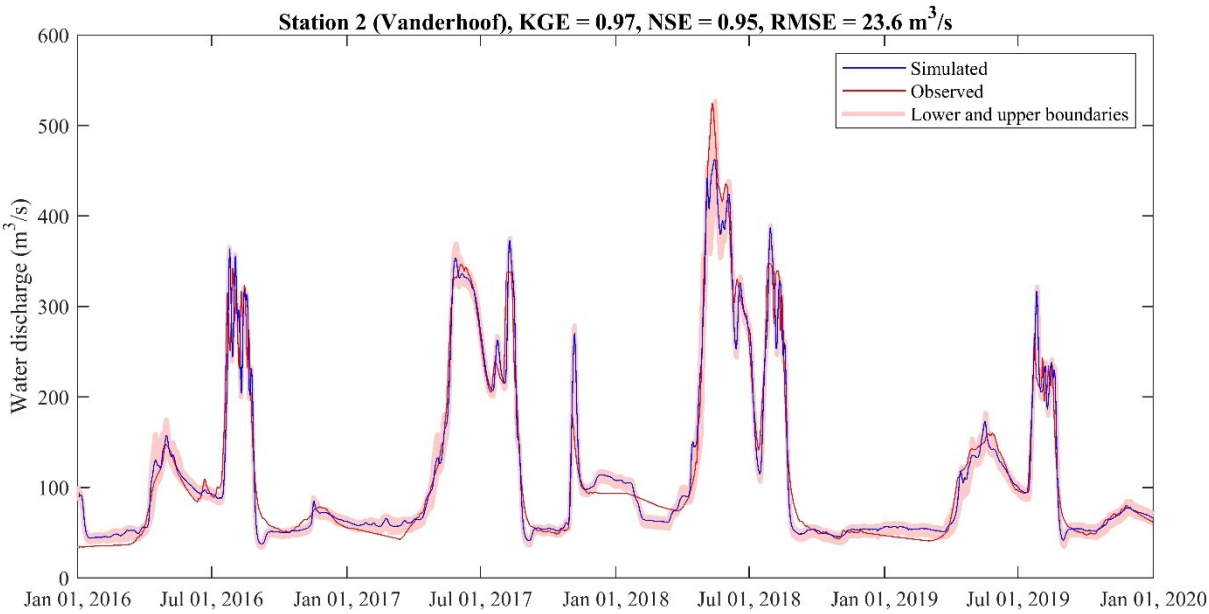

Figure 9.5 **Observed and simulated flow for the Vanderhoof station using calibrated hydrological parameters. The simulation curve represents the mean of eight sets of parameters resulting from iterative calibrations conducted for the entire basin. The shaded area denotes the upper and lower uncertainty bounds.**

If the water flow quantity for a CP is overestimated or underestimated, the value of V in Equation (9.9) will be biased. The overall error using the RMSE metric for hydrological calibration based on the Vanderhoof station (the only measuring site used for hydrological calibration) is 23.6 m^3/s (mean annual flow = 116.2 m^3/s, RRMSE ≈ 20%) (Figure 9.5). Considering no error in the thermal heat budget calculation for a CP, the ΔT is impacted by a $\pm 0.2 \times V$ in the flow (e.g., water volume) estimation. In this case, the relative error in the estimations for ΔT for one time step is:

$$\left|\frac{\frac{\Delta H}{V.\theta}-\frac{\Delta H}{(V+0.2V).\theta}}{\frac{\Delta H}{V.\theta}}\times 100\right| < |Relative\ Error\ for\ \Delta T| < \left|\frac{\frac{\Delta H}{V.\theta}-\frac{\Delta H}{(V-0.2V).\theta}}{\frac{\Delta H}{V.\theta}}\times 100\right| \quad (9.14)$$

$$16.6\% < |Relative\ Error\ for\ \Delta T| < 25\%$$

It is possible to estimate the consequent estimation error in ΔT using this relative error for each simulated time step. With the measured thermal range at Vanderhoof (0 to 20.8°C), the absolute error for ΔT is 5.2°C (the range for Nautley is 0 to 22°C and the absolute error in ΔT is 5.5°C).

The error in water temperature incremental change (ΔT) in a time step t then propagates in space to the downstream CPs and in time to the next time steps, and the estimation of this type of uncertainty cannot easily be formulated for a semi-distributed hydrological-thermal model.

By partitioning the H_{ups} from the total heat budget of the Vanderhoof station (for CP = 23) for calibration periods (April to November), we found that the contribution of H_{ups} is on average 96.6% (varies from 88.1% to 99.3% using the three summers in the 2016-2019 simulation) (Figure IV-F1 in Appendix IV-F). Therefore, a 20% error in estimating water flow (or water volume) means, on average, 19.2% (20% × 0.966) of the total heat budget in this CP for a time step. After H_{ups}, the main heat contribution is from the local terms, as shown in Table 9.5.

Table 9.5 **The mean contribution of each surface heat flux to the total local heat budget variations during the day using the simulated time series for the 2016-2019 period (The means are calculated using April- November simulated values). For this simulation, the ERA5 DLI (Model 10) is used**

Surface heat budget	Mean (%)	Min (%)	Max (%)
H_{short}	+50.7	+11.5	+72.1
H_{long}	-35.0	-73.5	-4.8
H_{ev}	-1.3	-5.2	0
H_{sens}	-6.4	-54.3	+75.5

Without discharge observations for other CPs, it is impossible to quantify this uncertainty for the whole watershed. However, performing the hydrological calibration once and using the same water quantities for all the scenarios in this paper, it is possible to assume this uncertainty is almost the same for all ten calibrations/simulations. We emphasize “almost the

same" since the evaporation is both a mass and an energy flux. Therefore, there is a parametric interdependency between water quality and quantity modules via water evaporation. However, since the amount of water evaporation depth is much less than the stream flow depth in Canadian rivers [less than 7 mm/day (Ouellet-Proulx *et al.*, 2019)], it is possible to assume there is negligible covariance between the water quantity calculated in the hydrological calibration and the evaporation heat term in the thermal calibration stage. Based on this assumption, the thermal calibration is almost independent of hydrologic calibration using CEQUEAU, and the residual uncertainty from hydrologic calibration is almost the same for all thermal calibration/simulation scenarios.

9.4 Discussion

Evaluating the CEQUEAU hydrological-thermal model's results entails grappling with various uncertainties (errors) from various sources affecting the longwave radiation heat budget. This section discusses our findings, with impacts on longwave radiation heat budget (H_{long}) in CEQUEAU.

9.4.1 Model parameter uncertainty

The CEQUEAU model contains a total of 34 parameters listed and categorized in Appendix IV-D. Of those, 24 parameters are exclusively for the hydrologic module, 8 are exclusively for the thermal module, and two are related to evaporation (Thornthwaite equation constants) and are shared between hydrologic and thermal modules (Tables IV-D1 to IV-D3 in Appendix IV-D). We used the CMA-ES algorithm to estimate these parameters using multisite calibration in each scenario. Regardless of the structure and formulation of the model, the estimation of model parameters is an ill-posed problem by nature (DeChant & Moradkhani, 2014).

Difficulties in parameter calibration are several and include those associated with different types of uncertainties: (1) model structural uncertainty, which results from simplifications and discretization in the model, and (2) observation and input data uncertainty. The combined effect of these factors makes the problem ill-posed because separating and quantifying these uncertainties is challenging, leading to less certain parameter estimates for the CEQUEAU model. In a study using a controlled environment to partition heat budget terms with high precision in Quebec, Canada, it was shown that when there is no heat advection, the primary heat source is incoming solar radiation together with DLI (shortwave radiation during the daytime and DLI during day and night), and the leading heat sinks were outgoing longwave radiation and evaporation (both during day and night time) (Ouellet *et al.*, 2014).

Our study can be considered as an extension of the work by Ouellet *et al.* (2014) using a hydrological-thermal model for a sub-watershed, including advection heat terms that were not present in the controlled experiment to study DLI impacts and the difference between DLI modeling alternatives. However, different possible sets of optimized parameters change the heat budget proportions, and without measurements in the Nechako river for each heat term, it is impossible to constrain them during calibration.

On the other hand, choosing the correct optimization algorithm for parameter estimation is challenging. There are several optimization algorithms with numerous applications in hydrology (Bozorg-Haddad *et al.*, 2017). This study used CMA-ES due to its strength in finding the global optimum and the convergence speed compared with other currently used optimization methods (Arsenault *et al.*, 2014). Therefore, this study cannot quantify the uncertainty due to this choice.

9.4.2 CEQUEAU model inadequacies affecting H_{long}

The CEQUEAU model structure, as a conceptual model, has several assumptions and simplifications; we try to discuss the inadequacies with impacts on H_{long} based on Equations (9.1) and (9.3) implemented in the CEQUEAU model structure.

The physical foundations for longwave radiation are well formulated using the Stephen-Boltzmann law in thermodynamics. The primary source of uncertainty for H_{long} is β_{air} estimation for all weather conditions to calculate the DLI; consequently, to better estimate the water temperature and finally to better calculate the OLR. The complex relationship between meteorological variables and β_{air}, and several heat budgets affecting the water temperature make it difficult to address this uncertainty with high precision. Although DLI and OLR are still subject to more research for single stations (Du *et al.*, 2023; Li *et al.*, 2017; Long *et al.*, 2021), this paper is one of the few studies trying to address DLI's impact on the watershed scale by implementing alternative DLI models into a semi-distributed hydrological model.

The ensemble of ten alternative DLI models can provide an estimate for H_{long} uncertainity (Dion *et al.*, 2021; Ouellet-Proulx *et al.*, 2017b). Figure 9.6 shows the mean simulated H_{long} using ten calibrated CEQUEAU models for ten DLI models with their minimum and maximum boundaries for CP = 23 (Vanderhoof station) for calibration periods. Figure 9.6 shows that the mean range of uncertainty (RMSE for minimum and maximum boundaries) equals 5.69E+6 MJ (RRMSE = 88%) (metrics are calculated using April to November periods).

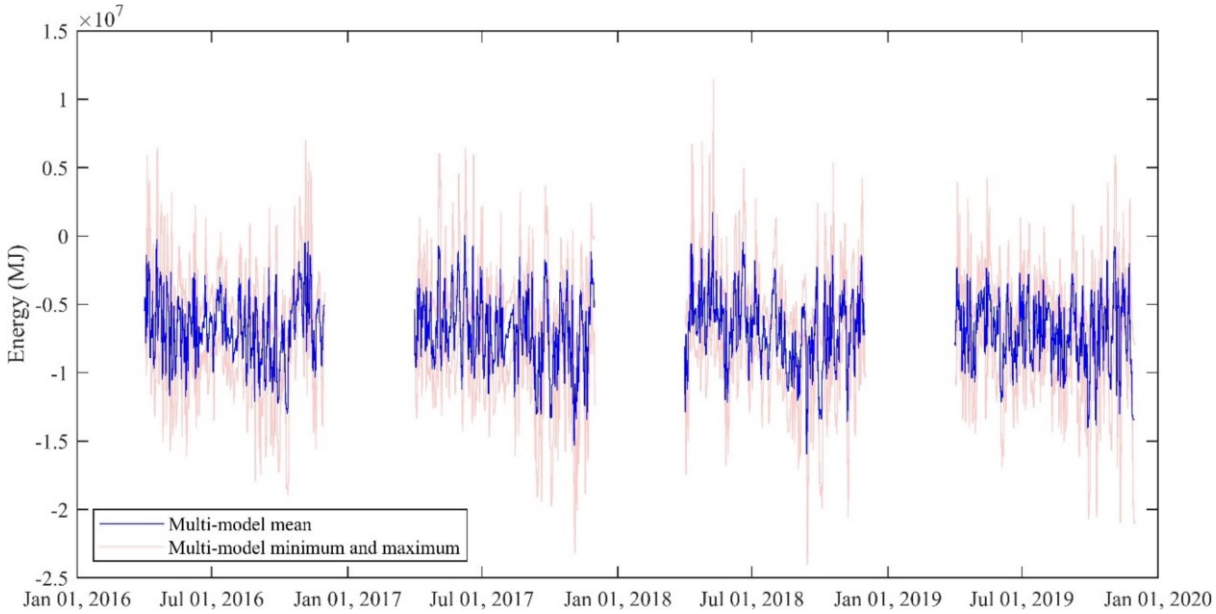

Figure 9.6 **Simulated H_{long} time series for the Vanderhoof station using calibrated CEQUEAU equipped with ten DLI model alternatives. The shaded area shows lower and upper bounds for each time step, and the blue line is the ensemble mean of simulated values.**

9.4.3 DLI model choice impact on model calibration parameters

The CEQUEAU model uses eight global parameters exclusively for the thermal module and two for latent heat shared between hydrological and thermal modules. The values of these parameters over ten multisite calibration scenarios are presented in Table 9.3. Among eight parameters, CRAYSO (≈ 0.2), CRAYIN (0.25-0.4, except 0.675 for Model 9), CCONVE (≈ 0.05, lower boundary for calibration, except for Models 5 and 9), CORSOL (≈ 0), and XAA (≈ 0, lower boundary for calibration) are relatively constant (less variable) in all scenarios.

The TNAP and BASSOL values are variable over scenarios and are within physically acceptable bounds. The BASSOL values range from 5-20 mm, suggesting that the days with this amount of precipitation or higher should be considered low solar radiation days. The CORSOL value is close to 0, meaning that the estimations for BASSOL are unbiased without the need for a correction factor. The correction function in the CEQUEAU source for the mean solar radiation (ASR) for rainless days is ASR x (1 + CORSOL) and for rainy days as ASR x (1 - CORSOL) ranging from 0 to 1.

The importance of negative values for H_{long} was corroborated by Khorsandi *et al.* (2022) (ARTICLE I), who showed that when using ERA5 meteorological data as input to the CEQUEAU model, the evaporation term is not the most critical heat loss budget term in the Nechako watershed, and the longwave heat term has a higher impact on cooling the water

temperature during summers. This higher importance is visible in absolute values compared to latent heat (Figure IV-F1 in Appendix IV-F).

9.4.4 The acceptable performance of models 2-8

The results obtained from our study demonstrate the better performance of models 2-8 and ERA5 data compared to Model 1 both at the station and watershed scales. This better performance can be attributed to incorporating an enhanced β_{air} estimation compared to Model 1 with a constant value for β_{air} and incorporating all climatic effective variables on β_{air} (Satterlund, 1979). By accurately estimating the β_{air} values, our models can better capture the underlying relationships between T_{air} and T_{water}, leading to more precise predictions for H_{long} and other heat budget terms. This enhanced estimation prevents flawed values in T_{water} with similar results compared to the current DLI model in the CEQUEAU model (Model 4 for DLI). Moreover, Model 9 consistently outperformed CEQUEAU's default option (Model 4) across single-site and multisite performance metrics. Considering the data-scarce watershed of Nechako, the current ten DLI model alternatives are only validated upstream of the watershed (only for the SLS station). The results in ARTICLE III showed the superiority of ERA5 at the station scale, followed by Model 9, while the results of this paper show the superiority of Model 9 at the watershed scale, followed by ERA5 DLI. However, their performances are close (Figure 9.2). The best performance among the nine DLI models (alternative mathematical formulations) at the watershed scale is the same as the site scale analysis (Model 9). We must mention that the results for ARTICLE III and this study should be validated using more observed data across the Nechako watershed, especially downstream of this watershed, since most measured water temperatures are collected in this region. It is essential to acknowledge that the generalizability of our models should be examined across diverse watersheds and measuring sites for DLI. While our experiments were conducted on the Nechako watershed, further investigations on several locations would provide a more comprehensive evaluation of the DLI model's performance, especially by measuring more heat budget terms.

9.4.5 ERA5 performance at the watershed scale

While the results for single sites (ARTICLE III) demonstrated the superiority of ERA5-based DLI data, our focus on multisite thermal performance metrics at the watershed scale again showed this high performance. Table 9.2 illustrates the commendable performance of ERA5's DLI data at individual stations. The excellence of ERA5's DLI data locally, particularly at individual stations before integrating into the hydrological model, can be attributed to its comprehensive representation of cloud impacts on DLI within the lowest atmospheric layer

(Hogan & Bozzo, 2016). This edge stems from the highly detailed modeling within the ECMWF global atmosphere model. The ECMWF Earth system model incorporates satellite data assimilation for cloud cover, assimilates observed land and ocean data, and excels in land-surface energy partitioning. As a result, this Earth system reanalysis output outperforms satellite-based DLI products for inland regions, showcasing the benefits of such extensive data assimilation. Furthermore, including a "Longwave layer-wise emission" scheme refines heat transfer among atmospheric layers, extensively improving DLI estimates over the watershed area (Hogan & Bozzo, 2016).

The three-dimensional DLI modeling approach employed by the ECMWF atmosphere model enriches the precision of DLI estimations over the watershed, primarily when focusing solely on meteorological conditions without involving land hydrological processes. At the local scale and without considering land hydrological processes, ERA5's DLI data outperform the mathematical models for DLI considered in our study, which predominantly rely on one-dimensional modeling approaches (as documented in ARTICLE III).

However, the intricacies grow when transitioning from local-scale station-based meteorological modeling to hydrological modeling encompassing heat budgets within water bodies and land-surface mechanisms impacting them (White *et al.*, 2023). It becomes evident that merely incorporating superior input data does not inherently guarantee more realistic estimations for individual surface heat budget components. These components are subject to interactions and multifaceted processes. While relying solely on high-quality input data may produce accurate results, it risks generating misinterpretations, primarily due to the lack of a comprehensive understanding of the underlying physical processes (White *et al.*, 2023).

9.4.6 Model 9 superiority

Model 9, an innovative derivative of Model 4, stands out as the best DLI alternative in our study, showcasing improvements in water temperature modeling for the Nechako watershed. The $RMSE_{multisite}$ for Model 4 for this region is 1.22°C, a respectable performance metric. However, introducing Model 9 elevates the performance, pushing the $RMSE_{multisite}$ down to 1.14°C. This reduction in RMSE enhances the model's predictive accuracy. Even more intriguing, it surpasses all other alternatives, including using the ERA5 DLI product.

Model 9 is the sole model in our study to incorporate the LAI as a distributed variable, forging a crucial link between canopy coverage and DLI. This innovative approach allows us to capture the interactions of canopy impacts with the radiant environment, reshaping our understanding of how DLI evolves in the watershed. The result is probably a more accurate

representation of DLI generation through the conversion of H_{short} to heat and its subsequent contribution to the longwave radiation component (H_{long}) of the heat budget.

An essential heat budget component is evaporation (H_{ev}), which shapes 10-15% of the total heat budget for controlled environments (Ouellet-Proulx *et al.*, 2019). Model 9 has better aligned these estimations with real-world scenarios, providing a more realistic representation of this negative heat budget term.

Although the ERA5 dataset provides more precise DLI data at the SLS station, it does not inherently alter the partitioning of the land heat budget, particularly changes in longwave radiation due to land cover. In contrast, Model 9 integrates a parameter of $\beta_{canopy} = 0.98$ for the canopy, thereby increasing its emissivity and relating it with LAI. Incorporating distributed LAI from the ERA5 dataset within Model 9 enabled it to accommodate other heat budget components flexibly, thereby elevating the level of realism in DLI modeling for hydrological-thermal studies.

Based on Pomeroy *et al.* (2009), who examined the impacts of coniferous forest temperatures on DLI and snowmelt, we incorporated Model 9 to partition DLI into two components. Initially, Pomeroy *et al.* (2009) partition heat inside the canopy into three critical components: direct incident DLI, DLI originating from the canopy, and shortwave extinction that converts to longwave radiation within the canopy. While we did not have access to the shortwave extinction function for the canopy effects in the Nechako watershed, we assumed T_{air} equals T_{canopy}, thus effectively modeling DLI from the canopy (as mentioned by Pomeroy *et al.* (2009)). This improvement marks a promising avenue for future research where we can refine our understanding by exploring shortwave extinction function modeling within the CEQUEAU model by further temperature measurements under canopy and developing CEQUEAU's source code to incorporate a shortwave extinction function.

9.5 Conclusion

The main objective of this study was to compare nine mathematical models for DLI plus the ERA5 DLI as a substitute for these models at the watershed scale. The Nechako watershed in British Columbia, Canada, and the CEQUEAU hydrological-thermal model was used. CEQUEAU's calibration for each DLI model at four water temperature measuring sites was performed using a multisite calibration method using all measuring data. Our results showed that the Stephan-Boltzmann equation and a suitable formulation for air emissivity could provide the best performance metrics for simulated water temperature estimates. Moreover, the results suggest using the DLI variable directly from the ERA5 reanalysis product, which provides high performance metrics. Our results show that using DLI from the ERA5 dataset

improves the model's performance at the whole watershed scale and provides data at a more extended period for long-term decision-making of interest to stakeholders in the Nechako watershed. This improved skill confirms the several previous studies recommending ERA5 data as a reliable source for data-scarce regions and partitioning DLI for the under canopy condition. Also, our results addressed sources of uncertainty due to variables impacting DLI in CEQUEAU. Due to improved results using ERA5 DLI data at the station level and the higher performance by Model 9, we recommend directly using this data or implementing the Model 9 formulation for thermal modeling using CEQUEAU and recommend future studies focusing more on other heat budget terms combined with these alternatives.

Acknowledgment

This work was funded by the Canadian Natural Sciences and Engineering Research Council (NSERC) and Rio Tinto as part of a Collaborative Research and Development grant (Grant Number: CRDPJ 523640-18). The authors would like to thank Richard Loubier, the programmer/developer of the CEQUEAU at Rio Tinto, for his valuable comments on the source code.

ARTICLE IV BIBLIOGRAPHIE

Arsenault R, Poulin A, Côté P, Brissette F (2014). Comparison of stochastic optimization algorithms in hydrological model calibration. *Journal of Hydrologic Engineering* 19(7):1374-1384.

Auffray M, Senécal J-F, Turgeon K, St-Hilaire A, Maheu A (2023). Reservoirs regulated by small dams have a similar warming effect than lakes on the summer thermal regime of streams. *Science of The Total Environment* 869:161445.

Bárbaro EW, Oliveira AP, Soares J, Ferreira MJ, Božnar MZ, Mlakar P, Escobedo JF (2009). Downward Atmospheric longwave radiation in the city of São Paulo. *AIP Conference Proceedings.* AIP, p 447-450.

Bozorg-Haddad O, Solgi M, Loï HA (2017). *Meta-heuristic and evolutionary algorithms for engineering optimization.* John Wiley & Sons,

Brown J, Rosolem R, Woods R, Rocha H, Roberti D (2021). Evaluation of high-resolution meteorological global data products using flux tower observations across Brazil.), p EGU21-15387.

Choi M, Jacobs JM, Kustas WP (2008). Assessment of clear and cloudy sky parameterizations for daily downwelling longwave radiation over different land surfaces in Florida, USA. *Geophysical Research Letters* 35(20).

de Kok RJ, Steiner JF, Litt M, Wagnon P, Koch I, Azam MF, Immerzeel WW (2020). Measurements, models and drivers of incoming longwave radiation in the Himalaya. *International Journal of Climatology* 40(2):942-956.

DeChant CM, Moradkhani H (2014). Hydrologic prediction and uncertainty quantification. *Handbook of Engineering Hydrology,* CRC press. p 387-414.

Dion P, Martel J-L, Arsenault R (2021). Hydrological ensemble forecasting using a multi-model framework. *Journal of hydrology* 600:126537.

Du Y, Wang T, Zhou Y, Li D, Wang S, Xian Y (2023). Upscaling of longwave downward radiation from instantaneous to any temporal scale: Algorithms, validation, and comparison. *International Journal of Applied Earth Observation and Geoinformation* 117:103196.

Duarte HF, Dias NL, Maggiotto SR (2006). Assessing daytime downward longwave radiation estimates for clear and cloudy skies in Southern Brazil. *Agricultural and Forest Meteorology* 139(3):171-181.

Dugdale SJ, Hannah DM, Malcolm IA (2017a). River temperature modelling: A review of process-based approaches and future directions. *Earth-Science Reviews* 175:97-113.

Ficklin DL, Luo Y, Stewart IT, Maurer EP (2012). Development and application of a hydroclimatological stream temperature model within the Soil and Water Assessment Tool. *Water Resources Research* 48(1):16.

Fniguire F, Laftouhi N-E, Al-Mahfadi AS, El Himer H, Khalil N, Saidi ME (2022). Hydrological modelling using the distributed hydrological model CEQUEAU in a semi-arid mountainous area: a case study of Ourika watershed, Marrakech Atlas, Morocco. *Euro-Mediterranean Journal for Environmental Integration* 7(1):89-102.

Gatien P, Arsenault R, Martel J-L, St-Hilaire A (2022). Using the ERA5 and ERA5-Land reanalysis datasets for river water temperature modelling in a data-scarce region. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 10.1080/07011784.2022.2113917:1-18.

Hani I, St-Hilaire A, Ouarda TBMJ (2023). Machine-learning modeling of hourly potential thermal refuge area: A case study from the Sainte-Marguerite River (Quebec, Canada). *River Research and Applications* 39(9):1763-1782.

Herrero J, Polo M (2012). Parameterization of atmospheric longwave emissivity in a mountainous site for all sky conditions. *Hydrology and Earth System Sciences* 16(9):3139.

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J-N (2020a). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146(730):1999-2049.

Hogan RJ, Bozzo A (2016). *ECRAD: A new radiation scheme for the IFS.* European Centre for Medium-Range Weather Forecasts, (787). 33 p

James CB, Mitchell AC, Buchak EM (1991). *Nechako Reservoir mathematical modelling studies.* Canada, (0-921303-23-8). 236-245 p

Jiang H, Yang Y, Bai Y, Wang H (2020). Evaluation of the Total, Direct, and Diffuse Solar Radiations From the ERA5 Reanalysis Data in China. *IEEE Geoscience and Remote Sensing Letters* 17(1):47-51.

Khorsandi M, St-Hilaire A, Arsenault R (2022). Multisite calibration of a semi-distributed hydrologic and thermal model in a large Canadian watershed. *Hydrological Sciences Journal* 67(14):2147-2174.

Larabi S, Schnorbus MA, Zwiers F (2022). A coupled streamflow and water temperature (VIC-RBM-CE-QUAL-W2) model for the Nechako Reservoir. *Journal of Hydrology: Regional Studies* 44:101237.

Li M, Jiang Y, Coimbra CFM (2017). On the determination of atmospheric longwave irradiance under all-sky conditions. *Solar Energy* 144:40-48.

Long Z, Ding L, Zhou J, Zhou T (2021). Toward the Estimation of All-Weather Daytime Downward Longwave Radiation over the Tibetan Plateau. *Atmosphere* 12(12):1692.

Mai J, Arsenault R, Tolson BA, Latraverse M, Demeester K (2020). Application of Parameter Screening to Derive Optimal Initial State Adjustments for Streamflow Forecasting. *Water Resources Research* 56(9).

McMillan HK, Westerberg IK, Krueger T (2018). Hydrological data uncertainty and its implications. *WIREs Water* 5(6):e1319.

Morin G, Couillard D (1990). Predicting river temperatures with a hydrological model. *Encyclopedia of fluid mechanic, surface and groundwater flow phenomena,* Volk Gulf Publishing Company, Houston, Texas Vol 10. p 171-209.

Morin G, Paquet P (2007). *Modèle hydrologique CEQUEAU.* INRS, Centre Eau, Terre et Environnement, Quebec, Canada. 458 p

Niemelä S, Räisänen P, Savijärvi H (2001). Comparison of surface radiative flux parameterizations: Part I: Longwave radiation. *Atmospheric research* 58(1):1-18.

Nussbaumer EA, Pinker RT (2012). Estimating surface long-wave radiative fluxes at global scale. *Quarterly Journal of the Royal Meteorological Society* 138(665):1083-1093.

Oke TR (1987). *Boundary layer climates.* Routledge,

Ouellet-Proulx S (2018). *Prévision thermique d'ensemble en rivière avec assimilation de données*. Doctoral (Université du Québec, Institut national de la recherche scientifique, Quebec, Canada). 222 p. https://espace.inrs.ca/7621/1/T750.pdf

Ouellet-Proulx S, Chiadjeu OC, Boucher M-A, St-Hilaire A (2017a). Assimilation of water temperature and discharge data for ensemble water temperature forecasting. *Journal of hydrology* 554:342-359.

Ouellet-Proulx S, St-Hilaire A, Boucher M-A (2017b). Water temperature ensemble forecasts: Implementation using the CEQUEAU model on two contrasted river systems. *Water* 9(7):457.

Ouellet-Proulx S, St-Hilaire A, Boucher MA (2019). Implication of evaporative loss estimation methods in discharge and water temperature modelling in cool temperate climates. *Hydrological Processes* 33(22):2867-2884.

Ouellet V, Secretan Y, St-Hilaire A, Morin J (2014). Water temperature modelling in a controlled environment: comparative study of heat budget equations. *Hydrological Processes* 28(2):279-292.

Parkes MW (2022). River conversations: A confluence of lessons and emergence from the Taieri River and the Nechako River. *River Research and Applications* 38(3):443-452.

Picketts IM, Parkes MW, Déry SJ (2017). Climate change and resource development impacts in watersheds: Insights from the Nechako River Basin, Canada. *The Canadian Geographer / Le Géographe canadien* 61(2):196-211.

Pomeroy JW, Marks D, Link T, Ellis C, Hardy J, Rowlands A, Granger R (2009). The impact of coniferous forest temperature on incoming longwave radiation to melting snow. *Hydrological Processes* 23(17):2513-2525.

Prata A (1996). A new long-wave formula for estimating downward clear-sky radiation at the surface. *Quarterly Journal of the Royal Meteorological Society* 122(533):1127-1151.

Sardar T, Xu A, Raziq A (2017). Downward shortwave radiation estimation and spatial assessment on sites over complex terrain applying integrative approach of MTCLIM-XL, interpolation, RS and GIS. *Environment Systems and Decisions* 37(2):198-213.

Satterlund DR (1979). An improved equation for estimating long-wave radiation from the atmosphere. *Water Resources Research* 15(6):1649-1650.

Sheedy B (2005). *Analysis of a cold water release facility in the Nechako Reservoir*. Master of Business Administration (Faculty of Business Administration-Simon Fraser University). 90 p. https://summit.sfu.ca/_flysystem/fedora/sfu_migrate/10292/etd2058.pdf

St-Hilaire A, Boucher M-A, Chebana F, Ouellet-Proulx S, Zhou QX, Larabi S, Dugdale S, Latraverse M (2015). Breathing a new life to an older model: The CEQUEAU tool for flow and water temperature simulations and forecasting. *Proceedings of the 22nd Canadian Hydrotechnical Conference.* (Montréal, Canada, April 29–May 2), p 10.

Taghizadeh E, Ahmadi-Givi F, Brocca L, Sharifi E (2021). Evaluation of satellite/reanalysis precipitation products over Iran. *International Journal of Remote Sensing* 42(9):3474-3497.

Tang W, Qin J, Yang K, Zhu F, Zhou X (2021). Does ERA5 outperform satellite products in estimating atmospheric downward longwave radiation at the surface? *Atmospheric research* 252:105453.

Tarek M, Brissette FP, Arsenault R (2020b). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.* 24(5):2527-2544.

Verseghy DL, McFarlane NA, Lazare M (1993). Class—A Canadian land surface scheme for GCMS, II. Vegetation model and coupled runs. *International Journal of Climatology* 13(4):347-370.

White JC, Khamis K, Dugdale S, Jackson FL, Malcolm IA, Krause S, Hannah DM (2023). Drought impacts on river water temperature: A process-based understanding from temperate climates. *Hydrological Processes* 37(10):e14958.

Yao B, Liu C, Yin Y, Liu Z, Shi C, Iwabuchi H, Weng F (2020). Evaluation of cloud properties from reanalyses over East Asia with a radiance-based approach. *Atmos. Meas. Tech.* 13(3):1033-1049.

APPENDIX I

Appendices for Article I: Multisite calibration of a semi-distributed hydrologic and thermal model in a large Canadian watershed

Appendix I-A: Understanding parameters in the CEQUEAU hydrological module

Table I-A1. Parameters used in the CEQUEAU hydrological module

No.	Definition	Lower Bound	Upper Bound	Calibrated Value	Unit	Symbol used in CEQUEAU
1	Percolation coefficient from the upper zone to the lower zone	0.01	0.7	0.01	dimensionless	CIN
2	Lakes and marshes drainage coefficient	0.01	1	0.1545	dimensionless	CVMAR
3	Lower-zone lower drainage coefficient	2E-5	8E-4	2E-5	dimensionless	CVNB
4	Lower-zone upper drainage coefficient	2E-5	1E-3	2.56E-5	dimensionless	CVNH
5	Upper-zone lower drainage coefficient	1E-4	0.005	1E-4	dimensionless	CVSB
6	Upper-zone intermediate drainage coefficient	1E-4	1	0.327	dimensionless	CVSI
7	Daily maximum infiltration	0	40	17.1	mm/day	XINFMA
8	Percolation threshold from the upper to the lower zone	0	100	7.2	mm	HINF
9	Upper-zone intermediate drainage threshold	10	100	100	mm	HINT
10	Lakes and marshes drainage threshold	100	500	460.2	mm	HMAR
11	Lower-zone upper threshold	100	500	258	mm	HNAP
12	Threshold of evaporation at the potential rate	20	200	20	mm	HPOT
13	Upper-zone runoff threshold	0	300	163.4	mm	HSOL
14	Upper-zone runoff threshold for impermeable surfaces	0	10	6.75	mm	HRIMP
15	Routing coefficient fitting parameter	0.002	0.01	0.0051	dimensionless	EXXKT
16	Rain-snow threshold	-2	3	3	°C	STRNE
17	Melting rate in a forested area	0	10	1.23	mm/°C/day	TFC
18	Melting rate in an open area	0	10	5.6	mm/°C/day	TFD
19	Melting threshold in a forested area	-2	3	-.86	°C	TSC
20	Melting threshold in an open area	-2	3	1.27	°C	TSD
21	The cold content coefficient for snowpack	0	3	0		TTD
22	Priming threshold of snowpack	-5	2	1.9	°C	TTS
23	Percentage of daily evapotranspiration taken from the Lower zone	0	0.4	0	%	EVNAP
24	The exponent of Thornthwaite formula	0	2	0.526	dimensionless	XAA
25	Thermal index of Thornthwaite formula	5	40	24.46	dimensionless	XIT
26	Concentration time, i.e. time for water to travel from the highest upstream point to the outlet	0	10	1.06	days	ZN

Appendix I-B: The four steps of the CMA-ES algorithm

Step 1- Initialization, which is defining initial values:

$$m \in R^n, \sigma \in R, \lambda \in \{2\ 3\ 4\ \ldots\} \tag{B1}$$

which m = initial solution, which after the first iteration would be the mean of best solutions' distribution of the last iteration; σ = initial step size, which will be updated for the next steps; λ = initial population size of points around m; and n = number of decision variables. m and σ are problem-dependent, and usually, $\lambda \geq 5$ is used. Some other values should be initialized using the above values in step 2.

Step 2- Initialization of algorithm setting values as follows:

$$p_c = 0 \tag{B2}$$

$$p_\sigma = 0 \tag{B3}$$

$$c_c \approx \frac{4}{n} \tag{B4}$$

$$c_\sigma \approx \frac{3}{n} \tag{B5}$$

$$c_1 \approx \frac{2}{n^2} \tag{B6}$$

$$c_\mu \approx \frac{\mu_w}{n^2} \tag{B7}$$

$$d_\sigma \approx 1 \tag{B8}$$

$$C = I \tag{B9}$$

where p_c = anisotropic evolution path; c_c = the backward time horizon for p_c evolution path; p_σ = isotropic evolution path; c_σ = the backward time horizon for the p_σ evolution path; c_1= the learning rate for the rank-one update of the covariance matrix; c_μ= the learning rate for the rank-μ update of the covariance matrix and must not exceed $1 - c_1$, d_σ = the damping parameter, and C = initial covariance matrix equal to the identity matrix (I).

Step 3- Set initial weights as follows:

$$w_{i=1,\ldots,\lambda} \text{ in a decreasing form while } \sum_i |w_i| = 1 \text{ and } \sum_i w_i^2 \approx \frac{3}{\lambda} \tag{B10}$$

$$\mu_w^{-1} = \sum_i w_i^2 \tag{B11}$$

Step 4- With all required values set as above and knowing inputs, the algorithm will perform the following calculation in an iterative form till stopping criteria will happen:

Sampling $$x_i = m + \sigma \times y_i \sim N(m\ \sigma^2 C) \quad for\ i = 1\ \dots\ \lambda \tag{B12}$$

Update mean $$m \leftarrow m + \sigma \sum_i w_{\rho(i)} y_i = m + \sigma y_w \tag{B13}$$

Path for σ $$p_\sigma \leftarrow (1 - c_\sigma) p_\sigma + \sqrt{1 - (1 - p_\sigma)^2} \sqrt{\mu_w} C^{-\frac{1}{2}} y_w \tag{B14}$$

Update for σ $$\sigma \leftarrow \sigma \times \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|p_\sigma\|}{E\|N(0\ 1)\|} - 1\right)\right) \tag{B15}$$

Path for C $$p_C \leftarrow (1 - c_c) p_C + 1_{[0\ 2n]}\{\|p_\sigma\|^2\} \sqrt{1 - (1 - c_c)^2} \sqrt{\mu_w}, y_w \tag{B16}$$

Update for C $$C \leftarrow \left(1 - c_1 - c_\mu\right) C + c_\mu \sum_{i=1}^{\lambda} w_{\rho(i)} y_i y_i^T + c_1 p_C p_C^T \tag{B17}$$

where $\rho(i)$ is the index after sorting the values. This algorithm, after reaching to stopping criteria returns $m_{optimal}$.

Appendix I-C: Simulated time series with single site calibration parameters for each station

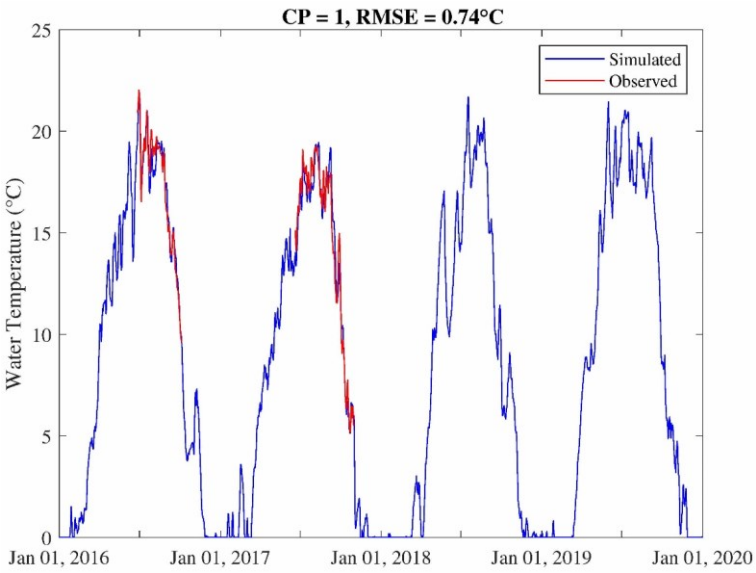

Figure I-C1 The simulated time series for station 1 using the single site calibration parameters for station 1

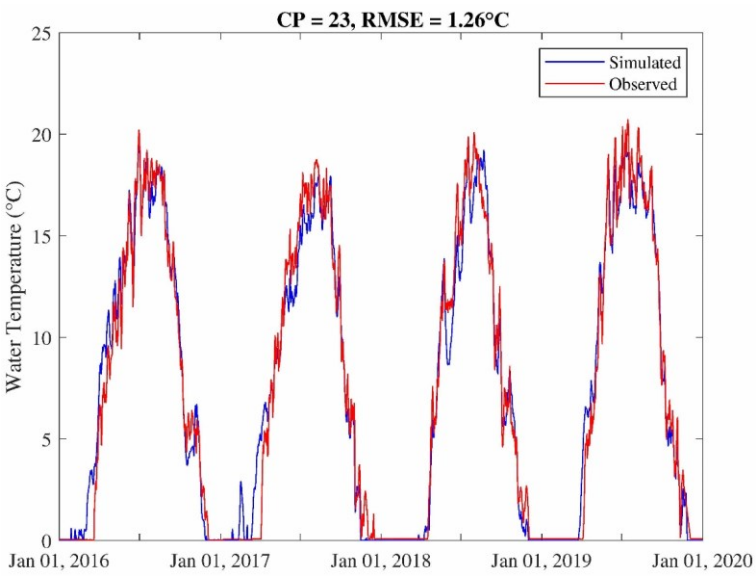

Figure I-C2 The simulated time series for station 2 using the single site calibration parameters for station 2

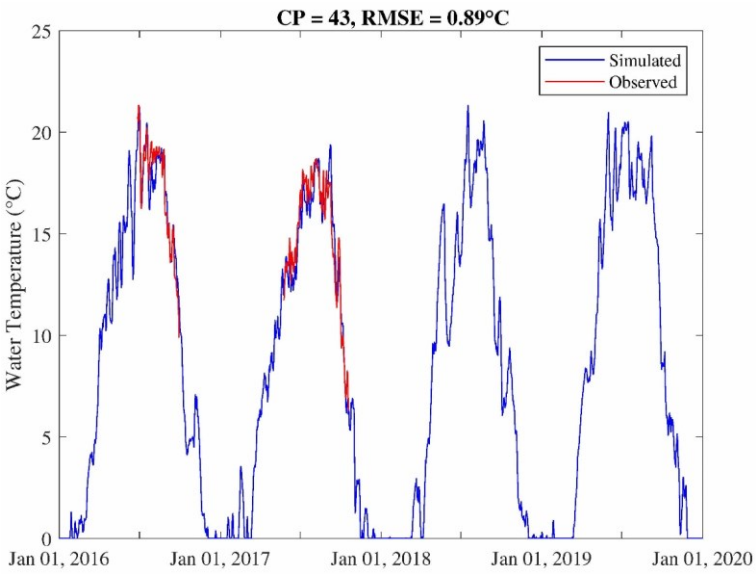

Figure I-C3 The simulated time series for station 3 using the single site calibration parameters for station 3

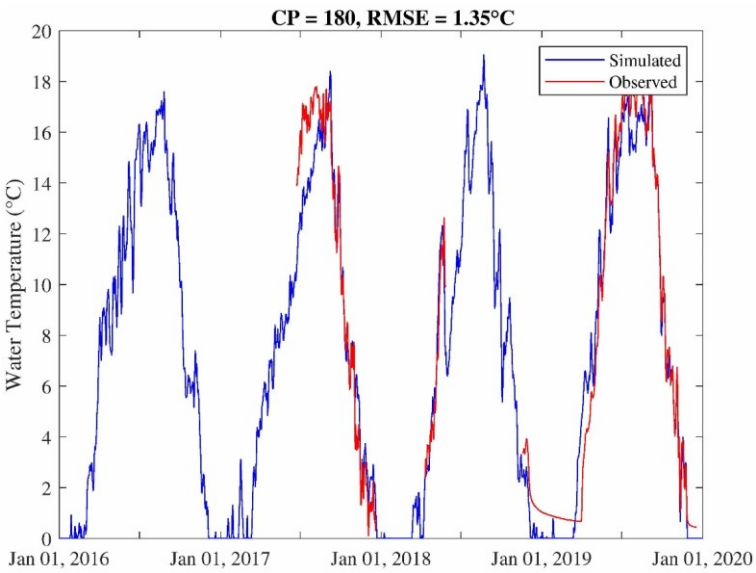

Figure I-C4 The simulated time series for station 4 using the single site calibration parameters for station 4

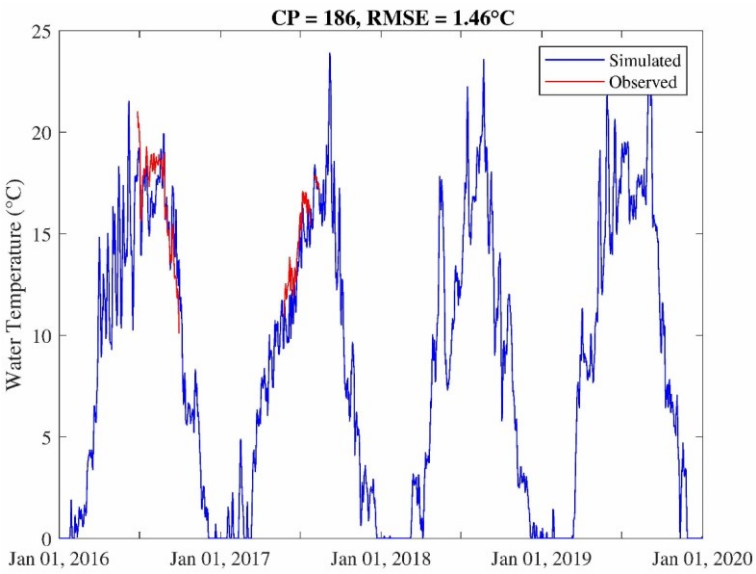

Figure I-C5 The simulated time series for station 5 using the single site calibration parameters for station 5

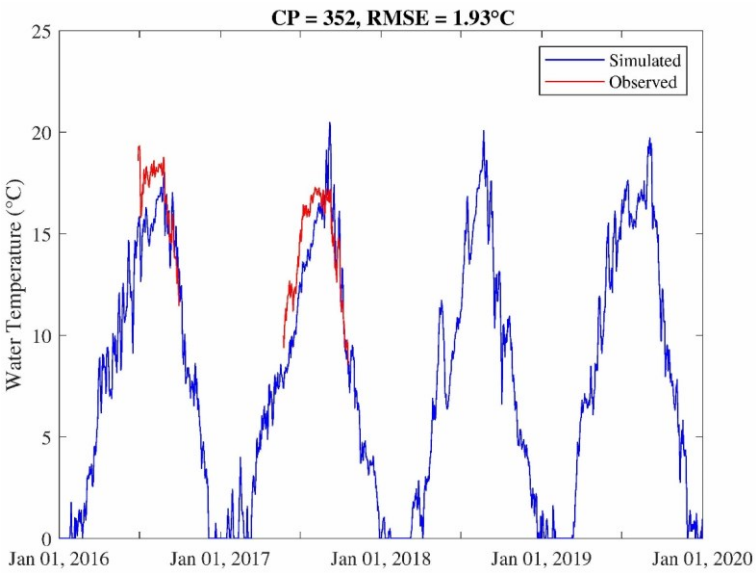

Figure I-C6 The simulated time series for station 6 using the single site calibration parameters for station 6

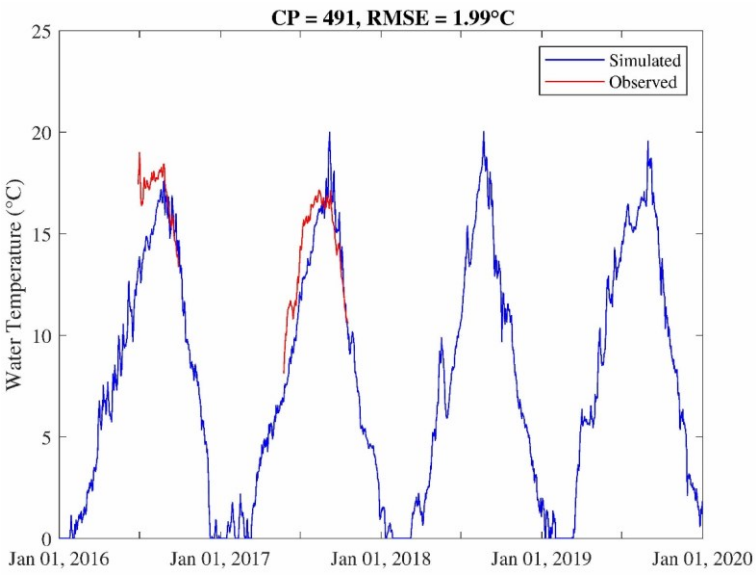

Figure I-C7 The simulated time series for station 7 using the single site calibration parameters for station 7

Appendix I-D: Simulated time series with multisite calibration parameters for each station

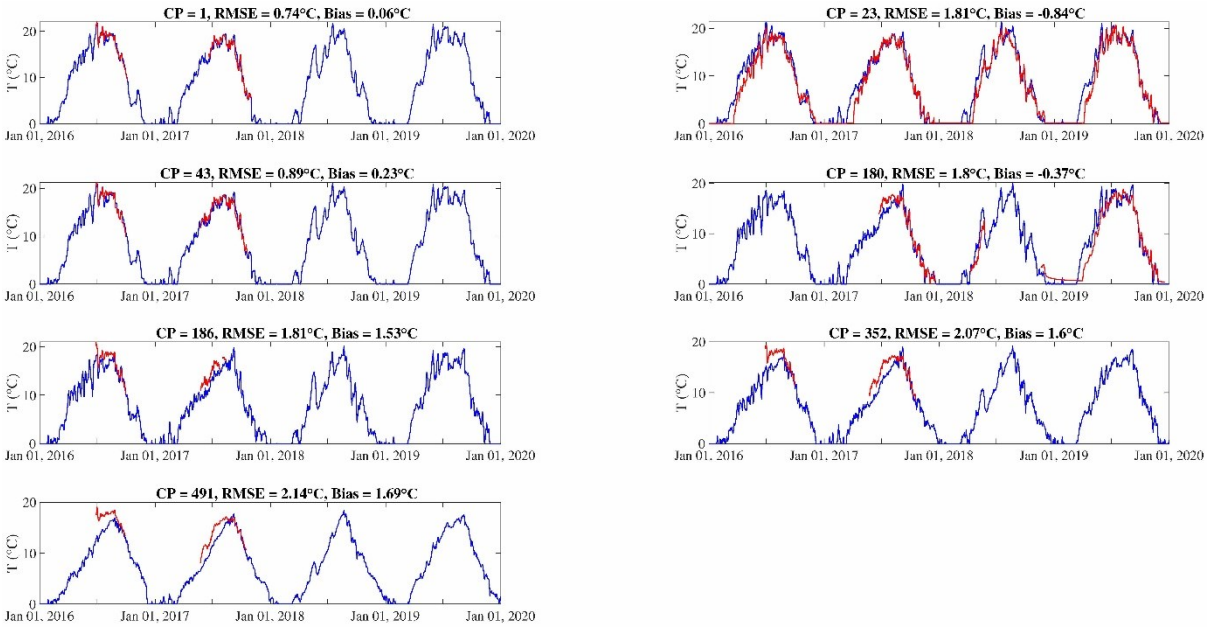

Figure I-D1 The simulated time series for all stations using the single site calibration parameters of station 1

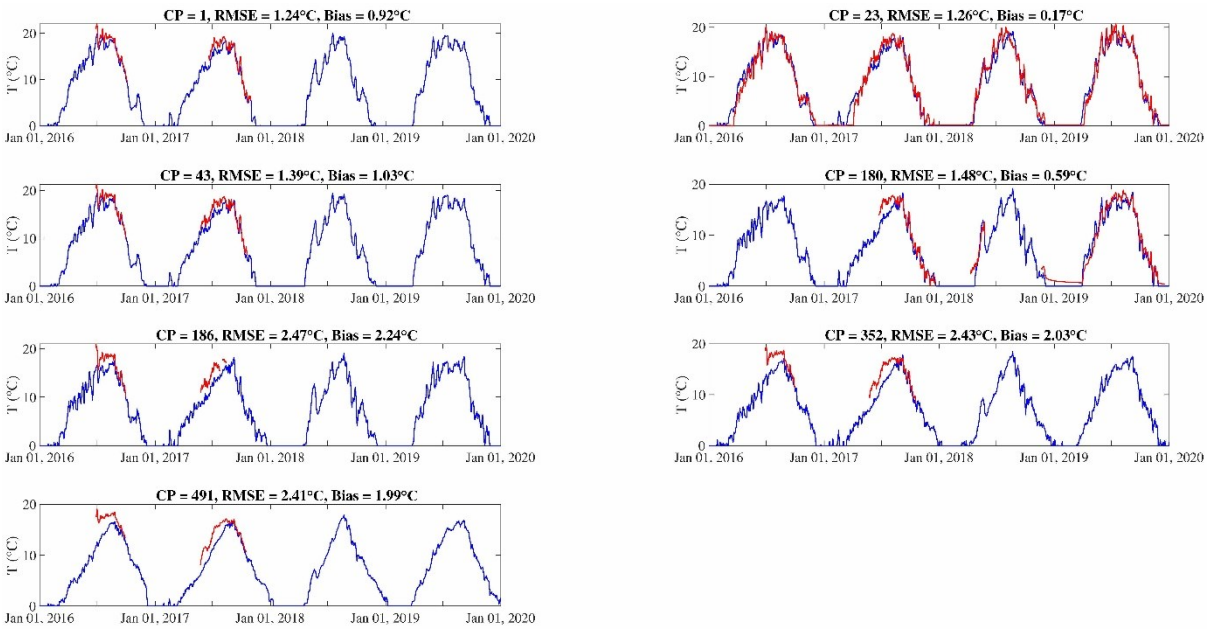

Figure I-D2 The simulated time series for all stations using the single site calibration parameters of station 2

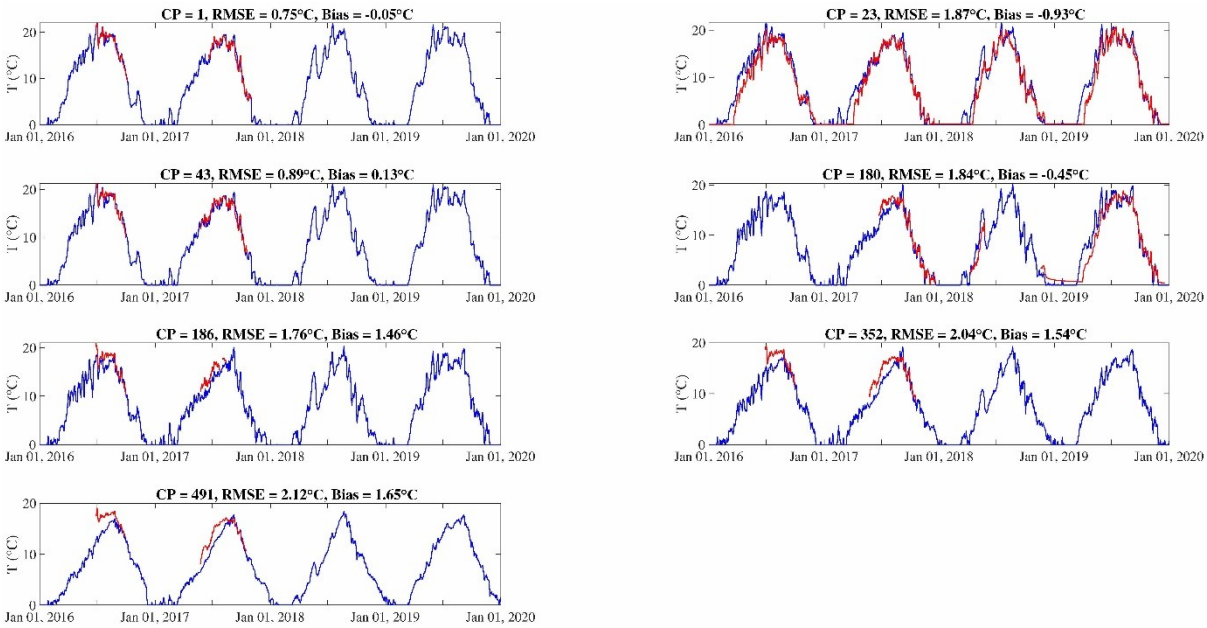

Figure I-D3 The simulated time series for all stations using the single site calibration parameters of station 3

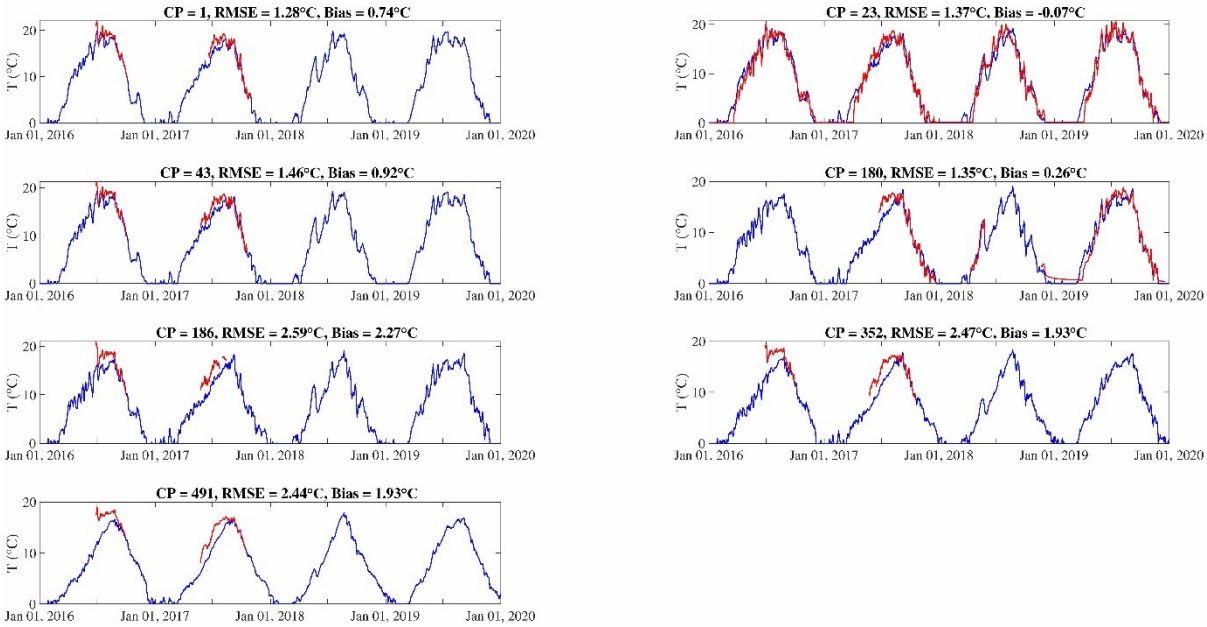

Figure I-D4 The simulated time series for all stations using the single site calibration parameters of station 4

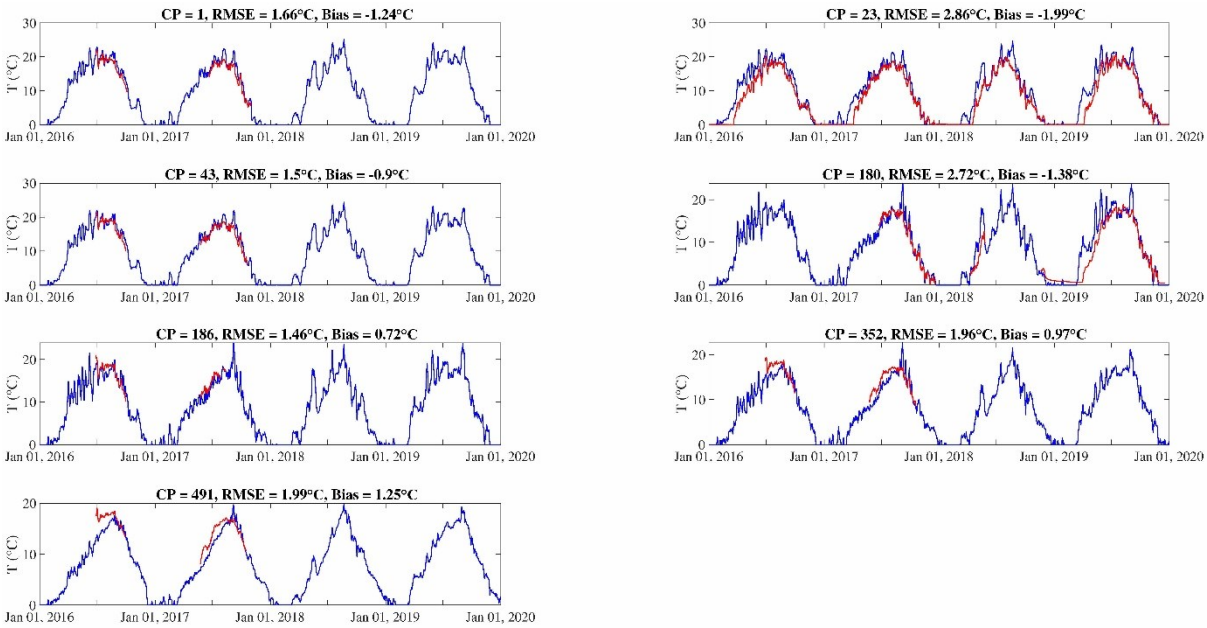

Figure I-D5 The simulated time series for all stations using the single site calibration parameters of station 5

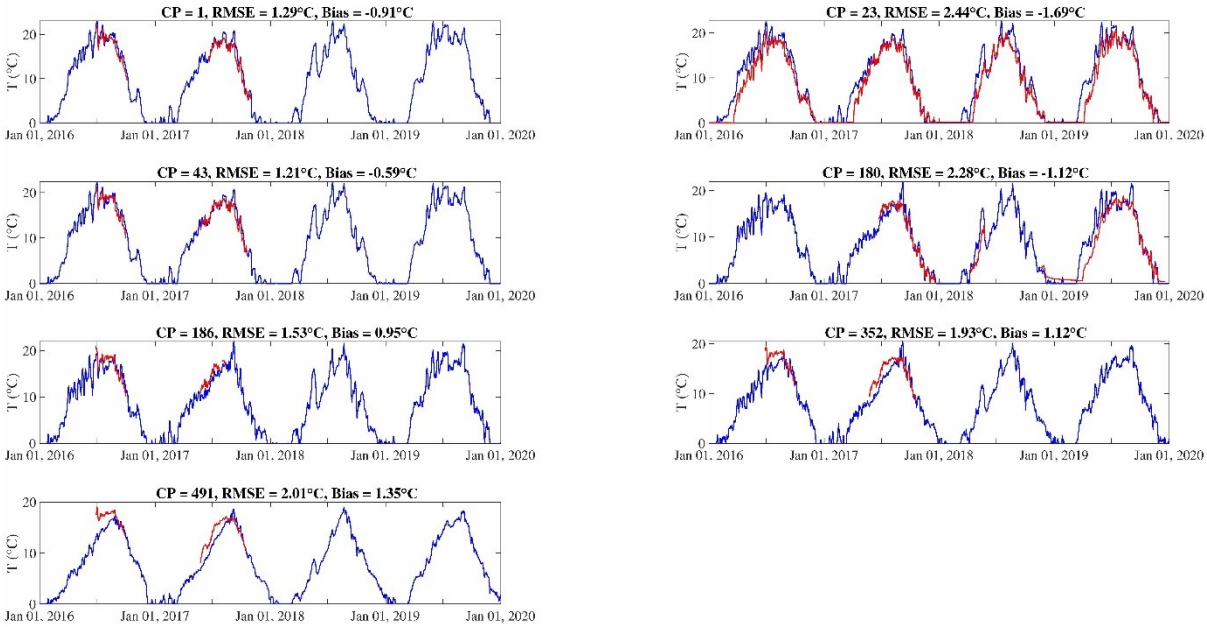

Figure I-D6 The simulated time series for all stations using the single site calibration parameters of station 6

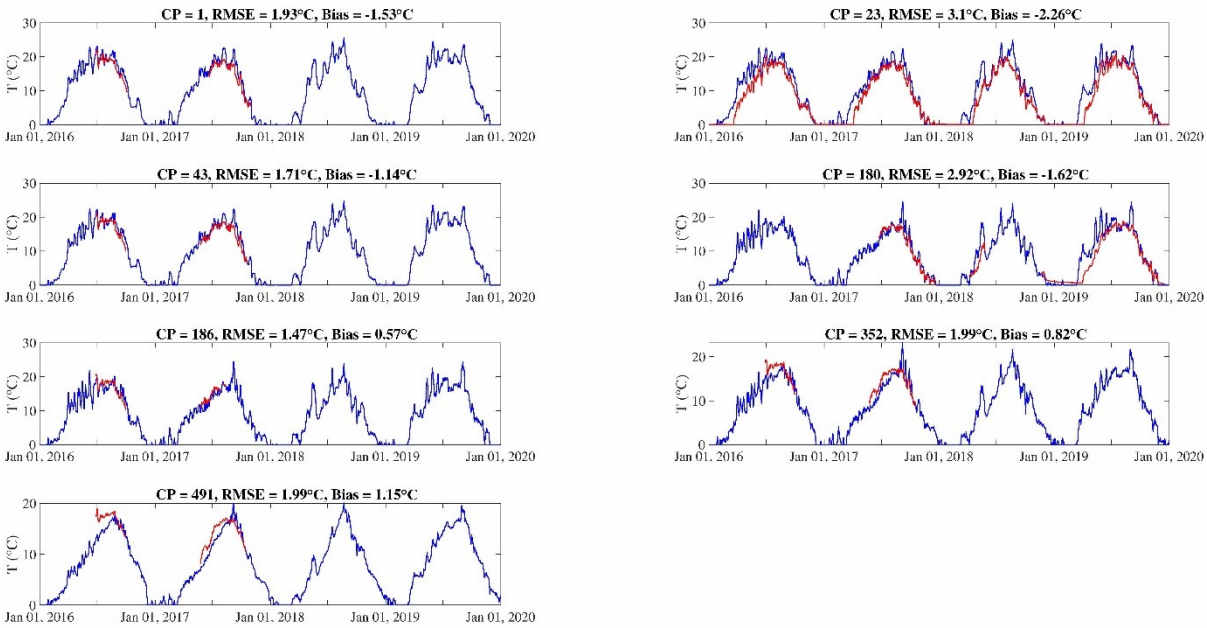

Figure I-D7 The simulated time series for all stations using the single site calibration parameters of station 7

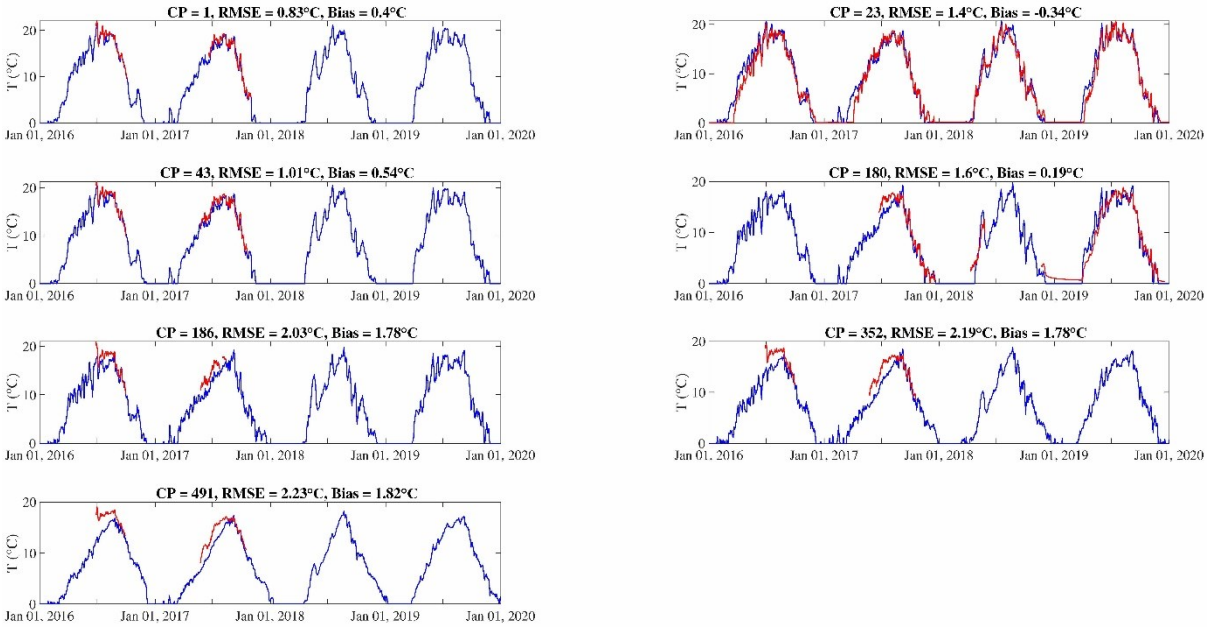

Figure I-D8 The simulated time series for all stations using the multisite calibration parameters

Appendix I-E: The MATLAB codes implemented in Article I

Single-site thermal calibration – CMAES setting and code

```
%% This code is for thermal Auto Calibration with CMAES method for 1 Station
clear all
clc
%% Declare and Load inputs
global InputStruct
load 'InputStruct.mat'
%% boundary conditions and observed data
for t = 1:1461
    InputStruct.assimilations.qualite(t).etatsCP.temperature(172) = InputStruct.NautleyTemp2016_2019(t, 2);
    InputStruct.assimilations.qualite(t).etatsCP.temperature(1109) = InputStruct.SpillwayTemp2016_2019(t, 2);
end
%% Hydrological parameters from previous optimization
load 'x_Hydro.mat'
%% Hydrological Optimization Parameters
InputStruct.parametres.sol.cin_s = x_Hydro(1);
InputStruct.parametres.sol.cvmar = x_Hydro(2);
InputStruct.parametres.sol.cvnb_s = x_Hydro(3);
InputStruct.parametres.sol.cvnh_s = x_Hydro(4);
InputStruct.parametres.sol.cvsb = x_Hydro(5);
InputStruct.parametres.sol.cvsi_s = x_Hydro(6);
InputStruct.parametres.sol.xinfma = x_Hydro(7);
InputStruct.parametres.sol.hinf_s = x_Hydro(8);
InputStruct.parametres.sol.hint_s = x_Hydro(9);
InputStruct.parametres.sol.hmar = x_Hydro(10);
InputStruct.parametres.sol.hnap_s = x_Hydro(11);
InputStruct.parametres.sol.hpot_s = x_Hydro(12);
InputStruct.parametres.sol.hsol_s = x_Hydro(13);
InputStruct.parametres.sol.hrimp_s = x_Hydro(14);
InputStruct.parametres.transfert.exxkt = x_Hydro(15);
InputStruct.parametres.neige.strne_s = x_Hydro(16);
InputStruct.parametres.fonte.cequeau.strne_s = x_Hydro(16);
InputStruct.parametres.neige.tfc_s = x_Hydro(17);
InputStruct.parametres.fonte.cequeau.tfc_s = x_Hydro(17);
InputStruct.parametres.neige.tfd_s = x_Hydro(18);
InputStruct.parametres.fonte.cequeau.tfd_s = x_Hydro(18);
InputStruct.parametres.neige.tsc_s = x_Hydro(19);
InputStruct.parametres.fonte.cequeau.tsc_s = x_Hydro(19);
InputStruct.parametres.neige.tsd_s = x_Hydro(20);
InputStruct.parametres.fonte.cequeau.tsd_s = x_Hydro(20);
InputStruct.parametres.neige.ttd = x_Hydro(21);
InputStruct.parametres.fonte.cequeau.ttd = x_Hydro(21);
InputStruct.parametres.neige.tts_s = x_Hydro(22);
InputStruct.parametres.fonte.cequeau.tts_s = x_Hydro(22);
InputStruct.parametres.evapo.cequeau.evnap = x_Hydro(23);
InputStruct.parametres.evapo.cequeau.xaa = x_Hydro(24);
InputStruct.parametres.evapo.cequeau.xit = x_Hydro(25);
InputStruct.parametres.transfert.zn = x_Hydro(26);
%% Setting the initials
InputStruct.parametres.solInitial.hsini = x_Hydro(9);
InputStruct.parametres.solInitial.hnini = x_Hydro(11);
InputStruct.parametres.solInitial.hmini = x_Hydro(10);
InputStruct.parametres.solInitial.q0    = 10;

lb = [0.05  0.05    0.05    0.05    0   5   5   0];
ub = [2 2   2   2   1000    10  20  1];
x  = (lb+0.5.*(ub-lb))';
sigma = (ub-lb)';
opts.LBounds = lb';
opts.UBounds = ub';
opts.MaxFunEvals  = 5000;
[xmin, fmin, counteval, stopflag, out, bestever]= cmaes('fun', x, sigma, opts);
```

Objective function for single site thermal calibration (for CP = 23)

```
%% this scrip is function for single staion thermal optimization
function f_RMSE = fun(x)
global InputStruct
InputStruct.parametres.qualite.cequeau.temperat.crayso = x(1);
InputStruct.parametres.qualite.cequeau.temperat.crayin = x(2);
InputStruct.parametres.qualite.cequeau.temperat.cevapo = x(3);
InputStruct.parametres.qualite.cequeau.temperat.cconve = x(4);
InputStruct.parametres.qualite.cequeau.temperat.crigel = x(5);
InputStruct.parametres.qualite.cequeau.temperat.tnap = x(6);
InputStruct.parametres.qualite.cequeau.temperat.bassol = x(7);
InputStruct.parametres.qualite.cequeau.temperat.corsol = x(8);
% model simulation
% Use of the thermal module:
thermie = 1;
% Dates for which we want to simulate, calibrate, validate and illustrate the results:
startSim = datenum(2016, 01, 01);
endSim = datenum(2019, 12, 31);
dates.sim = [startSim:endSim]';
% first and end row of calibration period
% Settings to adjust:
%valeursParametres_thermie;
InputStruct.execution.dateDebut = dates.sim(1);
InputStruct.execution.dateFin = dates.sim(end);
InputStruct.parametres.option.calculQualite = thermie;
%Simulation
[y.etatsCE, y.etatsCP, y.etatsFonte, y.etatsEvapo, y.etatsBarrage, y.pasDeTemps,...
    y.avantAssimilationsCE, y.avantAssimilationsFonte,
y.avantAssimilationsEvapo,y.etatsQualCP, y.avAssimQual] = ...
cequeauQuantiteMex_v461(InputStruct.execution, InputStruct.parametres,
InputStruct.bassinVersant, ...
     InputStruct.meteoPointGrille, [], InputStruct.assimilations);
timesteps = dates.sim(end)-dates.sim(1) + 2;
%% CP
simCP = 23; % our station
Qsim = zeros(timesteps-1, 1);
Tsim = zeros(timesteps-1, 1);
for t = 2:timesteps
    Qsim(t-1,1) = y.etatsCP(t).debit(simCP);
    Tsim(t-1,1) = y.etatsQualCP(t).temperature(simCP);
end
%% Observed Values
Tobs = InputStruct.temp_CP23(:,2);
temps(:,1) = [InputStruct.meteoStation.t];
temps(f_row:e_row,2) = Tobs;
temps(:,3) = Tsim;
cal_temps = temps(f_row:e_row,:);
%% extracting specific flow months
mah = datevec(cal_temps(:,1));
specific_temps = cal_temps(ismember(mah(:,2),[4:11]),:);
%%
f_RMSE  = RMSE(specific_temps(:,2), specific_temps(:,3))
end
```

Objective function for multisite thermal calibration

```
1-  %% this scrip is function for multisite thermal optimization
2-  function f_RMSE = fun(x)
3-  global InputStruct
4-  InputStruct.parametres.qualite.cequeau.temperat.crayso = x(1);
5-  InputStruct.parametres.qualite.cequeau.temperat.crayin = x(2);
6-  InputStruct.parametres.qualite.cequeau.temperat.cevapo = x(3);
7-  InputStruct.parametres.qualite.cequeau.temperat.cconve = x(4);
8-  InputStruct.parametres.qualite.cequeau.temperat.crigel = x(5);
9-  InputStruct.parametres.qualite.cequeau.temperat.tnap = x(6);
10- InputStruct.parametres.qualite.cequeau.temperat.bassol = x(7);
11- InputStruct.parametres.qualite.cequeau.temperat.corsol = x(8);
12- % model simulation. Use of the thermal module:
13- thermie = 1;
14- % Dates for which we want to simulate, calibrate, validate and illustrate the results:
15- startSim = datenum(2016, 01, 01);
16- endSim = datenum(2019, 12, 31);
17- dates.sim = [startSim:endSim]';
18- InputStruct.execution.dateDebut = dates.sim(1);
19- InputStruct.execution.dateFin = dates.sim(end);
20- InputStruct.parametres.option.calculQualite = thermie;
21- %Simulation
22- [y.etatsCE, y.etatsCP, y.etatsFonte, y.etatsEvapo, y.etatsBarrage, y.pasDeTemps,...
23-     y.avantAssimilationsCE, y.avantAssimilationsFonte,
    y.avantAssimilationsEvapo,y.etatsQualCP, y.avAssimQual] = ...
24- cequeauQuantiteMex_v461(InputStruct.execution, InputStruct.parametres,
    InputStruct.bassinVersant, ...
25-      InputStruct.meteoPointGrille, [], InputStruct.assimilations);
26- %%
27- timesteps = dates.sim(end)-dates.sim(1) + 2;
28- %% CP
29- Tsim1 = zeros(timesteps-1, 1);
30- Tsim23 = zeros(timesteps-1, 1);
31- Tsim43 = zeros(timesteps-1, 1);
32- Tsim180 = zeros(timesteps-1, 1);
33- Tsim186 = zeros(timesteps-1, 1);
34- Tsim352 = zeros(timesteps-1, 1);
35- Tsim491 = zeros(timesteps-1, 1);
36- for t = 2:timesteps
37-     Tsim1(t-1,1) = y.etatsQualCP(t).temperature(1);
38-     Tsim23(t-1,1) = y.etatsQualCP(t).temperature(23);
39-     Tsim43(t-1,1) = y.etatsQualCP(t).temperature(43);
40-     Tsim180(t-1,1) = y.etatsQualCP(t).temperature(180);
41-     Tsim186(t-1,1) = y.etatsQualCP(t).temperature(186);
42-     Tsim352(t-1,1) = y.etatsQualCP(t).temperature(352);
43-     Tsim491(t-1,1) = y.etatsQualCP(t).temperature(491);
44- end
45- Tsim = cat(1,Tsim1,Tsim23,Tsim43,Tsim180,Tsim186,Tsim352,Tsim491);
46- Tobs1 = InputStruct.temp_CP1(:,2);
47- Tobs23 = InputStruct.temp_CP23(:,2);
48- Tobs43 = InputStruct.temp_CP43(:,2);
49- Tobs180 = InputStruct.temp_CP180(:,2);
50- Tobs186 = InputStruct.temp_CP186(:,2);
51- Tobs352 = InputStruct.temp_CP352(:,2);
52- Tobs491 = InputStruct.temp_CP491(:,2);
53-
54- Tobs = cat(1,Tobs1,Tobs23,Tobs43,Tobs180,Tobs186,Tobs352,Tobs491);
55- T1 = InputStruct.temp_CP1(:,1);
56- T23 = InputStruct.temp_CP23(:,1);
57- T43 = InputStruct.temp_CP43(:,1);
58- T180 = InputStruct.temp_CP180(:,1);
59- T186 = InputStruct.temp_CP186(:,1);
60- T352 = InputStruct.temp_CP352(:,1);
61- T491 = InputStruct.temp_CP491(:,1);
62-
63- T = cat(1,T1,T23,T43,T180,T186,T352,T491);
64- temps(:,1) = T;
65- temps(:,2) = Tobs;
66- temps(:,3) = Tsim;
67- cal_temps = temps;
68- %% extracting specific flow months
69- mah = datevec(cal_temps(:,1));
70- specific_temps = cal_temps(ismember(mah(:,2),[4:11]),:);
71- f_RMSE  = RMSE(specific_temps(:,2), specific_temps(:,3))
72- end
```

APPENDIX II

Appendices for Article II: Future flow and water temperature scenarios in an impounded drainage basin, Implications for summer flow and temperature management downstream of the dam

Appendix II-A: Supplementary information, including data for species characteristics, climate, hydrology, and calibration in the study

Table II-A1. Characteristics of Sockeye salmon fish species in this study

Scientific name	*Oncorhynchus nerka*
Maximum tolerable temperature	19°C (20-30% mortality) 20°C (40% mortality) 21°C (100% mortality 72 hours following exhaustive exercise) (Middleton et al., 2018; Robinson et al., 2015)
Current Summer Temperature Monitoring Program (STMP) limit	20°C
Spawning period	August to September

Table II-A2. Parameters of the hydrologic module after calibration using the multisite calibration method

No.	Parameter	Unit	Value	Description
1	CIN	dimensionless	0.09	Percolation coefficient from the upper zone to the lower zone
2	CVMAR	dimensionless	0.04	Lakes and marshes drainage coefficient
3	CVNB	dimensionless	0.0002	Lower-zone lower drainage coefficient
4	CVNH	dimensionless	0.0008	Lower-zone upper drainage coefficient
5	CVSB	dimensionless	0.0002	Upper-zone lower drainage coefficient
6	CVSI	dimensionless	0.056	Upper-zone intermediate drainage coefficient
7	XINFMA	mm/day	2.30	Daily maximum infiltration
8	HINF	mm	49.13	Percolation threshold from the upper to the lower zone
9	HINT	mm	99.19	Upper-zone intermediate drainage threshold
10	HMAR	mm	263.24	Lakes and marshes drainage threshold
11	HNAP	mm	362.54	Lower-zone upper threshold
12	HPOT	mm	133.40	Threshold of evaporation at the potential rate
13	HSOL	mm	299.15	Upper-zone runoff threshold
14	HRIMP	mm	4.69	Upper-zone runoff threshold for impermeable surfaces
15	EXXKT	dimensionless	0.006	Routing coefficient fitting parameter
16	STRNE	°C	2.94	Rain-snow threshold
17	TFC	mm/°C/day	1.07	Melting rate in forested area
18	TFD	mm/°C/day	2.57	Melting rate in open area
19	TSC	°C	0.018	Melting threshold in forested area
20	TSD	°C	0.43	Melting threshold in open area
21	TTD		0.99	Cold content coefficient for snowpack
22	TTS	°C	-0.59	Priming threshold of snowpack
23	EVNAP	%	0.38	Percentage of daily evapotranspiration taken from Lower zone
24	XAA	dimensionless	1.42	Exponent of Thornthwaite formula
25	XIT	dimensionless	25.88	Thermal index of Thornthwaite formula
26	ZN	days	2.68	Concentration time, i.e. time for water to travel from the highest upstream point to the outlet

Table II-A3. Parameters of the thermal module after calibration using the multisite calibration method

No.	Parameter	Unit	Value	Description
1	CRAYSO	dimensionless	1.28	Coefficient for shortwave radiation (ratio)
2	CRAYIN	dimensionless	0.61	Coefficient for longwave radiation (ratio)
3	CEVAPO	dimensionless	0.22	Coefficient for evaporation (ratio)
4	CCONVE	dimensionless	2.00	Coefficient for sensible heat (ratio)
5	CRIGEL	mm	422.66	Threshold snow stock for controlling water temperature
6	TNAP	°C	6.99	Temperature of groundwater
7	BASSOL	dimensionless	6.89	Amount of precipitation that will identify days with low solar radiation
8	CORSOL	dimensionless	0.00	Correction factor for bassol (ratio)

Table II-A4. Eight global climate models used in this study to produce future climate data

No.	GCM	Full Name	Modeling Center	Spatial Resolution
1	BCC-CSM2-MR	Beijing Climate Center Climate System Model	Beijing Climate Center (BCC)	110 × 110 km
2	CMCC-CM2-SR5	Euro-Mediterranean Centre on Climate Change coupled climate model	Euro-Mediterranean Centre on Climate Change	0.9° lat × 1.25° lon 100 × 140 km
3	CMCC-ESM2	Second-generation CMCC Earth System Model	Euro-Mediterranean Centre on Climate Change	0.9° lat × 1.25° lon 100 × 140 km
4	EC-Earth3	European Centre Earth3 Model	European Centre for Medium-Range Weather Forecasts (ECMWF)	40 × 40 km
5	MIROC6	Model for Interdisciplinary Research on Climate	Japanese modeling community	1.4° lat × 1.4° lon 150 × 150 km
6	MPI-ESM1-2-HR	Max Planck Institute for Meteorology Earth System Higher-resolution Model	Max Planck Institute	100 × 100 km
7	MPI-ESM1-2-LR	Max Planck Institute for Meteorology Earth System Lower-resolution Model	Max Planck Institute	200 × 200 km
8	MRI-ESM2-0	Meteorological Research Institute Earth System Model version 2.0	Meteorological Research Institute	100 × 100 km

Table II-A5. Shared Socioeconomic Pathways (SSP) scenario descriptions

Scenario	Description
SSP2-4.5	Emissions in SSP2-4.5 will peak around 2040 at ~ 4.5 W/m^2, then decline and stabilize after the year 2100
SSP5- 8.5	Emissions in SSP5-8.5 emissions continue to rise throughout the 21st century to ~ 8.5 W/m^2 by the year 2100

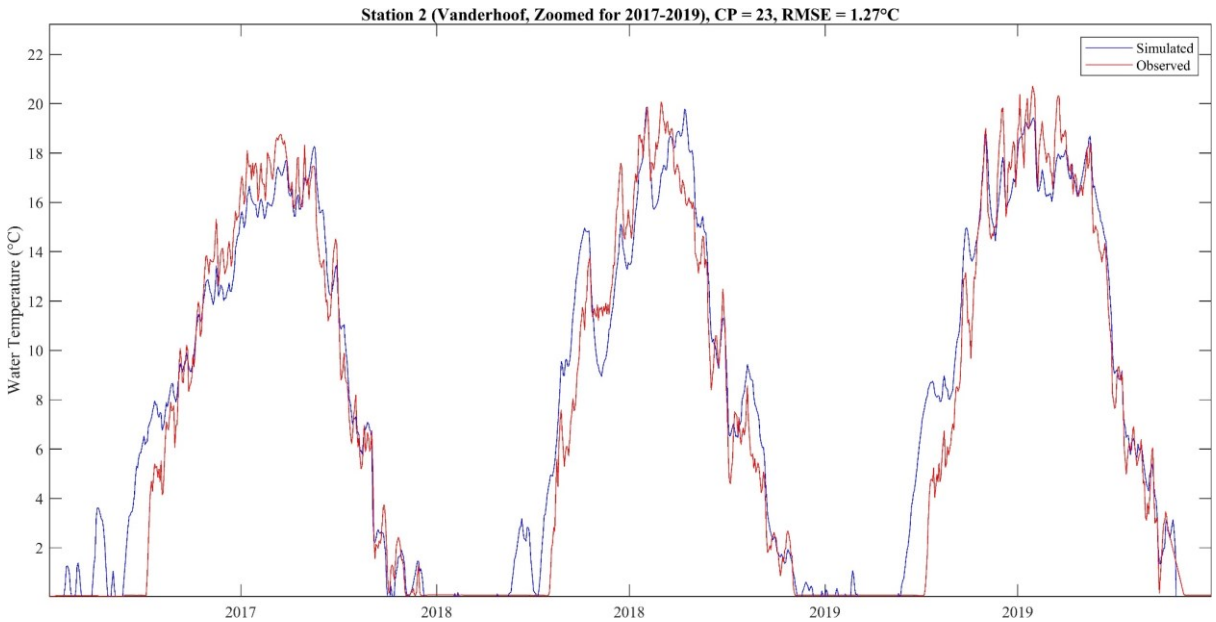

Figure II-A1 The simulated (blue) and observed (red) water temperature time series for Station 2 (Vanderhoof) with CP = 23, using the multisite calibration approach for the June-September period.

Appendix II-B: Climate simulation figures and duration curve analysis for Vanderhoof station

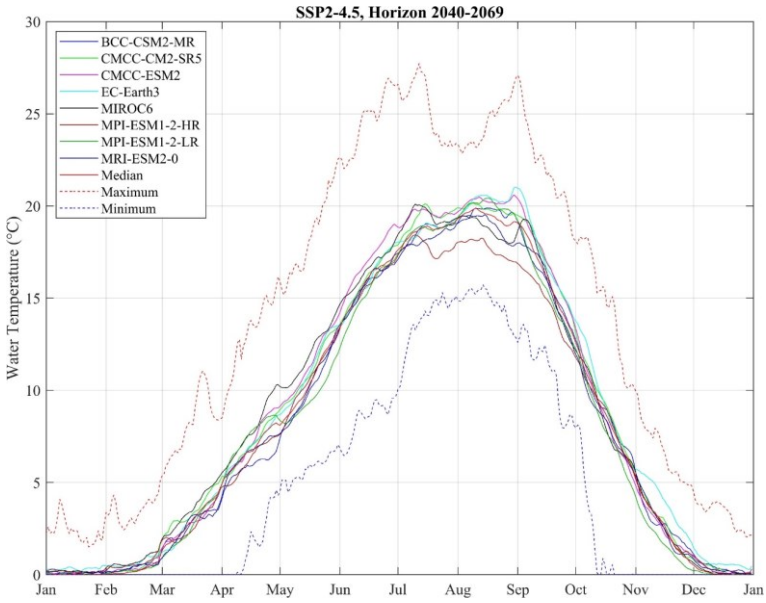

Figure II-B1 Simulated water temperature for Vanderhoof station using eight climate models with SSP2-4.5 for the 2040-2069 future time horizon and the median of eight-time series (red line) together with the lower and upper boundary of simulated values (blue and red dotted lines)

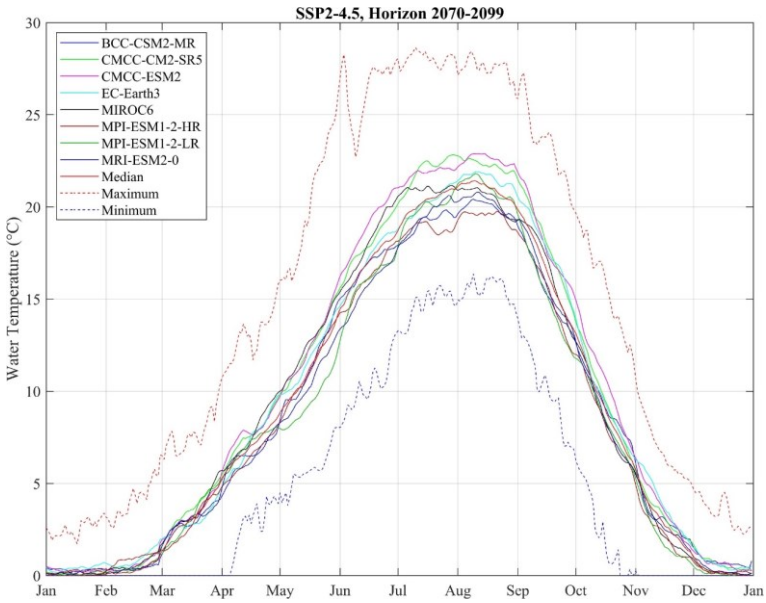

Figure II-B2 Same as Figure II-B1, but for the 2070–2099-time horizon

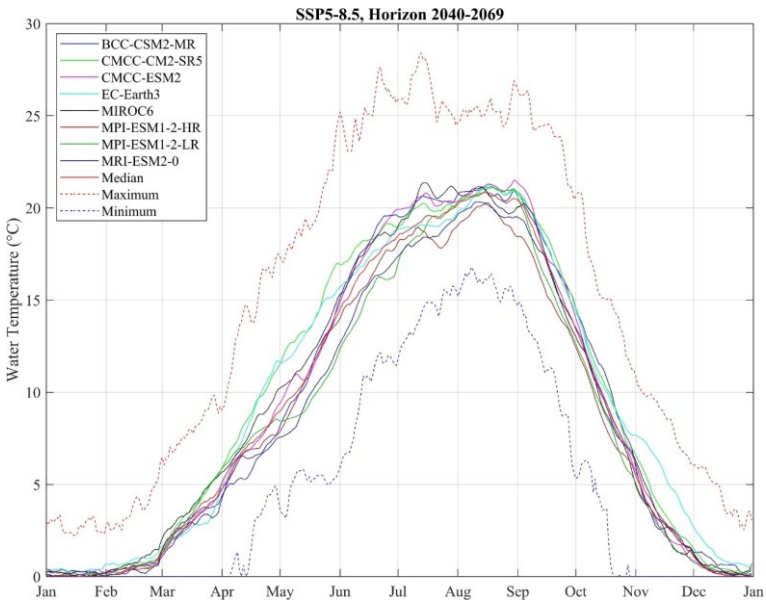

Figure II-B3 Same as Figure II-B1, but for the SSP5-8.5 scenario

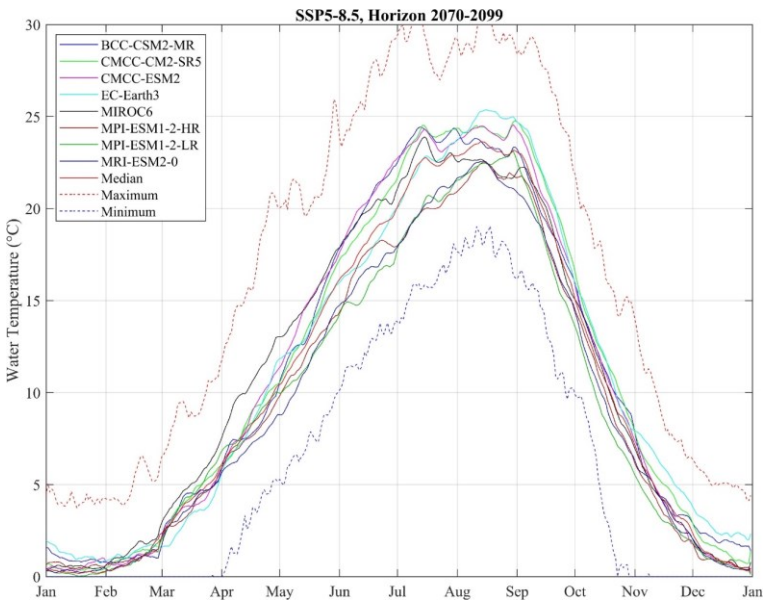

Figure II-B4 Same as Figure II-B1, but for the 2070–2099-time horizon and the SSP5-8.5 scenario

Appendix II-C: Water temperature variability figures under different climate scenarios and time horizons

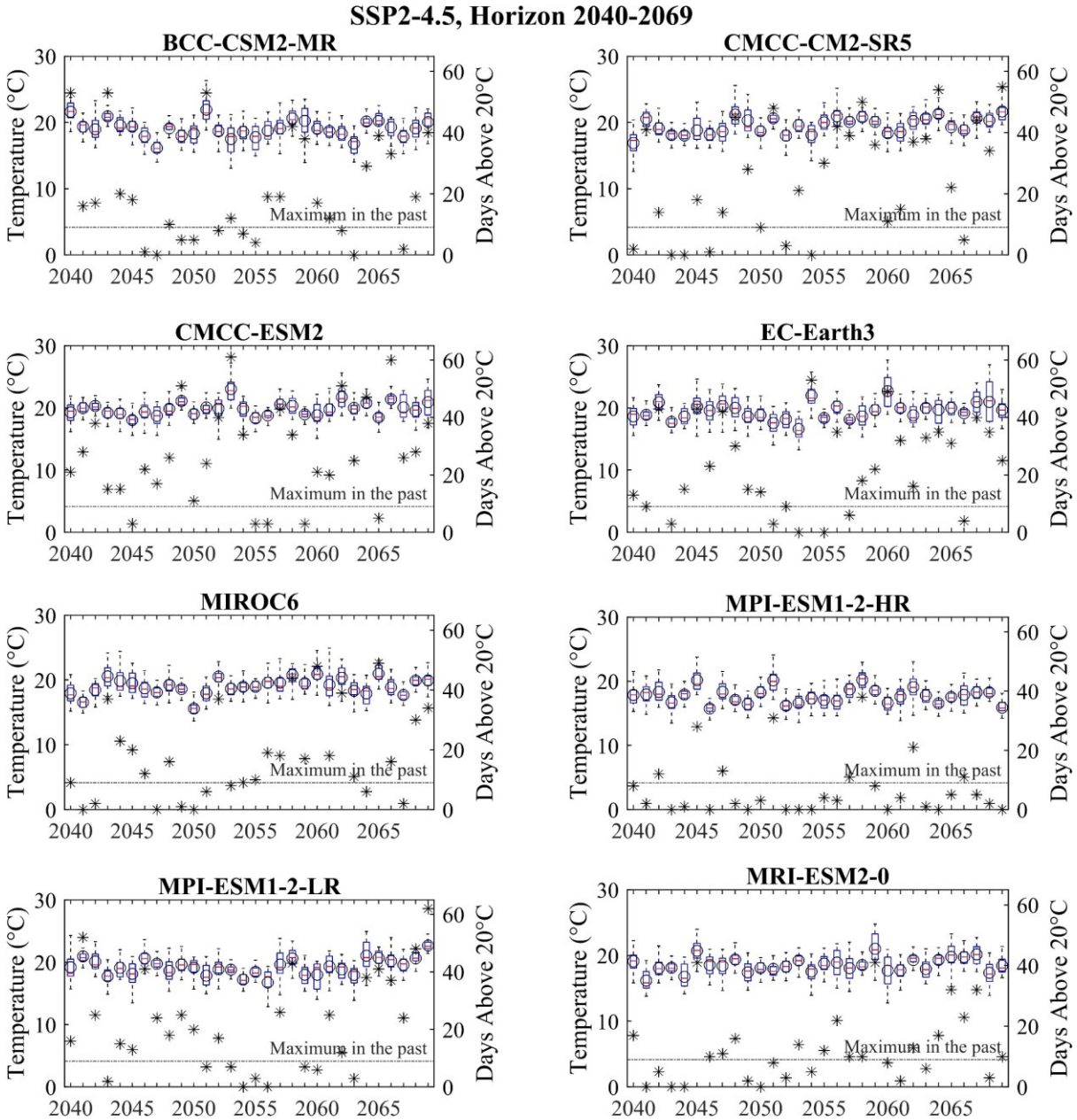

Figure II-C1 Yearly range of water temperature variability using boxplot for horizon 2040-2069 for eight climate models under SSP2-4.5 climate change scenario. The left Y-axis shows temperature (°C). For this axis, the median value is the red line, and the black circle presents the mean value. The right Y axis is the number of days. For this axis, asteriscs show the annual number of days on which temperature exceeds the critical limit of 20°C, also the "Maximum in the past" shows the maximum of days per year with temperature above 20°C. The upper and lower limits of boxes show 25th and 75th percentiles. Whiskers show the most extreme data.

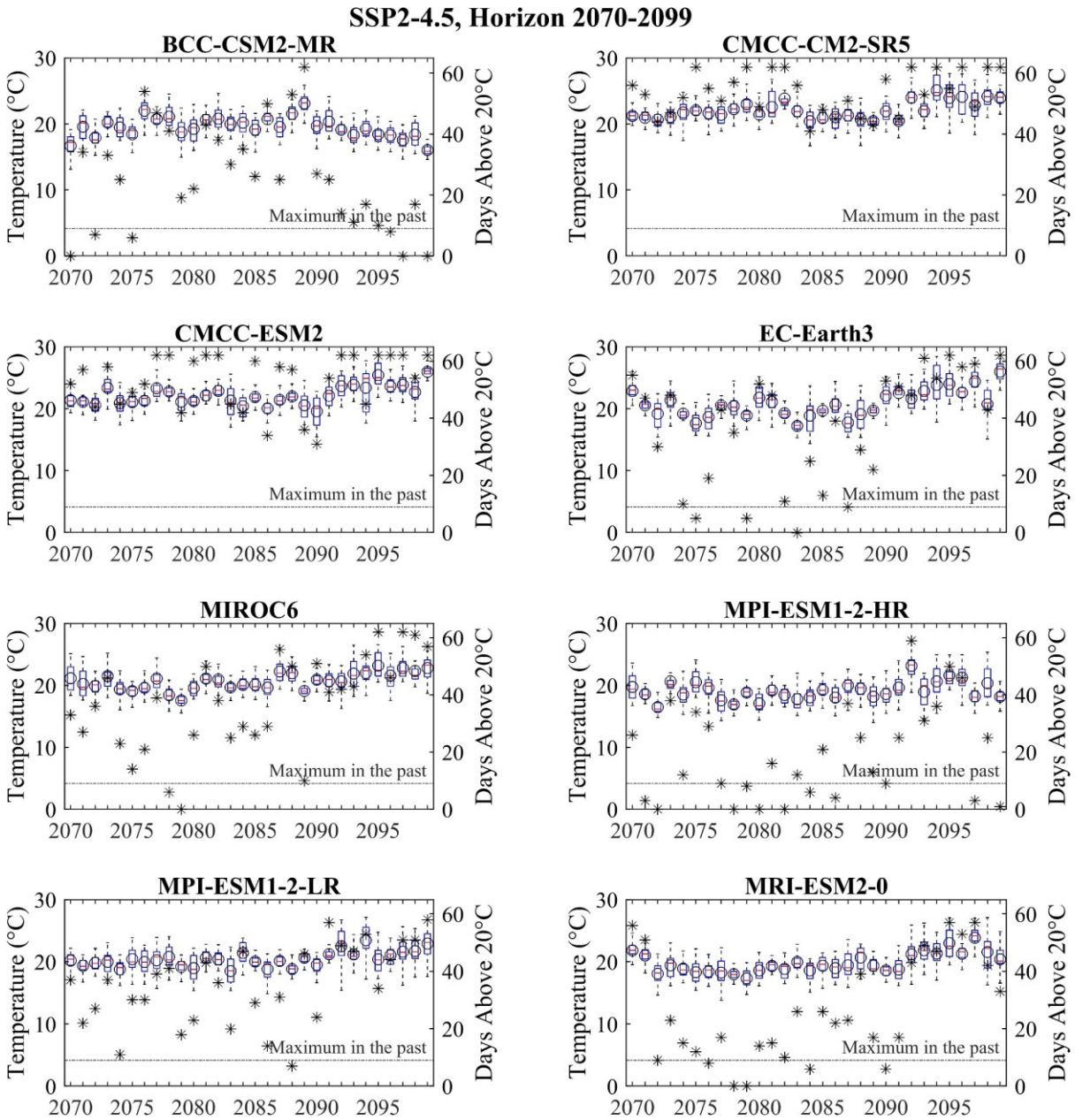

Figure II-C2 Same as Figure II-C1, but for the 2070-2099 time horizon

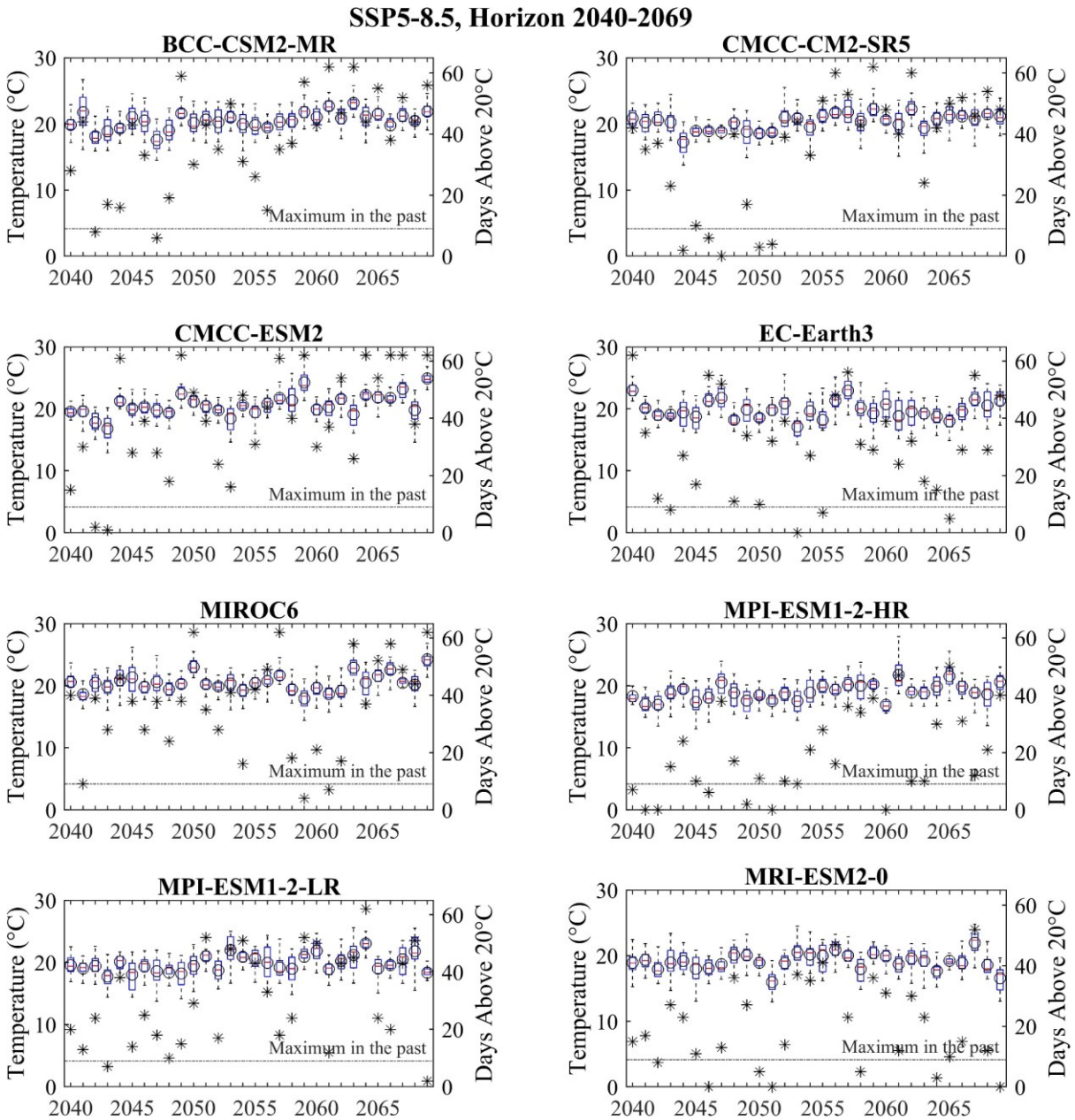

Figure II-C3 Same as Figure II-C1, but for the SSP5-8.5 scenario

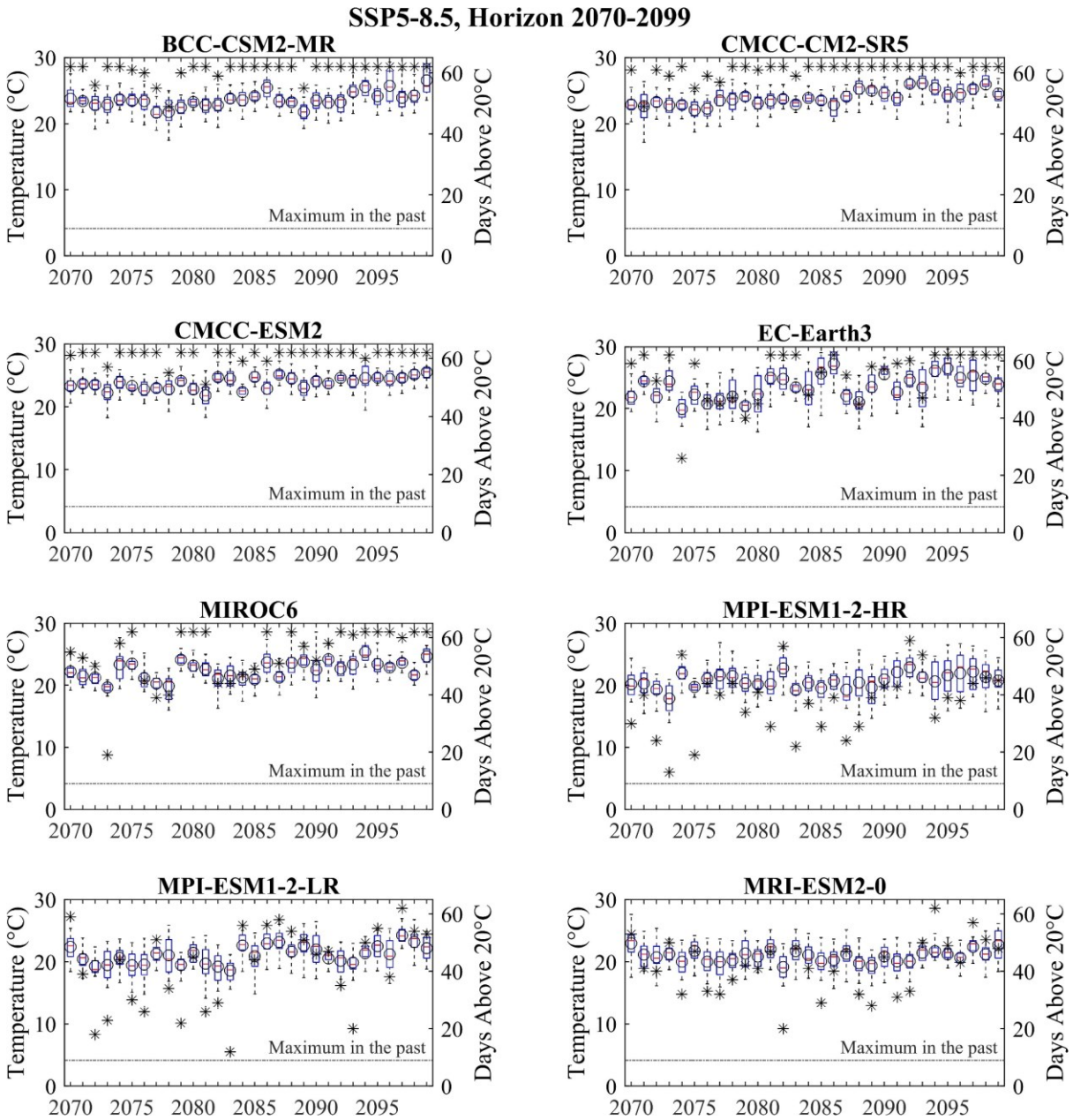

Figure II-C4 Same as Figure II-C1, but for the 2070-2099 time horizon and the SSP5-8.5 scenario

Table II-C1. Calibration/validation performance metrics in previous studies for several watersheds

Study (Authors, Date)	Variable	Performance Metric	Metric Value
Islam et al. (2019)	Water temperature	NSE (Calibration)	0.76 - 0.97
Islam et al. (2019)	Water temperature	Bias (Calibration)	-0.06 – 0.28°C
Islam et al. (2019)	Water temperature	RMSE (Calibration)	0.32 – 1.14°C
Islam et al. (2019)	Water temperature	Bias (Validation)	≤ 2.0°C
Islam et al. (2019)	Water temperature	RMSE (Validation)	1.09 – 2.16°C
Kwak et al. (2017b)	Water flow	NSE (Calibration)	0.82
Kwak et al. (2017b)	Water flow	NSE (Validation)	0.93
Kwak et al. (2017b)	Water temperature	RMSE (Calibration)	0.80°C
Kwak et al. (2017b)	Water temperature	RMSE (Validation)	0.77°C
Kwak et al. (2017b)	Water temperature	Bias (Calibration)	-0.36°C
Kwak et al. (2017b)	Water temperature	Bias (Validation)	-1.14°C
Ouellet-Proulx et al. (2017b)	Water flow	NSE (Calibration)	0.96
Ouellet-Proulx et al. (2017b)	Water flow	NSE (Validation)	0.86
Ouellet-Proulx et al. (2017b)	Water temperature	RMSE (Calibration)	1.38°C
Ouellet-Proulx et al. (2017b)	Water temperature	RMSE (Validation)	1.54°C
Ouellet-Proulx et al. (2017b)	Water temperature	Bias (Calibration)	0.24°C
Ouellet-Proulx et al. (2017b)	Water temperature	Bias (Validation)	0.2°C

Appendix II-D: The MATLAB codes implemented in Article II

ERA5 data preparation for 1980-2019

```
la0 = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','lat');
lo0 = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','lon');
%% 1 Time
d = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','time');
t = double(d)/24 + 722816; %d is : 'hours since 1979-01-01 00:00:00'
n_rows = size(d,1);
n_col = size(lo0,1)*size(la0,1);
%% 2 tMax
tMax = zeros(n_rows, n_col);
mx2t = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','t2m_ERA5');
%% 3 tMin
tMin = zeros(n_rows, n_col);
mn2t = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','t2m_ERA5');
%% 4 Total Precipitation
pTot = zeros(n_rows, n_col);
tp = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','tp_ERA5');
%% 5 rayonnement == 'Mean surface net short-wave radiation flux'
rayonnement = zeros(n_rows, n_col);
msdwswrf = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','msnswrf_ERA5');
%% 6 wind
vitesseVent = zeros(n_rows, n_col);
vitesseVent_u = zeros(n_rows, n_col);
vitesseVent_v = zeros(n_rows, n_col);
u10 = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','u10_ERA5');
v10 = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','v10_ERA5');
%% 7 nebulosite == 'Total cloud cover'
nebulosite = zeros(n_rows, n_col);
tcc = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','tcc_ERA5');
%% 8 Pressure (water vapour pressure)
temp_2m = zeros(n_rows, n_col);
t2m = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','t2m_ERA5');
dew_temp_2m = zeros(n_rows, n_col);
d2m = ncread('ds_Nechako_ERA5_3H_with_scaled_precip.nc','d2m_ERA5');
pression = zeros(n_rows, n_col);
for j = 1:size(la0,1)
    for i = 1:size(lo0,1)
        col = (j-1)*size(lo0,1) + i;
        locations (1, col) = lo0(i);
        locations (2, col) = la0(j);
        for k = 1:n_rows
            tMin(k, col)              = mn2t(i, j, k);
            tMax(k, col)              = mx2t(i, j, k);
            pTot(k, col)            = tp(i, j, k);
            rayonnement(k, col)     = msdwswrf(i, j, k);
            vitesseVent_u(k, col)   = u10(i, j, k);
            vitesseVent_v(k, col)   = v10(i, j, k);
            nebulosite(k, col)      = tcc(i, j, k);
            temp_2m(k, col)         = t2m(i, j, k);
            dew_temp_2m(k, col)     = d2m(i, j, k);
        end
    end
end
vitesseVent = (vitesseVent_u.^2 + vitesseVent_v.^2).^0.5;
%ref: Upreti H, Ojha CSP (2017). % ea (T) = es (Tdew)
for k = 1:n_rows
    for j = 1:n_col
        if dew_temp_2m(k, j) >= 0
            pression(k, j) = 0.61078*exp(17.27*dew_temp_2m(k, j)/(dew_temp_2m(k, j)+237.3));
        else
            pression(k, j) = 0.61078*exp(21.875*dew_temp_2m(k, j)/(dew_temp_2m(k, j)+265.5));
        end
    end
end
%%
ERA5Data = struct('lon_ERA5',  locations (1,:),'lat_ERA5', locations (2,:),...
    'Date', t, 'tMax', tMax,'tMin', ...
    tMin,'pTot', pTot, 'rayonnement', rayonnement,...
    'vitesseVent', vitesseVent, 'nebulosite', nebulosite,...
    'pression', pression);
```

Preparing CEQUEAU structure for 1970-2019

```
clear
clc
tic
%% Database 1
dataFolder  = 'ERA5';
prepare     = 'preparing_Data.m';
%% Static Names
commonDataFolder    = 'CommonFiles';
shapePolygon        = 'Buffered_Polygon_WGS84.shp';
studyArea           = 'NechakoPolygon.shp';
fishnetFile         = 'OriginalFishnetforOldNechako.shp';
demFile             = 'dem.tif';
CEQUEAU_structure   = 'InputStruct.mat';
%% Import Point Vector Data (= All the possible stations)
%% Import Polygon Vector Data (= the boundary I'm intersted to the points within)
polygon = shaperead(strcat(pwd,'\', commonDataFolder, '\', shapePolygon));
studyArea_plygon = shaperead(strcat(pwd,'\', commonDataFolder, '\', studyArea));
%% A check for cordinates systems
info = shapeinfo(strcat(pwd,'\', commonDataFolder, '\', shapePolygon));
run(strcat(pwd,'\', dataFolder, '\', prepare));
%% cor
x_points = ERA5Data.lon_ERA5;
y_points = ERA5Data.lat_ERA5;

x_polygon = polygon.X;
y_polygon = polygon.Y;
% this is the list of stations inside our polygon
station_selecttion = inpolygon(x_points, y_points, x_polygon, y_polygon);
stations_selected_lon = x_points(:, ismember(station_selecttion(1,:), [1]));
stations_selected_lat = y_points(:, ismember(station_selecttion(1,:), [1]));
%% Plotting the boundary and selected points
figure0 = figure;
mapshow(polygon);
hold on
mapshow(studyArea_plygon);
mapshow(stations_selected_lon, stations_selected_lat, 'DisplayType','point',
'Color',[1 0 1])
saveas(figure0,strcat(pwd,'\', dataFolder, '\', dataFolder, '_station_points.fig'));
close(figure0)
%% Preparing 5 columns for Each grid point, required for stations:
% name, i, j, tp, elevation,
ij_tp_elev = zeros(4, size(x_points, 2));
% Preparing i, j
fishnet = shaperead(strcat(pwd,'\', commonDataFolder, '\', fishnetFile));
info = shapeinfo(strcat(pwd,'\', commonDataFolder, '\', fishnetFile));
% info.CoordinateReferenceSystem
% mapshow(fishnet);
for p = 1:size(x_points, 2)
    for row = 1:size(fishnet, 1)
        if inpolygon(x_points(1, p), y_points(1, p), fishnet(row).X, fishnet(row).Y)
            ij_tp_elev(1, p) = fishnet(row).i;
            ij_tp_elev(2, p) = fishnet(row).j;
        end
    end
end
% Preparing tp (mm)
pTot = ERA5Data.pTot;
date1 = datetime(datevec(ERA5Data.Date));
TT_pTot = timetable(date1,pTot);

Ttp = retime(TT_pTot,'yearly','sum');
tp = Ttp{:,:};
tp = mean(tp);
ij_tp_elev(3, :) = tp;
% Preparing elevation (m.a.s.l)
[dem, R] = readgeoraster(strcat(pwd,'\', commonDataFolder, '\', demFile),
'OutputType', 'double');
info = georasterinfo(strcat(pwd,'\', commonDataFolder, '\', demFile));
info.RasterReference.CoordinateSystemType
elevations = geointerp(dem, R, y_points, x_points);
ij_tp_elev(4, :) = elevations;
%% Create Structure for selected stations
stations = struct('id', [],'nom', [],'i', [],'j', [], ...
    'tp', [], 'altitude', []);
```

```
73- stations_selected = ij_tp_elev(:, ismember(station_selecttion(1,:), [1]));
74- for z = 1:size(stations_selected, 2)
75-         stations (z).id = z;
76-         stations (z).nom = string(stations_selected(1, z))+",
    "+string(stations_selected(2, z));
77-         stations (z).i = stations_selected(1, z);
78-         stations (z).j = stations_selected(2, z);
79-         stations (z).tp = stations_selected(3, z);
80-         stations (z).altitude = stations_selected(4, z);
81- end
82- %% Create structure for meteorological data for ECCC
83- t           = ERA5Data.Date;
84- tMax        = ERA5Data.tMax(:, ismember(station_selecttion(1,:), [1]));
85- tMin        = ERA5Data.tMin(:, ismember(station_selecttion(1,:), [1]));
86- pTot        = ERA5Data.pTot(:, ismember(station_selecttion(1,:), [1]));
87- rayonnement = ERA5Data.rayonnement(:, ismember(station_selecttion(1,:), [1]));
88- rayonnement = rayonnement*3*3600/10^6;
89- vitesseVent = ERA5Data.vitesseVent(:, ismember(station_selecttion(1,:), [1]));
90- vitesseVent = 3.6*vitesseVent; %convert from m/s to km/hr
91- nebulosite  = ERA5Data.nebulosite(:, ismember(station_selecttion(1,:), [1]));
92- pression    = ERA5Data.pression(:, ismember(station_selecttion(1,:), [1]));
93- pression    = 7.50062*pression;%; %convert from kPa to mmHg
94- %% We need to add leap days
95- % this will be done using retime function and we need to do it for each var
96- % change to daily time step
97- t_sim_start = datenum(1980, 1, 1);%
98- t_sim_end   = datenum(2019, 12, 31);%
99-
100-        t_cequeau = (datenum(t_sim_start:hours(3):t_sim_end))';
101-        date1 = datetime(datevec(ERA5Data.Date));
102-        date2 = datetime(datevec(t_cequeau));
103-
104-        t = t_cequeau;
105-        t_daily = (datenum(t_sim_start:hours(24):t_sim_end))';
106-        NNN = size(t_cequeau, 1);
107-        NNN_daily = size(t_daily, 1);
108-        %tMax
109-        TT_era5_tMax = timetable(date1, tMax);
110-        TT2 = retime(TT_era5_tMax, date2, 'nearest');
111-        TT2_daily = retime(TT2,'daily','max');
112-        tMax = TT2{:,:};
113-        tMax_daily = TT2_daily{:,:};
114-        %tMin
115-        TT_era5_tMin = timetable(date1, tMin);
116-        TT2 = retime(TT_era5_tMin, date2, 'nearest');
117-        TT2_daily = retime(TT2,'daily','min');
118-        tMin = TT2{:,:};
119-        tMin_daily = TT2_daily{:,:};
120-        %pTot
121-        TT_era5_pTot = timetable(date1, pTot);
122-        TT2 = retime(TT_era5_pTot, date2, 'nearest');
123-        TT2_daily = retime(TT2,'daily','sum');
124-        pTot = TT2{:,:};
125-        pTot_daily = TT2_daily{:,:};
126-        %rayonnement
127-        TT_era5_rayonnement = timetable(date1, rayonnement);
128-        TT2 = retime(TT_era5_rayonnement, date2, 'nearest');
129-        TT2_daily = retime(TT2,'daily','sum');
130-        rayonnement = TT2{:,:};
131-        rayonnement_daily = TT2_daily{:,:};
132-        %vitesseVent
133-        TT_era5_vitesseVent = timetable(date1, vitesseVent);
134-        TT2 = retime(TT_era5_vitesseVent, date2, 'nearest');
135-        TT2_daily = retime(TT2,'daily','mean');
136-        vitesseVent = TT2{:,:};
137-        vitesseVent_daily = TT2_daily{:,:};
138-        %nebulosite
139-        TT_era5_nebulosite = timetable(date1, nebulosite);
140-        TT2 = retime(TT_era5_nebulosite, date2, 'nearest');
141-        TT2_daily = retime(TT2,'daily','mean');
142-        nebulosite = TT2{:,:};
143-        nebulosite_daily = TT2_daily{:,:};
144-        %pression
145-        TT_era5_pression = timetable(date1, pression);
146-        TT2 = retime(TT_era5_pression, date2, 'nearest');
147-        TT2_daily = retime(TT2,'daily','mean');
148-        pression = TT2{:,:};
```

```
149-        pression_daily = TT2_daily{:,:};
150-        %% for daily data
151-        meteoStation = struct('t', t_daily,'tMax', tMax_daily,'tMin', ...
152-            tMin_daily,'pTot', pTot_daily, 'rayonnement', rayonnement_daily,...
153-            'vitesseVent', vitesseVent_daily, 'nebulosite', nebulosite_daily,...
154-            'pression', pression_daily);
155-        %% Changing boundary conditions
156-        load(strcat(pwd,'\', commonDataFolder, '\', 'realFlow_SLS_1980_2019.mat'));
157-        % the name of this variable is "realFlow" with two columns "Date + Value"
158-
159-        % increasing the flow at SLS 25% or 50%
160-        ddate = datevec(realFlow(:,1));
161-        summerFlow = realFlow(and(ddate(:,2) >=6, ddate(:,2) <=9), :);
162-        summerFlow(:,2) = summerFlow(:,2).* 1;
163-        % summerFlow(:,2) = summerFlow(:,2).* 1.25;
164-        % summerFlow(:,2) = summerFlow(:,2).* 1.5;
165-
166-        realFlow(and(ddate(:,2) >=6, ddate(:,2) <=9), :) = summerFlow;
167-
168-        id = [1:1233];
169-        temperature = NaN*ones(1,1233);
170-        temperature (1,1109) = 0;
171-        etatsCP   = struct('id', id,'temperature', temperature);
172-
173-        load(strcat(pwd,'\', commonDataFolder, '\', CEQUEAU_structure))
174-        load(strcat(pwd,'\', commonDataFolder, '\',
    'SLS_Synthetic_temps1980_2019.mat'));
175-        %% +5 and -5C to see the impact
176-        % SLS_Synthetic_temps1981_2019 = SLS_Synthetic_temps1981_2019 + 5;
177-        % SLS_Synthetic_temps1981_2019 = SLS_Synthetic_temps1981_2019 - 5;
178-        InputStruct.bassinVersant.barrage(1).debit  = realFlow(:,2);
179-        InputStruct.assimilations.qualite = [];
180-        for zx = 1:NNN_daily
181-            InputStruct.assimilations.qualite(zx).pasDeTemps = datenum(1980,1,1,0,0,0)
    + zx -1;
182-            InputStruct.assimilations.qualite(zx).etatsCP = etatsCP;
183-            % below line is changing water temprature at SLS.
184-            % min value is column 2 and max value for each day is column 3
185-            InputStruct.assimilations.qualite(zx).etatsCP.temperature(1, 1109) =
    SLS_Synthetic_temps1980_2019(zx,1);
186-        end
187-        %% Placing new created data in the current structure
188-        InputStruct.stations = stations;
189-        InputStruct.meteoStation = meteoStation;
190-        % %% Excecution dates for interpolation
191-        % % change to 3 hours time step
192-        InputStruct.parametres.option.ipassim = 24;
193-        InputStruct.execution.dateDebut  =  t_sim_start;
194-        InputStruct.execution.dateFin    =  t_sim_end;
195-        %% Producing interpolated data for all variables
196-        InputStruct.meteoPointGrille = [];
197-        meteoInterpolee = cequeauInterpolationMex(InputStruct.execution,
    InputStruct.parametres, InputStruct.bassinVersant, InputStruct.stations,
    InputStruct.meteoStation);
198-
199-        InputStruct.meteoPointGrille = meteoInterpolee;
200-        InputStruct.meteoStation   = rmfield(InputStruct.meteoStation  , 'tMax');
201-        InputStruct.meteoStation   = rmfield(InputStruct.meteoStation  , 'tMin');
202-
203-        InputStruct.meteoStation = renameStructField(InputStruct.meteoStation,
    'rayonnement', 'tMax');
204-        InputStruct.meteoStation = renameStructField(InputStruct.meteoStation,
    'vitesseVent', 'tMin');
205-        meteoInterpolee = cequeauInterpolationMex(InputStruct.execution,
    InputStruct.parametres, InputStruct.bassinVersant, InputStruct.stations,
    InputStruct.meteoStation);
206-        meteoInterpolee = renameStructField(meteoInterpolee, 'tMax', 'rayonnement');
207-        meteoInterpolee = renameStructField(meteoInterpolee, 'tMin', 'vitesseVent');
208-
209-        InputStruct.meteoPointGrille.rayonnement = meteoInterpolee.rayonnement;
210-        InputStruct.meteoPointGrille.vitesseVent = meteoInterpolee.vitesseVent;
211-
212-        InputStruct.meteoStation = renameStructField(InputStruct.meteoStation,
    'nebulosite', 'tMax');
213-        InputStruct.meteoStation = renameStructField(InputStruct.meteoStation,
    'pression', 'tMin');
```

```
214-        meteoInterpolee = cequeauInterpolationMex(InputStruct.execution,
    InputStruct.parametres, InputStruct.bassinVersant, InputStruct.stations,
    InputStruct.meteoStation);
215-        meteoInterpolee = renameStructField(meteoInterpolee, 'tMax', 'nebulosite');
216-        meteoInterpolee = renameStructField(meteoInterpolee, 'tMin', 'pression');
217-
218-        InputStruct.meteoPointGrille.nebulosite = meteoInterpolee.nebulosite;
219-        InputStruct.meteoPointGrille.pression = meteoInterpolee.pression;
220-        % to keep all data in meteoStation field
221-        InputStruct.meteoStation = meteoStation;
222-        %% Add spin up period
223-        % this period is a repetition of first year for 3 times
224-        firstDate = datevec(InputStruct.meteoStation.t(1,1));
225-        firstYear = firstDate(1);
226-        % the number 10 was chosen after trial and error
227-        DELTA = 10;
228-        spinupStart = datenum(firstYear-DELTA,1,1);
229-        spinupEnd = datenum(firstYear-1,12,31);
230-        spinupTime = (spinupStart:spinupEnd)';
231-        spinLength = size(spinupTime, 1);
232-        % spin up structure building
233-        NewStruct = InputStruct;
234-        NewStruct.meteoStation.t = [spinupTime;InputStruct.meteoStation.t];
235-        NewStruct.meteoPointGrille.tMin =
    [InputStruct.meteoPointGrille.tMin(1:spinLength, :);
    InputStruct.meteoPointGrille.tMin];
236-        NewStruct.meteoPointGrille.tMax =
    [InputStruct.meteoPointGrille.tMax(1:spinLength, :);
    InputStruct.meteoPointGrille.tMax];
237-        NewStruct.meteoPointGrille.pTot =
    [InputStruct.meteoPointGrille.pTot(1:spinLength, :);
    InputStruct.meteoPointGrille.pTot];
238-        NewStruct.meteoPointGrille.rayonnement =
    [InputStruct.meteoPointGrille.rayonnement(1:spinLength, :);
    InputStruct.meteoPointGrille.rayonnement];
239-        NewStruct.meteoPointGrille.vitesseVent =
    [InputStruct.meteoPointGrille.vitesseVent(1:spinLength, :);
    InputStruct.meteoPointGrille.vitesseVent];
240-        NewStruct.meteoPointGrille.nebulosite =
    [InputStruct.meteoPointGrille.nebulosite(1:spinLength, :);
    InputStruct.meteoPointGrille.nebulosite];
241-        NewStruct.meteoPointGrille.pression =
    [InputStruct.meteoPointGrille.pression(1:spinLength, :);
    InputStruct.meteoPointGrille.pression];
242-        % boundary condition (water flow)
243-        NewStruct.bassinVersant.barrage(1).debit  =
    [InputStruct.bassinVersant.barrage(1).debit(1:spinLength, :);
    InputStruct.bassinVersant.barrage(1).debit];
244-
245-        % new simulation start time
246-        NewStruct.execution.dateDebut = NewStruct.meteoStation.t (1,1);
247-        InputStruct = NewStruct;
248-        %% Simulation
249-        % best set of parameters for Nautley (all flows)
250-        %% Hydrological Optimization Parameters
251-        load(strcat(pwd,'\', commonDataFolder, '\', 'x_hydro.mat'));
252-        load(strcat(pwd,'\', commonDataFolder, '\', 'x_thermal.mat'));
253-        %% Setting the initials
254-        InputStruct.parametres.solInitial.hsini = x_hydro(9);
255-        InputStruct.parametres.solInitial.hnini = x_hydro(11);
256-        InputStruct.parametres.solInitial.hmini = x_hydro(10);
257-        InputStruct.parametres.solInitial.q0    = 100;
258-        %%
259-        InputStruct.parametres.sol.cin_s = x_hydro(1);
260-        InputStruct.parametres.sol.cvmar = x_hydro(2);
261-        InputStruct.parametres.sol.cvnb_s = x_hydro(3);
262-        InputStruct.parametres.sol.cvnh_s = x_hydro(4);
263-        InputStruct.parametres.sol.cvsb = x_hydro(5);
264-        InputStruct.parametres.sol.cvsi_s = x_hydro(6);
265-        InputStruct.parametres.sol.xinfma = x_hydro(7);
266-        InputStruct.parametres.sol.hinf_s = x_hydro(8);
267-        InputStruct.parametres.sol.hint_s = x_hydro(9);
268-        InputStruct.parametres.sol.hmar = x_hydro(10);
269-        InputStruct.parametres.sol.hnap_s = x_hydro(11);
270-        InputStruct.parametres.sol.hpot_s = x_hydro(12);
271-        InputStruct.parametres.sol.hsol_s = x_hydro(13);
272-        InputStruct.parametres.sol.hrimp_s = x_hydro(14);
```

```
273-        InputStruct.parametres.transfert.exxkt = x_hydro(15);
274-        InputStruct.parametres.neige.strne_s = x_hydro(16);
275-        InputStruct.parametres.fonte.cequeau.strne_s = x_Hydro(16);
276-        InputStruct.parametres.neige.tfc_s = x_hydro(17);
277-        InputStruct.parametres.fonte.cequeau.tfc_s = x_Hydro(17);
278-        InputStruct.parametres.neige.tfd_s = x_hydro(18);
279-        InputStruct.parametres.fonte.cequeau.tfd_s = x_Hydro(18);
280-        InputStruct.parametres.neige.tsc_s = x_hydro(19);
281-        InputStruct.parametres.fonte.cequeau.tsc_s = x_Hydro(19);
282-        InputStruct.parametres.neige.tsd_s = x_hydro(20);
283-        InputStruct.parametres.fonte.cequeau.tsd_s = x_Hydro(20);
284-        InputStruct.parametres.neige.ttd = x_hydro(21);
285-        InputStruct.parametres.fonte.cequeau.ttd = x_Hydro(21);
286-        InputStruct.parametres.neige.tts_s = x_hydro(22);
287-        InputStruct.parametres.fonte.cequeau.tts_s = x_Hydro(22);
288-        InputStruct.parametres.evapo.cequeau.evnap = x_hydro(23);
289-        InputStruct.parametres.evapo.cequeau.xaa = x_hydro(24);
290-        InputStruct.parametres.evapo.cequeau.xit = x_hydro(25);
291-        InputStruct.parametres.transfert.zn = x_hydro(26);
292-        %% Thermal Optimization Parameters
293-        InputStruct.parametres.qualite.cequeau.temperat.crayso = x_thermal(1);
294-        InputStruct.parametres.qualite.cequeau.temperat.crayin = x_thermal(2);
295-        InputStruct.parametres.qualite.cequeau.temperat.cevapo = x_thermal(3);
296-        InputStruct.parametres.qualite.cequeau.temperat.cconve = x_thermal(4);
297-        InputStruct.parametres.qualite.cequeau.temperat.crigel = x_thermal(5);
298-        InputStruct.parametres.qualite.cequeau.temperat.tnap = x_thermal(6);
299-        InputStruct.parametres.qualite.cequeau.temperat.panap = 0;
300-        InputStruct.parametres.qualite.cequeau.temperat.tinit = 0;
301-        InputStruct.parametres.qualite.cequeau.temperat.bassol = x_thermal(7);
302-        InputStruct.parametres.qualite.cequeau.temperat.corsol = x_thermal(8);
303-        %%
304-        [y.etatsCE, y.etatsCP, y.etatsFonte, y.etatsEvapo, y.etatsBarrage,
    y.pasDeTemps,...
305-            y.avantAssimilationsCE, y.avantAssimilationsFonte,
    y.avantAssimilationsEvapo,y.etatsQualCP, y.avAssimQual] = ...
306-        cequeauQuantiteMex_v461(InputStruct.execution, InputStruct.parametres,
    InputStruct.bassinVersant, ...
307-             InputStruct.meteoPointGrille, [], InputStruct.assimilations);
```

Model calibration

```
1-  clear all
2-  clc
3-  %% Declare and Load inputs
4-  global InputStruct
5-  load 'InputStruct.mat'
6-  %% Hydrological parameters from previous optimization
7-  load('HydrologicalCalibrationResultMultisite.mat')
8-  %% Hydrological Optimization Parameters
9-  InputStruct.parametres.sol.cin_s = x_Hydro(1);
10- InputStruct.parametres.sol.cvmar = x_Hydro(2);
11- InputStruct.parametres.sol.cvnb_s = x_Hydro(3);
12- InputStruct.parametres.sol.cvnh_s = x_Hydro(4);
13- InputStruct.parametres.sol.cvsb = x_Hydro(5);
14- InputStruct.parametres.sol.cvsi_s = x_Hydro(6);
15- InputStruct.parametres.sol.xinfma = x_Hydro(7);
16- InputStruct.parametres.sol.hinf_s = x_Hydro(8);
17- InputStruct.parametres.sol.hint_s = x_Hydro(9);
18- InputStruct.parametres.sol.hmar = x_Hydro(10);
19- InputStruct.parametres.sol.hnap_s = x_Hydro(11);
20- InputStruct.parametres.sol.hpot_s = x_Hydro(12);
21- InputStruct.parametres.sol.hsol_s = x_Hydro(13);
22- InputStruct.parametres.sol.hrimp_s = x_Hydro(14);
23- InputStruct.parametres.transfert.exxkt = x_Hydro(15);
24- InputStruct.parametres.neige.strne_s = x_Hydro(16);
25- InputStruct.parametres.fonte.cequeau.strne_s = x_Hydro(16);
26- InputStruct.parametres.neige.tfc_s = x_Hydro(17);
27- InputStruct.parametres.fonte.cequeau.tfc_s = x_Hydro(17);
28- InputStruct.parametres.neige.tfd_s = x_Hydro(18);
29- InputStruct.parametres.fonte.cequeau.tfd_s = x_Hydro(18);
30- InputStruct.parametres.neige.tsc_s = x_Hydro(19);
31- InputStruct.parametres.fonte.cequeau.tsc_s = x_Hydro(19);
32- InputStruct.parametres.neige.tsd_s = x_Hydro(20);
33- InputStruct.parametres.fonte.cequeau.tsd_s = x_Hydro(20);
34- InputStruct.parametres.neige.ttd = x_Hydro(21);
```

```
35- InputStruct.parametres.fonte.cequeau.ttd = x_Hydro(21);
36- InputStruct.parametres.neige.tts_s = x_Hydro(22);
37- InputStruct.parametres.fonte.cequeau.tts_s = x_Hydro(22);
38- InputStruct.parametres.evapo.cequeau.evnap = x_Hydro(23);
39- InputStruct.parametres.evapo.cequeau.xaa = x_Hydro(24);
40- InputStruct.parametres.evapo.cequeau.xit = x_Hydro(25);
41- InputStruct.parametres.transfert.zn = x_Hydro(26);
42- %% Setting the initials
43- InputStruct.parametres.solInitial.hsini = x_Hydro(9);
44- InputStruct.parametres.solInitial.hnini = x_Hydro(11);
45- InputStruct.parametres.solInitial.hmini = x_Hydro(10);
46- InputStruct.parametres.solInitial.q0    = 10;
47-
48- lb = [0.05  0.05    0.05    0.05    0   5   5   0];
49- ub = [2 2   2   2   1000    10  20  1];
50- % ub = [1.95 1.92 0.83 2.17 75.29 5.38 16.94 0.25];
51- x  = (lb+0.5.*(ub-lb))';
52- sigma = (ub-lb)';
53- opts.LBounds = lb';
54- opts.UBounds = ub';
55- opts.MaxFunEvals  = 5000;
56- [xmin, fmin, counteval, stopflag, out, bestever]= cmaes('fun', x, sigma, opts);
57-
```

Objective function

```
1-  function F = fun(x)
2-  global InputStruct
3-  InputStruct.parametres.qualite.cequeau.temperat.crayso = x(1);
4-  InputStruct.parametres.qualite.cequeau.temperat.crayin = x(2);
5-  InputStruct.parametres.qualite.cequeau.temperat.cevapo = x(3);
6-  InputStruct.parametres.qualite.cequeau.temperat.cconve = x(4);
7-  InputStruct.parametres.qualite.cequeau.temperat.crigel = x(5);
8-  InputStruct.parametres.qualite.cequeau.temperat.tnap = x(6);
9-  InputStruct.parametres.qualite.cequeau.temperat.panap = 0;
10- InputStruct.parametres.qualite.cequeau.temperat.tinit = 0;
11- InputStruct.parametres.qualite.cequeau.temperat.bassol = x(7);
12- InputStruct.parametres.qualite.cequeau.temperat.corsol = x(8);
13- % Use of the thermal module:
14- thermie = 1;
15- % Dates for which we want to simulate, calibrate, validate and illustrate the results:
16- startSim = datenum(1970, 01, 01);
17- endSim = datenum(2019, 12, 31);
18- dates.sim = [startSim:endSim]';
19- % Settings to adjust:
20- %valeursParametres_thermie;
21- InputStruct.execution.dateDebut = dates.sim(1);
22- InputStruct.execution.dateFin = dates.sim(end);
23- InputStruct.parametres.option.calculQualite = thermie;
24- %Simulation
25- [y.etatsCE, y.etatsCP, y.etatsFonte, y.etatsEvapo, y.etatsBarrage, y.pasDeTemps,...
26-     y.avantAssimilationsCE, y.avantAssimilationsFonte,
    y.avantAssimilationsEvapo,y.etatsQualCP, y.avAssimQual] = ...
27- cequeauQuantiteMex_v461(InputStruct.execution, InputStruct.parametres,
    InputStruct.bassinVersant, ...
28-      InputStruct.meteoPointGrille, [], InputStruct.assimilations);
29- %%
30- timesteps = dates.sim(end)-dates.sim(1) + 2;
31- %% CP
32- Tsim1 = zeros(timesteps-1, 1);
33- Tsim23 = zeros(timesteps-1, 1);
34- Tsim43 = zeros(timesteps-1, 1);
35- Tsim180 = zeros(timesteps-1, 1);
36- Tsim181 = zeros(timesteps-1, 1);
37- Tsim186 = zeros(timesteps-1, 1);
38- Tsim352 = zeros(timesteps-1, 1);
39- Tsim491 = zeros(timesteps-1, 1);
40-
41- for t = 2:timesteps
42-     Tsim1(t-1,1) = y.etatsQualCP(t).temperature(1);
43-     Tsim23(t-1,1) = y.etatsQualCP(t).temperature(23);
44-     Tsim43(t-1,1) = y.etatsQualCP(t).temperature(43);
45-     Tsim180(t-1,1) = y.etatsQualCP(t).temperature(180);
46-     Tsim181(t-1,1) = y.etatsQualCP(t).temperature(181);
47-     Tsim186(t-1,1) = y.etatsQualCP(t).temperature(186);
```

```
    Tsim352(t-1,1) = y.etatsQualCP(t).temperature(352);
    Tsim491(t-1,1) = y.etatsQualCP(t).temperature(491);
end
spinupStart = datenum(1970,1,1);
spinupEnd = datenum(1979,12,31);
spinupTime = (spinupStart:spinupEnd)';
spinLength = size(spinupTime, 1);

Tsim1_e = Tsim1(spinLength+1:end,1);
Tsim23_e = Tsim23(spinLength+1:end,1);
Tsim43_e = Tsim43(spinLength+1:end,1);
Tsim180_e = Tsim180(spinLength+1:end,1);
Tsim181_e = Tsim181(spinLength+1:end,1);
Tsim186_e = Tsim186(spinLength+1:end,1);
Tsim352_e = Tsim352(spinLength+1:end,1);
Tsim491_e = Tsim491(spinLength+1:end,1);

Tsim = cat(1,Tsim1_e,Tsim23_e,Tsim43_e,Tsim180_e,...
    Tsim181_e,Tsim186_e,Tsim352_e,Tsim491_e);
Tobs1 = InputStruct.temp_CP1(:,1);
Tobs23 = InputStruct.temp_CP23(:,1);
Tobs43 = InputStruct.temp_CP43(:,1);
Tobs180 = InputStruct.temp_CP180(:,1);
Tobs181 = InputStruct.temp_CP181(:,1);
Tobs186 = InputStruct.temp_CP186(:,1);
Tobs352 = InputStruct.temp_CP352(:,1);
Tobs491 = InputStruct.temp_CP491(:,1);

Tobs = cat(1,Tobs1,Tobs23,Tobs43,Tobs180,Tobs181,Tobs186,Tobs352,Tobs491);
T1 = InputStruct.SpillwayFlow1980_2019(:,1);
T23 = T1;
T43 = T1;
T180 = T1;
T181 = T1;
T186 = T1;
T352 = T1;
T491 = T1;

T = cat(1,T1,T23,T43,T180,T181,T186,T352,T491);

temps(:,1) = T;
temps(:,2) = Tobs;
temps(:,3) = Tsim;
cal_temps = temps;
%% extracting specific flow months
mah = datevec(cal_temps(:,1));
specific_temps = cal_temps(ismember(mah(:,2),[6:9]),:);
%% Bias for vanderhoof
t_vanderhoof(:,1) = T23;
t_vanderhoof(:,2) = Tobs23;
t_vanderhoof(:,3) = Tsim23_e;
cal_t_vanderhoof = t_vanderhoof;
        mah_vanderhoof = datevec(cal_t_vanderhoof(:,1));
        specific_temps_vanderhoof =
    cal_t_vanderhoof(ismember(mah_vanderhoof(:,2),[6:9]),:);
        %%
        F  = RMSE(specific_temps(:,2), specific_temps(:,3))
        end
```

Climate change simulations (SSP2-4.5 for 2040-2069)

```
clear
clc
tic
for database = 1:8
    switch database
        case 1
            dataFolder  = 'BCC-CSM2-MR';
            MBCnData    = 'BCC-CSM2-MR_ERA5_MBCn_data.mat';
        case 2
            dataFolder  = 'CMCC-CM2-SR5';
            MBCnData    = 'CMCC-CM2-SR5_ERA5_MBCn_data.mat';
        case 3
            dataFolder  = 'CMCC-ESM2';
            MBCnData    = 'CMCC-ESM2_ERA5_MBCn_data.mat';
```

```
15-         case 4
16-             dataFolder  = 'EC-Earth3';
17-             MBCnData    = 'EC-Earth3_ERA5_MBCn_data.mat';
18-         case 5
19-             dataFolder  = 'MIROC6';
20-             MBCnData    = 'MIROC6_ERA5_MBCn_data.mat';
21-         case 6
22-             dataFolder  = 'MPI-ESM1-2-HR';
23-             MBCnData    = 'MPI-ESM1-2-HR_ERA5_MBCn_data.mat';
24-         case 7
25-             dataFolder  = 'MPI-ESM1-2-LR';
26-             MBCnData    = 'MPI-ESM1-2-LR_ERA5_MBCn_data.mat';
27-         case 8
28-             dataFolder  = 'MRI-ESM2-0';
29-             MBCnData    = 'MRI-ESM2-0_ERA5_MBCn_data.mat';
30-     end
31- %% Static Names
32- commonDataFolder    = 'CommonFiles';
33- shapePolygon        = 'Polygon.shp';
34- studyArea           = 'NechakoPolygon.shp';
35- fishnetFile         = 'OriginalFishnetforOldNechako.shp';
36- demFile             = 'dem.tif';
37- CEQUEAU_structure   = 'InputStruct.mat';
38- %% Import Polygon Vector Data (= the boundary I'm intersted to the points within)
39- polygon = shaperead(strcat(pwd,'\', commonDataFolder, '\', shapePolygon));
40- studyArea_plygon = shaperead(strcat(pwd,'\', commonDataFolder, '\', studyArea));
41- %% A check for cordinates systems
42- info = shapeinfo(strcat(pwd,'\', commonDataFolder, '\', shapePolygon));
43- %% Load past and future data = Provided by Jean-Luc
44- load(strcat(pwd,'\', dataFolder, '\', MBCnData))
45- x_points = watershed_data_raw.lon_CMIP6;
46- y_points = watershed_data_raw.lat_CMIP6;
47-
48- x_polygon = polygon.X;
49- y_polygon = polygon.Y;
50- % this is the list of stations inside our polygon
51- station_selecttion = inpolygon(x_points, y_points, x_polygon, y_polygon);
52- stations_selected_lon = x_points(:, ismember(station_selecttion(1,:), [1]));
53- stations_selected_lat = y_points(:, ismember(station_selecttion(1,:), [1]));
54- %% Plotting the boundary and selected points
55- figure0 = figure;
56- mapshow(polygon);
57- hold on
58- mapshow(studyArea_plygon);
59- % all stations are like below
60- % mapshow(watershed_data_raw.lon_CMIP6, watershed_data_raw.lat_CMIP6,
    'DisplayType','point', 'Color',[1 1 1])
61- % selected points inside the boundary = Stations
62- mapshow(stations_selected_lon, stations_selected_lat, 'DisplayType','point',
    'Color',[1 0 1])
63- saveas(figure0,strcat(pwd,'\', dataFolder, '\', dataFolder, '_station_points.fig'));
64- close(figure0)
65- %% Preparing 5 columns for Each grid point, required for stations:
66- % name, i, j, tp, elevation,
67- ij_tp_elev = zeros(4, size(x_points, 2));
68- % Preparing i, j
69- fishnet = shaperead(strcat(pwd,'\', commonDataFolder, '\', fishnetFile));
70- info = shapeinfo(strcat(pwd,'\', commonDataFolder, '\', fishnetFile));
71- % info.CoordinateReferenceSystem
72- % mapshow(fishnet);
73- for p = 1:size(x_points, 2)
74-     for row = 1:size(fishnet, 1)
75-         if inpolygon(x_points(1, p), y_points(1, p), fishnet(row).X, fishnet(row).Y)
76-             ij_tp_elev(1, p) = fishnet(row).i;
77-             ij_tp_elev(2, p) = fishnet(row).j;
78-         end
79-     end
80- end
81- % Preparing tp (mm)
82- pTot = watershed_data_MBCn.pr_ssp245;
83- date1 = datetime(watershed_data_MBCn.dates_ssp245(:,:));
84- TT_pTot = timetable(date1,pTot);
85-
86- Ttp = retime(TT_pTot,'yearly','sum');
87- tp = Ttp{:,:};
88- tp = mean(tp);
89- ij_tp_elev(3, :) = tp;
```

```
90- % Preparing elevation (m.a.s.l)
91- [dem, R] = readgeoraster(strcat(pwd,'\', commonDataFolder, '\', demFile),
    'OutputType', 'double');
92- info = georasterinfo(strcat(pwd,'\', commonDataFolder, '\', demFile));
93- elevations = geointerp(dem, R, y_points, x_points);
94- ij_tp_elev(4, :) = elevations;
95- %% Create Structure for selected stations
96- stations = struct('id', [],'nom', [],'i', [],'j', [], ...
97-     'tp', [], 'altitude', []);
98- stations_selected = ij_tp_elev(:, ismember(station_selecttion(1,:), [1]));
99- for z = 1:size(stations_selected, 2)
100-                stations (z).id = z;
101-                stations (z).nom = string(stations_selected(1, z))+",
    "+string(stations_selected(2, z));
102-                stations (z).i = stations_selected(1, z);
103-                stations (z).j = stations_selected(2, z);
104-                stations (z).tp = stations_selected(3, z);
105-                stations (z).altitude = stations_selected(4, z);
106-        end
107-        % save('stations', 'stations');
108-        %% Create structure for meteorological data for ECCC
109-        t = datenum(watershed_data_MBCn.dates_ssp245);
110-        tMax        = watershed_data_MBCn.tas_ssp245(:,
    ismember(station_selecttion(1,:), [1]));
111-        tMin        = watershed_data_MBCn.tas_ssp245(:,
    ismember(station_selecttion(1,:), [1]));
112-        pTot        = watershed_data_MBCn.pr_ssp245(:,
    ismember(station_selecttion(1,:), [1]));
113-        rayonnement = watershed_data_MBCn.msnswrf_ssp245(:,
    ismember(station_selecttion(1,:), [1]));
114-        rayonnement = rayonnement*3*3600/10^6; %convert from W/m2 to MJ/m2 for 3hours
115-        vitesseVent = watershed_data_MBCn.sfcWind_ssp245(:,
    ismember(station_selecttion(1,:), [1]));
116-        vitesseVent = 3.6*vitesseVent; %convert from m/s to Km/hr
117-        nebulosite  = watershed_data_MBCn.clt_ssp245(:,
    ismember(station_selecttion(1,:), [1]));
118-
119-        d2m        = watershed_data_MBCn.d2m_ssp245(:,
    ismember(station_selecttion(1,:), [1]));
120-        %% vapour pressure calculation using titan's equation.
121-        temp_2m     = 0.5*(tMax + tMin);
122-        n_rows = size(temp_2m, 1);
123-        n_col = size(temp_2m, 2);
124-        pression = zeros(n_rows, n_col);
125-        for k = 1:n_rows
126-            for j = 1:n_col
127-                if d2m(k, j) >= 0
128-                    pression(k, j) = 0.61078*exp(17.27*d2m(k, j)/(d2m(k, j)+237.3));
129-                else
130-                    pression(k, j) = 0.61078*exp(21.875*d2m(k, j)/(d2m(k, j)+265.5));
131-                end
132-            end
133-        end
134-        pression    = 7.50062*pression; %convert from kPa to mmHg
135-        %% Loading the structure
136-        load(strcat(pwd,'\', commonDataFolder, '\', CEQUEAU_structure))
137-        %% We need to add leap days
138-        % this will be done using retime function and we need to do it for each var
139-        % change to daily time step
140-        InputStruct.parametres.option.ipassim = 24;
141-        t_start = datenum(2040,1,1,0,0,0);
142-        t_end   = datenum(2069,12,31,21,0,0);
143-        InputStruct.execution.dateDebut  =  t_start;% row =
144-        InputStruct.execution.dateFin    =  t_end;
145-        t_cequeau = (datenum(t_start:hours(3):t_end))';
146-        date1 = datetime(watershed_data_MBCn.dates_ssp245(:,:));
147-        date2 = datetime(datevec(t_cequeau));
148-
149-        t = t_cequeau;
150-        t_daily = (datenum(t_start:hours(24):t_end))';
151-        NNN = size(t_cequeau, 1);
152-        NNN_daily = size(t_daily, 1);
153-        %tMax
154-        TT_eccc_tMax = timetable(date1, tMax);
155-        TT2 = retime(TT_eccc_tMax, date2, 'nearest');
156-        TT2_daily = retime(TT2,'daily','max');
157-        tMax = TT2{:,:};
```

```
158-        tMax_daily = TT2_daily{:,:};
159-        %tMin
160-        TT_eccc_tMin = timetable(date1, tMin);
161-        TT2 = retime(TT_eccc_tMin, date2, 'nearest');
162-        TT2_daily = retime(TT2,'daily','min');
163-        tMin = TT2{:,:};
164-        tMin_daily = TT2_daily{:,:};
165-        %pTot
166-        TT_eccc_pTot = timetable(date1, pTot);
167-        TT2 = retime(TT_eccc_pTot, date2, 'nearest');
168-        TT2_daily = retime(TT2,'daily','sum');
169-        pTot = TT2{:,:};
170-        pTot_daily = TT2_daily{:,:};
171-        %rayonnement
172-        TT_eccc_rayonnement = timetable(date1, rayonnement);
173-        TT2 = retime(TT_eccc_rayonnement, date2, 'nearest');
174-        TT2_daily = retime(TT2,'daily','sum');
175-        rayonnement = TT2{:,:};
176-        rayonnement_daily = TT2_daily{:,:};
177-        %vitesseVent
178-        TT_eccc_vitesseVent = timetable(date1, vitesseVent);
179-        TT2 = retime(TT_eccc_vitesseVent, date2, 'nearest');
180-        TT2_daily = retime(TT2,'daily','mean');
181-        vitesseVent = TT2{:,:};
182-        vitesseVent_daily = TT2_daily{:,:};
183-        %nebulosite
184-        TT_eccc_nebulosite = timetable(date1, nebulosite);
185-        TT2 = retime(TT_eccc_nebulosite, date2, 'nearest');
186-        TT2_daily = retime(TT2,'daily','mean');
187-        nebulosite = TT2{:,:};
188-        nebulosite_daily = TT2_daily{:,:};
189-        %pression
190-        TT_eccc_pression = timetable(date1, pression);
191-        TT2 = retime(TT_eccc_pression, date2, 'nearest');
192-        TT2_daily = retime(TT2,'daily','mean');
193-        pression = TT2{:,:};
194-        pression_daily = TT2_daily{:,:};
195-        %%
196-        meteoStation = struct('t', t_daily,'tMax', tMax_daily,'tMin', ...
197-            tMin_daily,'pTot', pTot_daily, 'rayonnement', rayonnement_daily,...
198-            'vitesseVent', vitesseVent_daily, 'nebulosite', nebulosite_daily,...
199-            'pression', pression_daily);
200-        %% Changing boundary conditions
201-        load(strcat(pwd,'\', commonDataFolder, '\', 'qwo_2040_2069_ssp245_daily.mat'));
202-        % the name of this variable is "qwo_2040_2069_ssp245_daily" with 8 columns
203-        % for 8 climate models
204-
205-        %% increasing the flow at SLS 25% or 50%
206-        ddate = datevec(t_daily(:,1));
207-        summerFlow = qwo_2040_2069_ssp245_daily(and(ddate(:,2) >=6, ddate(:,2) <=9),
    :);
208-        summerFlow(:,:) = summerFlow(:,:).* 1;
209-        % summerFlow(:,:) = summerFlow(:,:).* 1.25;
210-        % summerFlow(:,:) = summerFlow(:,:).* 1.5;
211-        qwo_2040_2069_ssp245_daily(and(ddate(:,2) >=6, ddate(:,2) <=9), :) =
    summerFlow;
212-        %%
213-        id = [1:1233];
214-        temperature = NaN*ones(1,1233);
215-        temperature (1,1109) = 0;
216-        etatsCP   = struct('id', id,'temperature', temperature);
217-
218-        load(strcat(pwd,'\', commonDataFolder, '\', 'two_2040_2069_ssp245_daily.mat'));
219-        % the name of this variable is "two_2040_2069_ssp245_daily" with 8 columns
220-        % for 8 climate models
221-        InputStruct.bassinVersant.barrage(1).debit  = qwo_2040_2069_ssp245_daily(:,
    database);
222-        InputStruct.assimilations.qualite = [];
223-        for zx = 1:NNN_daily
224-            InputStruct.assimilations.qualite(zx).pasDeTemps = t_start + zx -1;
225-            InputStruct.assimilations.qualite(zx).etatsCP = etatsCP;
226-            % below line is changing water temprature at SLS.
227-            % min value is column 2 and max value for each day is column 3
228-            InputStruct.assimilations.qualite(zx).etatsCP.temperature(1, 1109) =
    two_2040_2069_ssp245_daily(zx, database);
229-        end
230-
```

```
231-        %% Placing new created data in the current structure
232-        InputStruct.stations = stations;
233-        InputStruct.meteoStation = meteoStation;
234-        %% Producing interpolated data for all variables
235-        InputStruct.meteoPointGrille = [];
236-        meteoInterpolee = cequeauInterpolationMex(InputStruct.execution,
    InputStruct.parametres, InputStruct.bassinVersant, InputStruct.stations,
    InputStruct.meteoStation);
237-
238-        InputStruct.meteoPointGrille = meteoInterpolee;
239-        InputStruct.meteoStation   = rmfield(InputStruct.meteoStation  , 'tMax');
240-        InputStruct.meteoStation   = rmfield(InputStruct.meteoStation  , 'tMin');
241-
242-        InputStruct.meteoStation = renameStructField(InputStruct.meteoStation,
    'rayonnement', 'tMax');
243-        InputStruct.meteoStation = renameStructField(InputStruct.meteoStation,
    'vitesseVent', 'tMin');
244-        meteoInterpolee = cequeauInterpolationMex(InputStruct.execution,
    InputStruct.parametres, InputStruct.bassinVersant, InputStruct.stations,
    InputStruct.meteoStation);
245-        meteoInterpolee = renameStructField(meteoInterpolee, 'tMax', 'rayonnement');
246-        meteoInterpolee = renameStructField(meteoInterpolee, 'tMin', 'vitesseVent');
247-
248-        InputStruct.meteoPointGrille.rayonnement = meteoInterpolee.rayonnement;
249-        InputStruct.meteoPointGrille.vitesseVent = meteoInterpolee.vitesseVent;
250-
251-        InputStruct.meteoStation = renameStructField(InputStruct.meteoStation,
    'nebulosite', 'tMax');
252-        InputStruct.meteoStation = renameStructField(InputStruct.meteoStation,
    'pression', 'tMin');
253-        meteoInterpolee = cequeauInterpolationMex(InputStruct.execution,
    InputStruct.parametres, InputStruct.bassinVersant, InputStruct.stations,
    InputStruct.meteoStation);
254-        meteoInterpolee = renameStructField(meteoInterpolee, 'tMax', 'nebulosite');
255-        meteoInterpolee = renameStructField(meteoInterpolee, 'tMin', 'pression');
256-
257-        InputStruct.meteoPointGrille.nebulosite = meteoInterpolee.nebulosite;
258-        InputStruct.meteoPointGrille.pression = meteoInterpolee.pression;
259-        % to keep all data in meteoStation field
260-        InputStruct.meteoStation = meteoStation;
261-        % save(strcat(pwd,'\', dataFolder, '\', 'struct_forSimulation.mat'),
    'InputStruct');
262-        %% Add spin up period
263-        % this period is a repetition of first 10 years
264-        firstDate = datevec(InputStruct.meteoStation.t(1,1));
265-        firstYear = firstDate(1);
266-        % the number 10 was chosen after trial and error
267-        DELTA = 10;
268-        spinupStart = datenum(firstYear-DELTA,1,1);
269-        spinupEnd = datenum(firstYear-1,12,31);
270-        spinupTime = (spinupStart:spinupEnd)';
271-        spinLength = size(spinupTime, 1);
272-        % spin up structure building
273-        NewStruct = InputStruct;
274-        NewStruct.meteoStation.t = [spinupTime;InputStruct.meteoStation.t];
275-        NewStruct.meteoPointGrille.tMin =
    [InputStruct.meteoPointGrille.tMin(1:spinLength, :);
    InputStruct.meteoPointGrille.tMin];
276-        NewStruct.meteoPointGrille.tMax =
    [InputStruct.meteoPointGrille.tMax(1:spinLength, :);
    InputStruct.meteoPointGrille.tMax];
277-        NewStruct.meteoPointGrille.pTot =
    [InputStruct.meteoPointGrille.pTot(1:spinLength, :);
    InputStruct.meteoPointGrille.pTot];
278-        NewStruct.meteoPointGrille.rayonnement =
    [InputStruct.meteoPointGrille.rayonnement(1:spinLength, :);
    InputStruct.meteoPointGrille.rayonnement];
279-        NewStruct.meteoPointGrille.vitesseVent =
    [InputStruct.meteoPointGrille.vitesseVent(1:spinLength, :);
    InputStruct.meteoPointGrille.vitesseVent];
280-        NewStruct.meteoPointGrille.nebulosite =
    [InputStruct.meteoPointGrille.nebulosite(1:spinLength, :);
    InputStruct.meteoPointGrille.nebulosite];
281-        NewStruct.meteoPointGrille.pression =
    [InputStruct.meteoPointGrille.pression(1:spinLength, :);
    InputStruct.meteoPointGrille.pression];
282-
```

```
283-        NewStruct.bassinVersant.barrage(1).debit  =
    [InputStruct.bassinVersant.barrage(1).debit(1:spinLength, :);
    InputStruct.bassinVersant.barrage(1).debit];
284-        % new simulation start time
285-        NewStruct.execution.dateDebut = NewStruct.meteoStation.t (1,1);
286-        InputStruct = NewStruct;
287-        %% Simulation
288-        % after several tries for calibration, we agreed on this set
289-        %% Hydrological Optimization Parameters
290-        load(strcat(pwd,'\', commonDataFolder, '\', 'x_hydro.mat'));
291-        load(strcat(pwd,'\', commonDataFolder, '\', 'x_thermal.mat'));
292-
293-        InputStruct.parametres.sol.cin_s = x_hydro(1);
294-        InputStruct.parametres.sol.cvmar = x_hydro(2);
295-        InputStruct.parametres.sol.cvnb_s = x_hydro(3);
296-        InputStruct.parametres.sol.cvnh_s = x_hydro(4);
297-        InputStruct.parametres.sol.cvsb = x_hydro(5);
298-        InputStruct.parametres.sol.cvsi_s = x_hydro(6);
299-        InputStruct.parametres.sol.xinfma = x_hydro(7);
300-        InputStruct.parametres.sol.hinf_s = x_hydro(8);
301-        InputStruct.parametres.sol.hint_s = x_hydro(9);
302-        InputStruct.parametres.sol.hmar = x_hydro(10);
303-        InputStruct.parametres.sol.hnap_s = x_hydro(11);
304-        InputStruct.parametres.sol.hpot_s = x_hydro(12);
305-        InputStruct.parametres.sol.hsol_s = x_hydro(13);
306-        InputStruct.parametres.sol.hrimp_s = x_hydro(14);
307-        InputStruct.parametres.transfert.exxkt = x_hydro(15);
308-        InputStruct.parametres.neige.strne_s = x_hydro(16);
309-        InputStruct.parametres.fonte.cequeau.strne_s = x_hydro(16);
310-        InputStruct.parametres.neige.tfc_s = x_hydro(17);
311-        InputStruct.parametres.fonte.cequeau.tfc_s = x_hydro(17);
312-        InputStruct.parametres.neige.tfd_s = x_hydro(18);
313-        InputStruct.parametres.fonte.cequeau.tfd_s = x_hydro(18);
314-        InputStruct.parametres.neige.tsc_s = x_hydro(19);
315-        InputStruct.parametres.fonte.cequeau.tsc_s = x_hydro(19);
316-        InputStruct.parametres.neige.tsd_s = x_hydro(20);
317-        InputStruct.parametres.fonte.cequeau.tsd_s = x_hydro(20);
318-        InputStruct.parametres.neige.ttd = x_hydro(21);
319-        InputStruct.parametres.fonte.cequeau.ttd = x_hydro(21);
320-        InputStruct.parametres.neige.tts_s = x_hydro(22);
321-        InputStruct.parametres.fonte.cequeau.tts_s = x_hydro(22);
322-        InputStruct.parametres.evapo.cequeau.evnap = x_hydro(23);
323-        InputStruct.parametres.evapo.cequeau.xaa = x_hydro(24);
324-        InputStruct.parametres.evapo.cequeau.xit = x_hydro(25);
325-        InputStruct.parametres.transfert.zn = x_hydro(26);
326-        %% Thermal Optimization Parameters
327-        InputStruct.parametres.qualite.cequeau.temperat.crayso = x_thermal(1);
328-        InputStruct.parametres.qualite.cequeau.temperat.crayin = x_thermal(2);
329-        InputStruct.parametres.qualite.cequeau.temperat.cevapo = x_thermal(3);
330-        InputStruct.parametres.qualite.cequeau.temperat.cconve = x_thermal(4);
331-        InputStruct.parametres.qualite.cequeau.temperat.crigel = x_thermal(5);
332-        InputStruct.parametres.qualite.cequeau.temperat.tnap = x_thermal(6);
333-        InputStruct.parametres.qualite.cequeau.temperat.panap = 0;
334-        InputStruct.parametres.qualite.cequeau.temperat.tinit = 0;
335-        InputStruct.parametres.qualite.cequeau.temperat.bassol = x_thermal(7);
336-        InputStruct.parametres.qualite.cequeau.temperat.corsol = x_thermal(8);
337-        %%
338-        [y.etatsCE, y.etatsCP, y.etatsFonte, y.etatsEvapo, y.etatsBarrage,
    y.pasDeTemps,...
339-            y.avantAssimilationsCE, y.avantAssimilationsFonte,
    y.avantAssimilationsEvapo,y.etatsQualCP, y.avAssimQual] = ...
340-        cequeauQuantiteMex_v461(InputStruct.execution, InputStruct.parametres,
    InputStruct.bassinVersant, ...
341-             InputStruct.meteoPointGrille, [], InputStruct.assimilations);
342-         %% Plotting the result
```

APPENDIX III

Appendices for Article III: Assessing the surface downward longwave irradiance models using ERA5 input data in Canada

Appendix III-A: Sensor locations in the Nechako and BVESM watersheds

(a) (b)

(c) (d)

Figure III-A1: The location of two installed sensors (a) SLS at the Nechako watershed within the ERA5 grid cell boundaries, (b) Zoomed image of SLS location. (c) and (d) similar to (a), and (b) just for BVESM station at the watershed with the same name

Appendix III-B: Parameter calibration results for DLI models at the SLS station

Table III-B1: Calibrated parameter values after the calibration (each calibration was repeated 10 times) step for the nine DLI models in comparison with observed data at the SLS station (a) mean, (b) standard deviation, and (c) coefficient of variations for each calibrated parameter.

(a) mean of resulted parameters for each DLI model									
Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
u	-	0.222	0.173	0.170	0.221	0.130	0.185	0.156	0.344
v	-	2.662	4.000	2.000	1.762	4.000	3.742	4.000	3.996
a	0.818	-	-	-	-	-	-	-	-
α	-	-	-	-	-	-	-	-	0.242

(b) standard deviation of resulted parameters for each DLI model									
Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
u	-	0.002	0.001	0.000	0.002	0.001	0.002	0.001	0.003
v	-	0.109	0.000	0.000	0.069	0.000	0.176	0.000	0.012
a	0.001	-	-	-	-	-	-	-	-
α	-	-	-	-	-	-	-	-	0.002

(c) coefficient of variation of resulted parameters for each DLI model (%)									
Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
u	-	0.856	0.731	0.000	0.898	0.983	0.849	0.826	0.794
v	-	4.087	0.000	0.000	3.926	0.000	4.693	0.000	0.305
a	0.088	-	-	-	-	-	-	-	-
α	-	-	-	-	-	-	-	-	0.687

Table III-B2: Calibrated parameter values after the calibration (each calibration was repeated 10 times) step for the nine DLI models in comparison with observed data at the BVESM station (a) mean, (b) standard deviation, and (c) coefficient of variations for each calibrated parameter.

(a) mean of resulted parameters for each DLI model									
Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
u	-	0.200	0.172	0.170	0.194	0.140	0.169	0.140	0.303
v	-	1.745	2.520	2.000	1.462	4.000	1.892	4.000	2.312
a	0.856	-	-	-	-	-	-	-	-
α	-	-	-	-	-	-	-	-	0.235

(b) standard deviation of resulted parameters for each DLI model									
Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
u	-	0.002	0.002	0.000	0.002	0.001	0.002	0.002	0.008
v	-	0.116	0.197	0.000	0.105	0.000	0.149	0.000	0.591
a	0.001	-	-	-	-	-	-	-	-
α	-	-	-	-	-	-	-	-	0.005

(c) coefficient of variation of resulted parameters for each DLI model (%)									
Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
u	-	1.078	1.130	0.000	1.108	0.933	1.075	1.376	2.734
v	-	6.626	7.831	0.000	7.204	0.000	7.862	0.000	25.550
a	0.071	-	-	-	-	-	-	-	-
α	-	-	-	-	-	-	-	-	2.063

Table III-B3: Summary of performance metric values after the calibration (each calibration was repeated 10 times) step for the nine DLI models in comparison with observed data at the SLS station (a) mean, (b) standard deviation, and (c) coefficient of variations of performance metrics for each calibrated model.

(a) mean of resulted performance metrics for each DLI model								
Model Name	RMSE (W/m^2)	RRMSE	Correlation	R^2	NSE	Bias (W/m^2)	AICc	BICc
ERA5	32.02	11.81	0.78	0.60	0.55	5.45	101815.95	101815.95
Model 1	37.58	13.87	0.62	0.39	0.38	0.48	105155.79	105163.04
Model 2	37.02	13.66	0.73	0.53	0.40	4.22	104844.14	104851.39
Model 3	33.42	12.33	0.73	0.53	0.51	-0.66	102709.28	102716.53
Model 4	32.07	11.84	0.75	0.56	0.55	-0.82	101853.98	101853.98
Model 5	35.92	13.25	0.77	0.59	0.44	5.12	104214.90	104222.15
Model 6	33.57	12.39	0.75	0.56	0.51	-2.27	102804.42	102811.67
Model 7	32.37	11.94	0.76	0.58	0.54	1.80	102045.58	102052.83
Model 8	32.53	12.00	0.75	0.56	0.54	-1.15	102149.12	102156.37
Model 9	32.29	11.92	0.74	0.55	0.54	0.67	101998.36	102005.61

(b) standard deviation of resulted performance metrics for each DLI model								
Model Name	RMSE (W/m^2)	RRMSE	Correlation	R^2	NSE	Bias (W/m^2)	AICc	BICc
ERA5	0.13	0.05	0.00	0.00	0.01	0.15	302.58	302.58
Model 1	0.12	0.04	0.00	0.01	0.01	0.03	309.74	309.75
Model 2	0.09	0.04	0.00	0.00	0.01	0.13	304.04	304.04
Model 3	0.10	0.03	0.00	0.00	0.00	0.09	316.60	316.61
Model 4	0.09	0.03	0.00	0.00	0.00	0.15	313.15	313.15
Model 5	0.09	0.04	0.00	0.00	0.01	0.11	296.81	296.81
Model 6	0.07	0.03	0.00	0.00	0.01	0.10	302.97	302.97
Model 7	0.08	0.03	0.00	0.00	0.00	0.12	310.05	310.06
Model 8	0.08	0.03	0.00	0.00	0.00	0.10	304.33	304.33
Model 9	0.10	0.03	0.00	0.00	0.00	0.04	319.46	319.46

(c) coefficient of variation of resulted performance metrics for each DLI model (%)								
Model Name	RMSE	RRMSE	Correlation	R^2	NSE	Bias	AICc	BIC
ERA5	0.42	0.40	0.32	0.64	1.01	2.69	0.30	0.30
Model 1	0.32	0.31	0.65	1.31	1.53	6.10	0.29	0.29
Model 2	0.24	0.26	0.30	0.60	1.65	3.13	0.29	0.29
Model 3	0.30	0.26	0.29	0.58	0.73	-13.95	0.31	0.31
Model 4	0.28	0.26	0.29	0.58	0.75	-18.02	0.31	0.31
Model 5	0.24	0.29	0.28	0.55	1.68	2.23	0.28	0.28
Model 6	0.21	0.26	0.35	0.69	1.13	-4.24	0.29	0.29
Model 7	0.23	0.23	0.29	0.57	0.88	6.81	0.30	0.30
Model 8	0.26	0.27	0.30	0.61	0.83	-8.35	0.30	0.30
Model 9	0.31	0.27	0.33	0.66	0.73	5.61	0.31	0.31

Table III-B4: Summary of performance metric values after the calibration (each calibration was repeated 10 times) step for the nine DLI models in comparison with observed data at the BVESM station (a) mean, (b) standard deviation, and (c) coefficient of variations of performance metrics for each calibrated model.

(a) mean of resulted performance metrics for each DLI model								
Model Name	RMSE (W/m^2)	RRMSE	Correlation	R^2	NSE	Bias (W/m^2)	AICc	BICc
ERA5	30.37	9.89	0.88	0.78	0.76	4.98	49680.61	49680.61
Model 1	36.02	11.73	0.81	0.66	0.66	-0.22	51436.68	51443.22
Model 2	38.36	12.49	0.85	0.72	0.61	6.69	52082.47	52089.02
Model 3	35.29	11.50	0.84	0.70	0.67	3.11	51226.52	51233.06
Model 4	32.02	10.43	0.87	0.75	0.73	6.55	50222.84	50222.84
Model 5	37.74	12.29	0.88	0.78	0.63	8.75	51916.14	51922.69
Model 6	31.66	10.31	0.88	0.77	0.74	2.55	50111.04	50117.58
Model 7	32.80	10.69	0.88	0.77	0.72	4.70	50474.32	50480.86
Model 8	34.40	11.21	0.86	0.74	0.69	2.50	50963.36	50969.91
Model 9	30.76	10.02	0.87	0.75	0.75	0.34	49812.18	49818.72

(b) standard deviation of resulted performance metrics for each DLI model								
Model Name	RMSE (W/m^2)	RRMSE	Correlation	R^2	NSE	Bias (W/m^2)	AICc	BICc
ERA5	0.18	0.06	0.00	0.00	0.00	0.23	376.75	376.75
Model 1	0.10	0.04	0.00	0.00	0.00	0.06	378.67	378.67
Model 2	0.26	0.08	0.00	0.00	0.01	0.20	414.36	414.37
Model 3	0.18	0.06	0.00	0.00	0.01	0.23	391.32	391.33
Model 4	0.16	0.05	0.00	0.00	0.00	0.25	381.30	381.30
Model 5	0.25	0.08	0.00	0.00	0.01	0.16	420.10	420.11
Model 6	0.19	0.06	0.00	0.00	0.00	0.18	390.40	390.41
Model 7	0.19	0.06	0.00	0.00	0.01	0.17	394.67	394.68
Model 8	0.20	0.06	0.00	0.00	0.00	0.23	393.83	393.84
Model 9	0.14	0.05	0.00	0.00	0.00	0.06	372.62	372.63

(c) coefficient of variation of resulted performance metrics for each DLI model (%)								
Model Name	RMSE	RRMSE	Correlation	R2	NSE	Bias	AICc	BIC
ERA5	0.58	0.58	0.29	0.58	0.65	4.61	0.76	0.76
Model 1	0.28	0.38	0.32	0.63	0.61	-25.56	0.74	0.74
Model 2	0.67	0.65	0.28	0.56	1.27	2.98	0.80	0.80
Model 3	0.52	0.53	0.26	0.52	0.82	7.50	0.76	0.76
Model 4	0.49	0.48	0.24	0.48	0.66	3.81	0.76	0.76
Model 5	0.66	0.63	0.21	0.43	1.11	1.83	0.81	0.81
Model 6	0.59	0.59	0.23	0.47	0.66	6.91	0.78	0.78
Model 7	0.57	0.53	0.22	0.43	0.72	3.56	0.78	0.78
Model 8	0.59	0.57	0.19	0.37	0.71	9.13	0.77	0.77
Model 9	0.47	0.49	0.26	0.52	0.54	18.50	0.75	0.75

Appendix III-C: The MATLAB codes implemented in Article III

The DLI functions in paper III:

```
function [DLI_sim, Metrics] = DLI_functions(ModelNumber, x, INPUT_OBS, INPUT_ERA5)
% Col:      1       2           3           4           5           6       7       8           9
% INPUT_OBS:    TIME    INCOMING_LW OUTGOING_LW INCOMING_SW OUTGOING_SW TAIR(K)
% INPUT_ERA5:   Time    Ta(K)       Ta(?C)      Tdw(K)      Tdw(?C)     B(DL)       ea(mmHg)    LAIhv(DL)   DLI_EAR5
    s = 5.670373*10^-8;
    Name = strcat('Model', num2str(ModelNumber));
    DLI_obs = INPUT_OBS(:, 2);
    switch ModelNumber
        case 0
            DLI_sim = INPUT_ERA5(:, 9);
        case 1
            DLI_sim = x(1).*s.*(INPUT_ERA5(:,2)).^4;
        case 2
            DLI_sim = (1+x(1).*(INPUT_ERA5(:,6)).^x(2)).*...
                (9.3645*10^-6.*(INPUT_ERA5(:,2)).^2).*...
                s.*(INPUT_ERA5(:,2)).^4;
        case 3
            DLI_sim = (1+x(1).*(INPUT_ERA5(:,6)).^x(2)).*...
                (1-0.261*exp(-7.77*10^(-4)*(273 - INPUT_ERA5(:,2)).^2)).*...
                s.*(INPUT_ERA5(:,2)).^4;
        case 4
            DLI_sim = (1+0.17.*(INPUT_ERA5(:,6)).^2).*...
                0.97.*(0.74+0.0065.*(INPUT_ERA5(:,7))).*...
                s.*(INPUT_ERA5(:,2)).^4;

        case 5
            DLI_sim = (1+x(1).*(INPUT_ERA5(:,6)).^x(2)).*...
                (1.24.*((1.3332239 * INPUT_ERA5(:,7)./ INPUT_ERA5(:,2))).^(1/7)).*...
                s.*(INPUT_ERA5(:,2)).^4;
        case 6
            DLI_sim = (1+x(1).*(INPUT_ERA5(:,6)).^x(2)).*...
                (1.08.*(1-exp(-(1.3332239 * INPUT_ERA5(:,7)).^(INPUT_ERA5(:,2)./2016)))).*...
                s.*(INPUT_ERA5(:,2)).^4;

        case 7
            DLI_sim = (1+x(1).*(INPUT_ERA5(:,6)).^x(2)).*...
                (1-(1+46.5.*((1.3332239 * INPUT_ERA5(:,7))./INPUT_ERA5(:,2))).*...
                exp(-1*(1.2 + 3*(46.5*((1.3332239 * INPUT_ERA5(:,7))...
                ./INPUT_ERA5(:,2)))).^(1/2))).*...
                s.*(INPUT_ERA5(:,2)).^4;
        case 8
            ea = 1.3332239 * INPUT_ERA5(:, 7);
            DLI_sim(or(ea>=2,isnan(ea)), :) = (1 + x(1).*(INPUT_ERA5(or(ea>=2,isnan(ea)), 6)).^x(2)).*...
                (0.72 + 0.009 .* (ea(or(ea>=2,isnan(ea)), :) - 2)).*...
                s.*(INPUT_ERA5(or(ea>=2,isnan(ea)),2)).^4;
            DLI_sim(ea<2, :) = (1 + x(1).*(INPUT_ERA5(ea<2, 6)).^x(2)).*...
                (0.72 - 0.076 .* (ea(ea<2, :) - 2)).*...
                s.*(INPUT_ERA5(ea<2,2)).^4;
        case 9
            DLI_sim = ((1 + x(1).*(INPUT_ERA5(:,6)).^x(2)) .* ...
                ((exp(-1*x(1)*INPUT_ERA5(:,8))).*0.97.*(0.74 + 0.0065.*INPUT_ERA5(:,7))) + ...
                0.98.*(1 - exp(-1*x(3)*INPUT_ERA5(:,8)))) .* ...
            s.*(INPUT_ERA5(:,2)).^4;
    end
    k = size(x, 2);
    N =  sum(~isnan(DLI_obs));
    Metrics(1, 1) = RMSE(DLI_obs, DLI_sim);
    Metrics(1, 2) = RRMSE(DLI_obs, DLI_sim);
    Metrics(1, 3) = nancorr(DLI_obs, DLI_sim);
    Metrics(1, 4) = R2(DLI_obs, DLI_sim);
    Metrics(1, 5) = NSE(DLI_obs, DLI_sim);
    Metrics(1, 6) = BIAS(DLI_obs, DLI_sim);
    Metrics(1, 7) = AIC(DLI_obs, DLI_sim, k);
    Metrics(1, 8) = BIC(DLI_obs, DLI_sim, k);
end
```

APPENDIX IV

Appendices for Article IV: Modeling the impact of downward longwave irradiance on water temperature at the watershed scale using ERA5 input data in western Canada.

Appendix IV-A: Cooling and warming impacts of DLI formulation on water's longwave heat budget (H_{long})

$$H_{long} = H_{long\downarrow} - H_{long\uparrow} = CRAYIN.A.\sigma(\beta_{air}T_{air}^4 - \beta_{water}T_{water}^4) \quad \text{(IV-A-1)}$$

$$\beta_{air}T_{air}^4 - \beta_{water}T_{water}^4 = (\sqrt[4]{\beta_{air}}T_{air} - \sqrt[4]{\beta_{water}}T_{water})(\sqrt[4]{\beta_{air}}T_{air} + \sqrt[4]{\beta_{water}}T_{water})(\sqrt{\beta_{air}}T_{air}^2 + \sqrt{\beta_{water}}T_{water}^2) \quad \text{(IV-A-2)}$$

The expression is equal to zero when $T_{air} = \sqrt[4]{\frac{\beta_{water}}{\beta_{air}}}T_{water}$

where $H_{long\downarrow}$ and $H_{long\uparrow}$ are DLI and OLR respectively (MJ); A is the water surface area in the CP (m^2); $CRAYIN$, is empirical dimensionless coefficient for thermal module calibration in CEQUEAU (presented in Table IV-D3 in Appendix IV-D); β_{air} and β_{water} are the sky and water emissivities (dimensionless, $\beta_{water} = 0.97$ in CEQUEAU); T_{air} and T_{water} are air and water temperatures, respectively (K) for the current CP.

Appendix IV-B: Process-based river temperature models and references

Table IV-B1: List of process-based river temperature models [based on Norton and Bradford (2009), Ficklin, Luo, Stewart, and Maurer (2012), Dugdale, Hannah, and Malcolm (2017), and Ouellet et al. (2020)]

No.	Model Name	Reference	Model Heat Fluxes	Minimum Time Step	Programming Language	Availability	Source Code Availability
1	BasinTemp	Allen (2008)	Surface, groundwater	Daily	N/a	Proprietary (Stillwater Sciences)	No
2	CE-QUAL-W2	Cole and Wells (2015)	Surface, groundwater, bed conduction	1 second	Fortran/visual basic	Free download	Yes
3	CEQUEAU	Morin and Paquet (2007)	Surface, groundwater	Sub daily	MATLAB/C++	Available on request	Yes
4	CrUSTe	LeBlanc et al. (1997)	Surface, groundwater	Hourly	STELLA	N/a	N/a
5	Delft3D-FLOW	Deltares (2014)	Surface water, bed conduction	1 second	Fortran	Free download	Yes
6	Heat source	Boyd and Kasper (2003)	Surface, groundwater, bed conduction	Hourly	Python/visual basic	Free download	Yes
7	DHVSM-RBM	Sun et al. (2015), Yearsley et al. (2001)	Surface, subsurface water	Hourly	Fortran	Free download	Yes
8	GAWSER/ GRIFFS	Schroeter (1996)	Surface water	1 min	N/a	N/a	No
9	GIS-STRTemp	Sansone (2001)	Surface, groundwater, bed conduction	Sub daily	N/a	N/a	No
10	HEC-RAS	Brunner (2016)	Surface, groundwater, bed conduction	1 second	Java	Free download	No
11	MIKE 11	DHI (2016)	Surface, groundwater	1 second	N/a	Commercially available	No
12	MNSTREM	Sinokrot and Stefan (1993)	Surface, groundwater, bed conduction	Daily	Fortran	Free download	Yes
13	MRSTM	Issak et al. [2009]	Stream network, geomorphology, climate, landscape features, fire effects	-	N/a	N/a	N/a
14	Qual2K	Chapra et al. (2012)	Surface, groundwater, bed conduction	Sub hourly	Fortran / Visual Basic	Free download	Yes
15	RAFT	Pike et al. (2013)	Surface water, bed conduction	1 min, sub hourly	N/a	N/a	N/a
16	RMA-11	King (2016)	Surface water	1 sec or less	Fortran	Proprietary (Resource Modelling Associates)	N/a
17	SHADE-HSPF	Becknell et al. (1997)	Surface, groundwater, bed conduction	Hourly	Fortran	Free download	Yes
18	SNTemp	Theurer et al. (1984)	Surface, groundwater, bed conduction, friction	Daily	Basic, Fortran	Free download	Yes
19	Streamline	Rutherford et al. (1997)	Surface, groundwater, bed conduction	15 min	Fortran / Visual Basic	Available on request	N/a
20	SWAT	Arnold et al. [1998]	Surface, groundwater,	Sub daily	Fortran	Free download	Yes
21	TEMP-84	Beschta and Weatherred (1984)	N/a	Daily		N/a	N/a
22	TVA-RMS	Deas et al. (2003)	Surface water	Hourly	C	Available on request	Yes
23	WAIORA	Jowett et al. (2004)	Surface water	Hourly	Delphi	Free download	No
24	WASP7	Wool et al. (2008)	Surface water	Hourly	Fortran	Free download	Yes
25	WET-Temp	Cox and Bolte (2007)	Surface water	Hourly	C++	Available on request	Yes

Appendix IV-C: Modular code diagram of CEQUEAU model

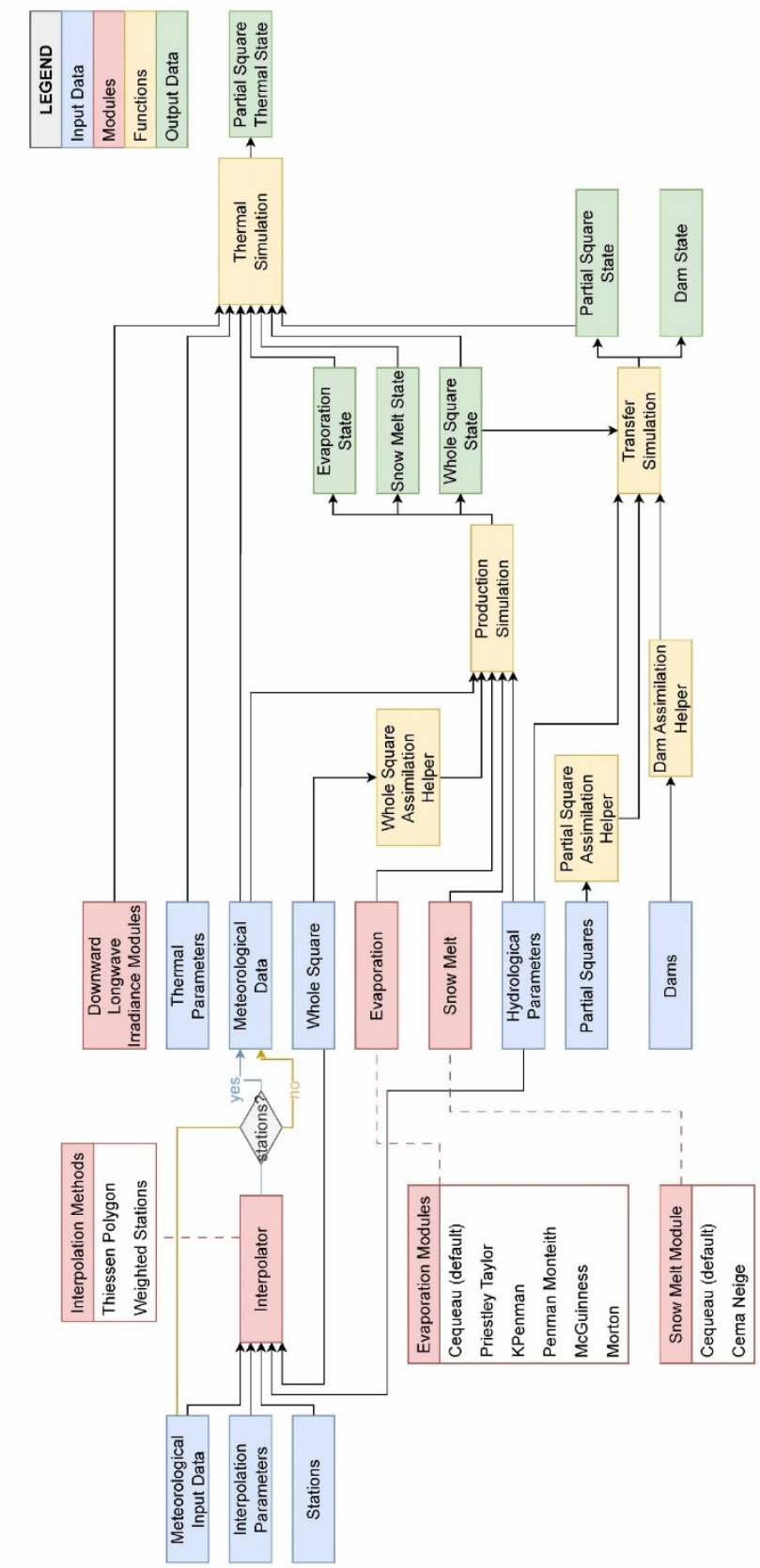

Figure IV-C1: Modular code diagram of CEQUEAU, including coding modules for (1) input data, (2) model modules (subroutines), (3) functions, and (4) output data. Each sub-module refers to one or several C++ file that calls the connected other sub-modules [from left to right]

Appendix IV-D: All parameters used in the CEQUEAU model calibrations

This section lists all calibration parameters used in the CEQUEAU model. The parameters include (1) hydrological calibration parameters, and (2) thermal calibration parameters.

Table IV-D1: List of hydrological calibration parameters used in the CEQUEAU model

No.	Parameter	Unit	Calibrated value	Description
1	CIN	dimensionless	0.468	Percolation coefficient from the upper zone soil reservoir to the lower zone.
2	CVMAR	dimensionless	0.475	Drainage coefficient of the lakes and marshes reservoir.
3	CVNB	dimensionless	0.000086	Lower-zone soil reservoir's lower drainage coefficient
4	CVNH	dimensionless	0.000297	Lower-zone soil reservoir's upper drainage coefficient
5	CVSB	dimensionless	0.000838	Upper-zone soil reservoir's lower drainage coefficient
6	CVSI	dimensionless	0.496	Upper-zone soil reservoir's intermediate drainage coefficient
7	XINFMA	mm/day	10.606	Daily maximum infiltration
8	HINF	mm	13.539	Percolation threshold from the upper to the lower zone reservoir
9	HINT	mm	93.901	Upper-zone reservoir intermediate drainage threshold.
10	HMAR	mm	268.707	Drainage threshold of the lakes and marshes reservoir.
11	HNAP	mm	314.815	Lower-zone reservoir upper threshold
12	HPOT	mm	82.673	Threshold of evaporation at the potential rate
13	HSOL	mm	194.234	Upper-zone reservoir runoff threshold
14	HRIMP	mm	4.155	Upper-zone reservoir runoff threshold for impermeable surfaces
15	EXXKT	dimensionless	0.005	Parameter for calculating the routing coefficient of partial squares
16	STRNE	°C	0.638	Rain-snow transition threshold
17	TFC	mm/°C/day	5.135	Melting rate in the forested area
18	TFD	mm/°C/day	7.453	Melting rate in open area
19	TSC	°C	-0.033	Melting threshold in the forested area
20	TSD	°C	0.843	Melting threshold in the open area

Table IV-D1: List of hydrological calibration parameters used in the CEQUEAU model (Continued)

No.	Parameter	Unit	Calibrated value	Description
21	TTD	-	1.207	Heat deficit coefficient
22	TTS	°C	-0.635	Minimum temperature for snow stock ripening
23	EVNAP	%	0.354	Percentage of daily evapotranspiration taken from Lower zone reservoir
24	ZN	in simulation time steps (days)	1.527	Concentration time, i.e., time for water to travel from the highest upstream point to the outlet.

Table IV-D2: List of evaporation module parameters for Thornthwaite equation in the CEQUEAU model

No.	Parameter	Unit	Lower Bound	Upper Bound	Description
1	XAA	dimensionless	0	5	Exponent in the Thornthwaite formula.
2	XIT	dimensionless	5	40	Value of Thornthwaite's thermal index.

Table IV-D3: List of thermal calibration parameters for CEQUEAU

No.	Parameter	Unit	Lower Bound	Upper Bound	Description
1	CRAYSO	dimensionless	0.05	2	Solar radiation coefficient. The coefficient that allows the increase or decrease level of solar radiation in the energy balance.
2	CRAYIN	dimensionless	-1	2	Infrared radiation coefficient. The coefficient that allows the increase or decrease level of infrared radiation in the energy balance.
3	CEVAPO	dimensionless	0.05	2	Evaporation coefficient. The coefficient that allows the increase or decrease level of evaporation in the energy balance.
4	CCONVE	dimensionless	0.05	2	Sensible heat coefficient. The coefficient that allows the increase or decrease level of sensible heat in the energy balance.
5	CRIGEL	mm	0	1000	Threshold snow stock for controlling water temperature
6	TNAP	°C	5	10	Temperature of groundwater
7	BASSOL	mm	5	20	Total precipitation depth used to detect days with low solar radiation.
8	CORSOL	dimensionless	0	1	Correction factor for bassol

Appendix IV-E: Simulated water temperature time series for DLI models at all stations using multisite calibration parameters

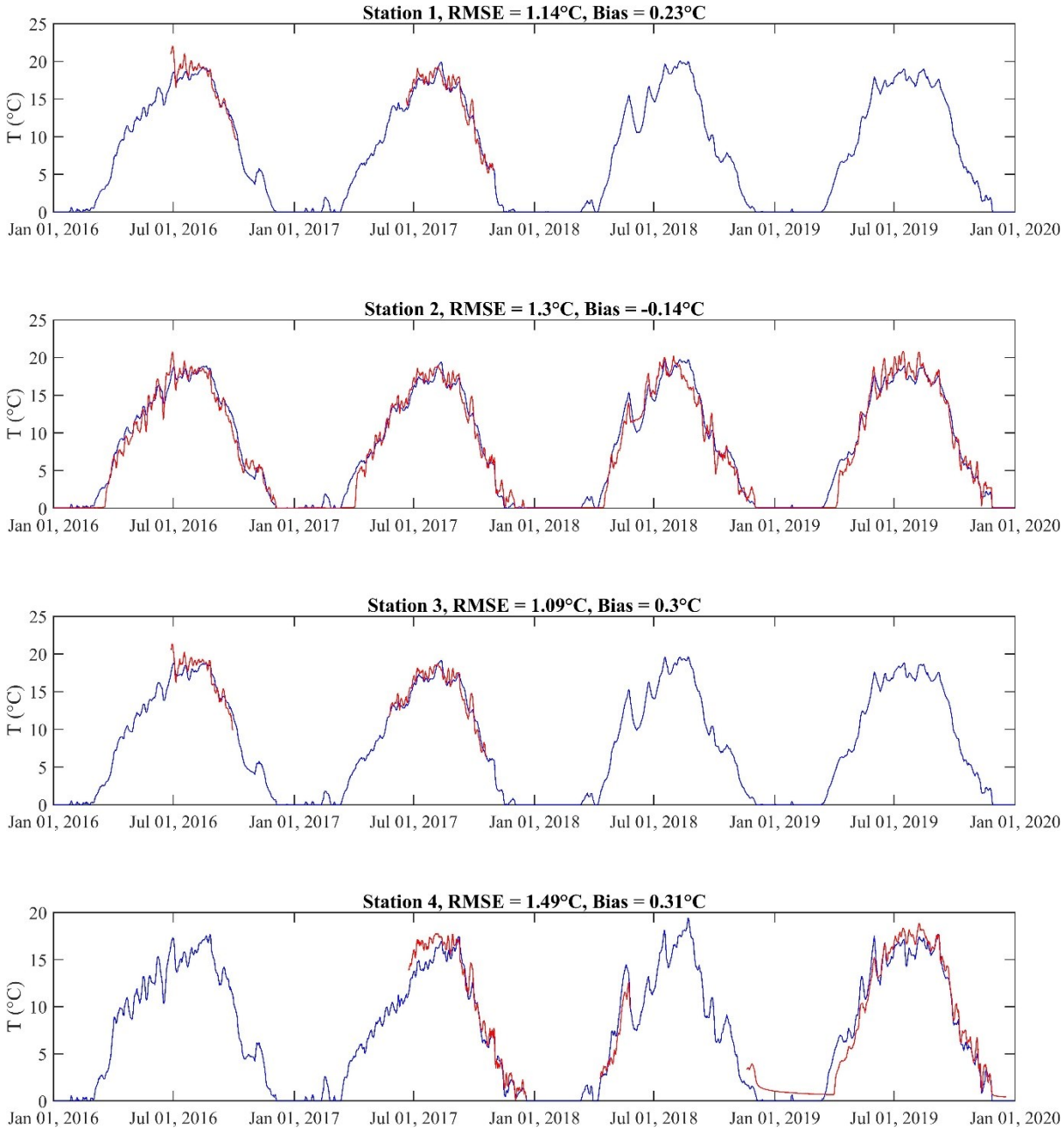

Figure IV-E1: The simulated time series for all stations using the multisite calibration parameters from Model 1 for DLI

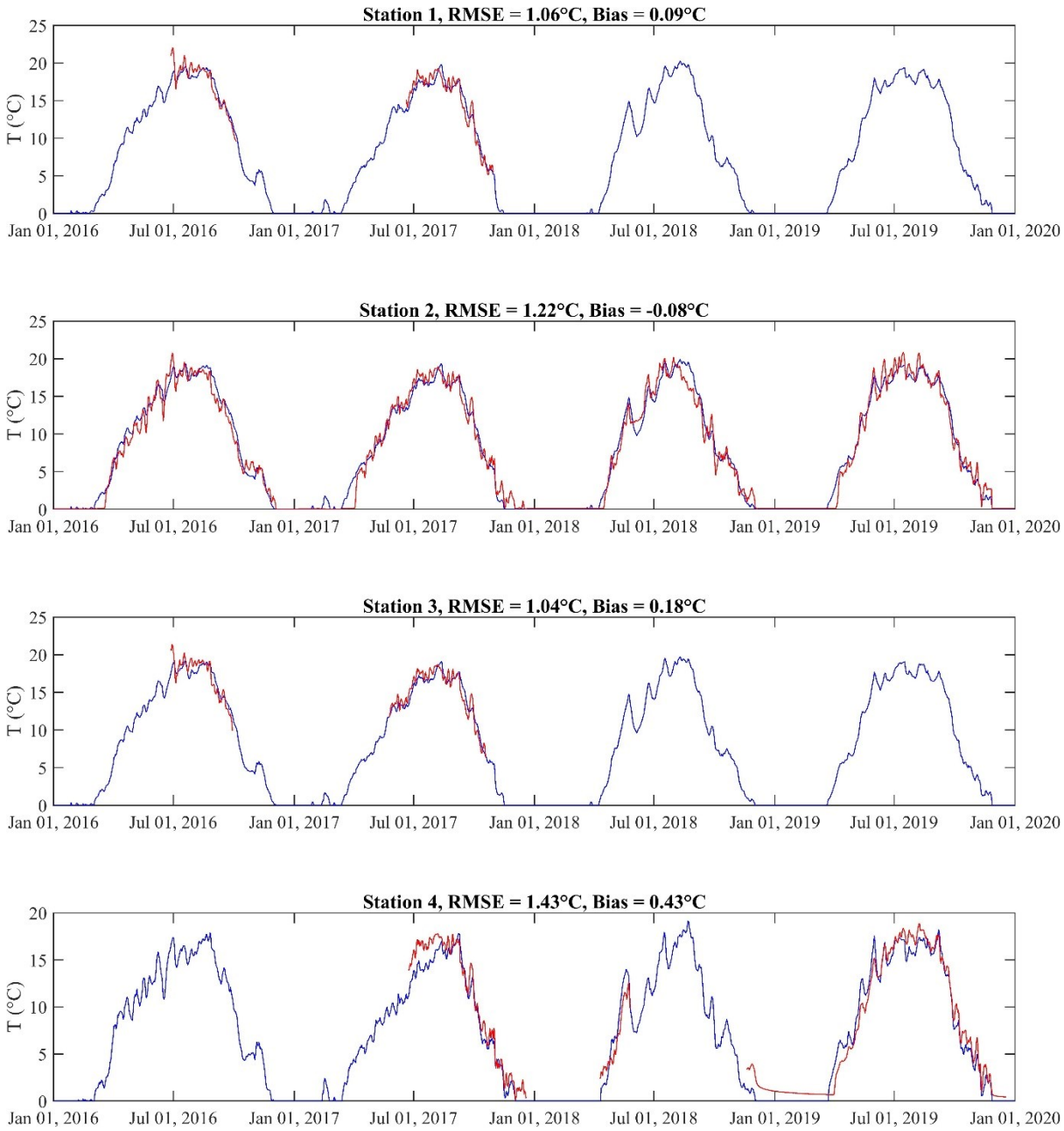

Figure IV-E2: The simulated time series for all stations using the multisite calibration parameters from Model 2 for DLI

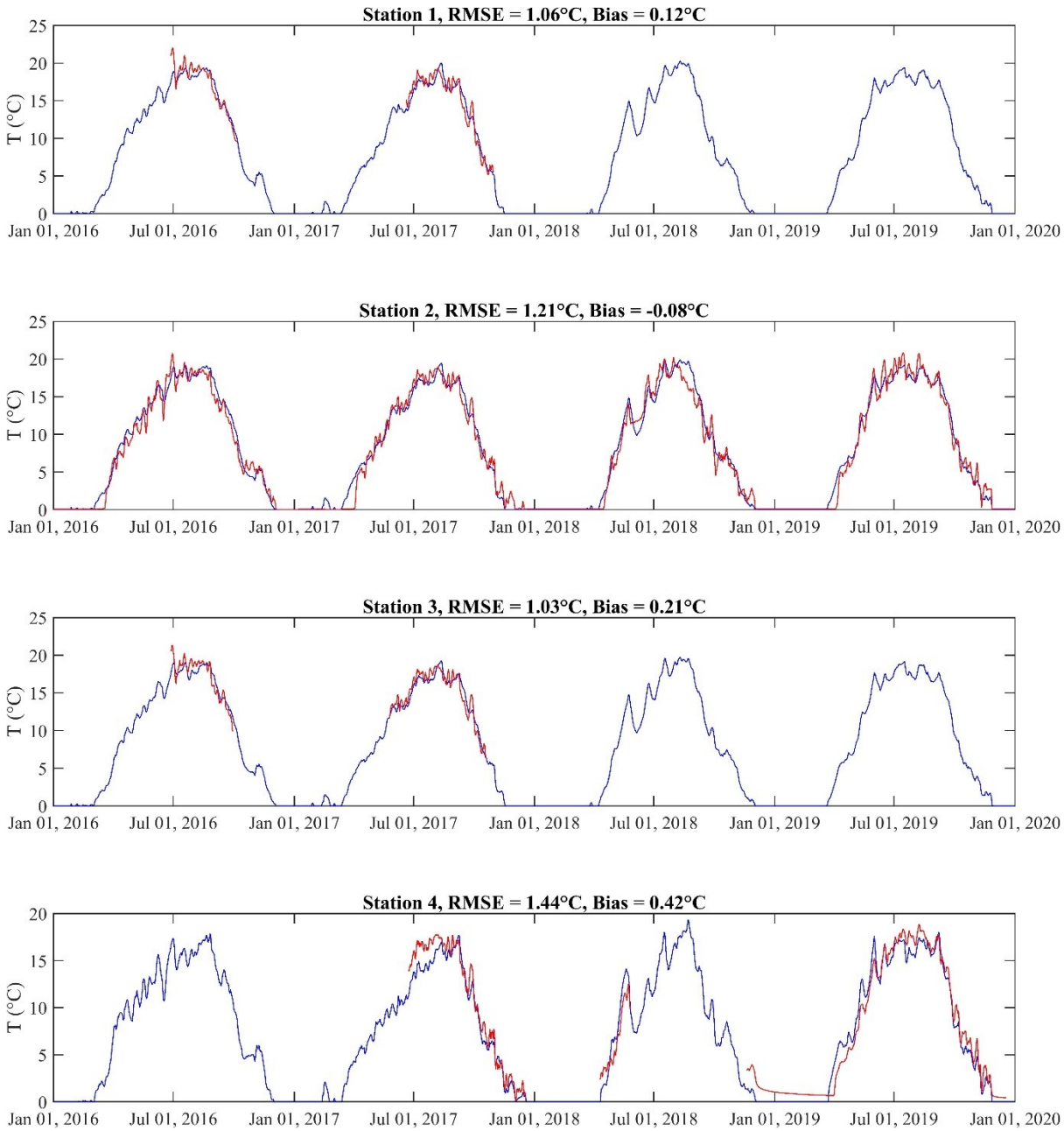

Figure IV-E3: The simulated time series for all stations using the multisite calibration parameters from Model 3 for DLI

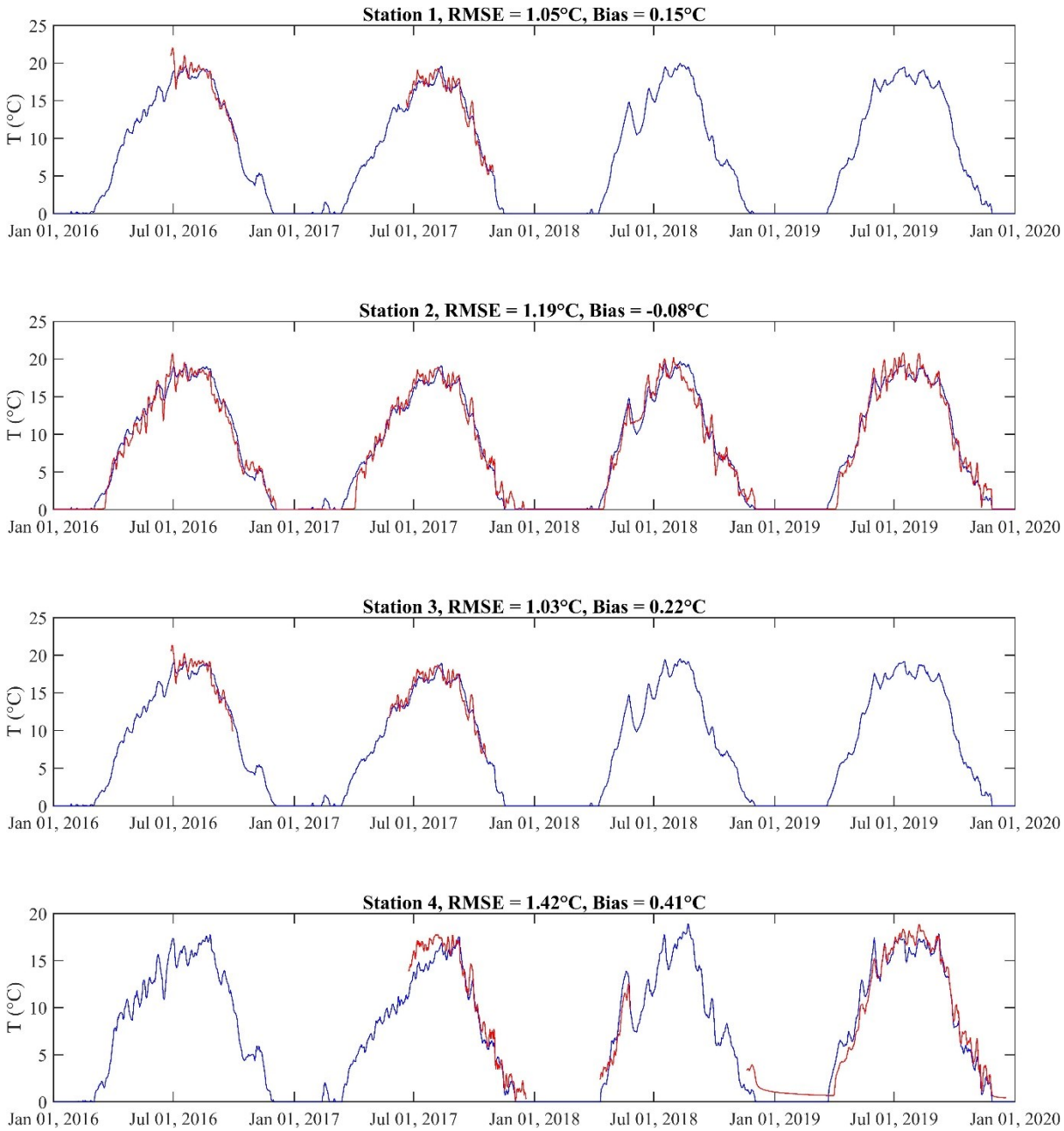

Figure IV-E4: The simulated time series for all stations using the multisite calibration parameters from Model 4 for DLI

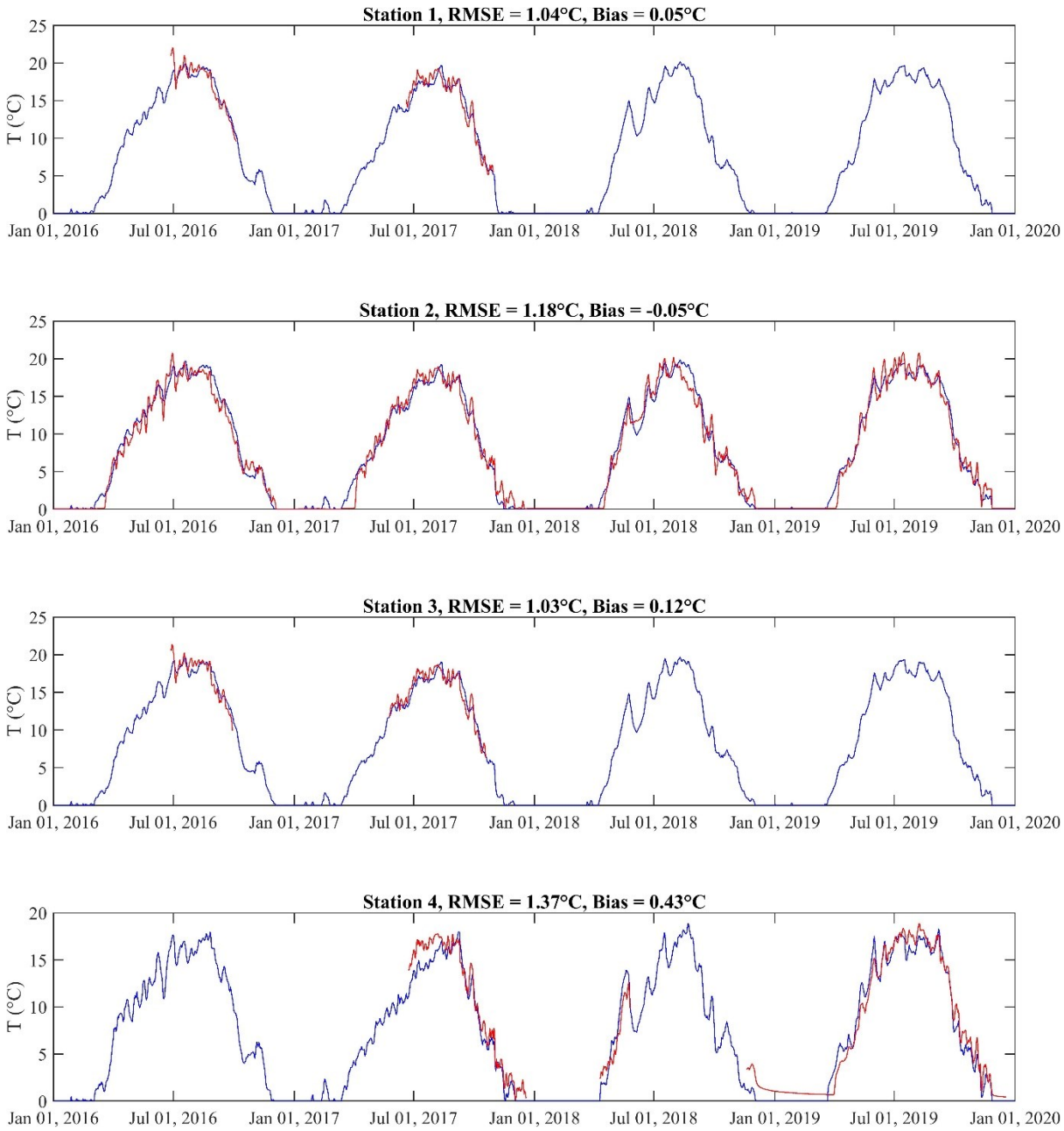

Figure IV-E5: The simulated time series for all stations using the multisite calibration parameters from Model 5 for DLI

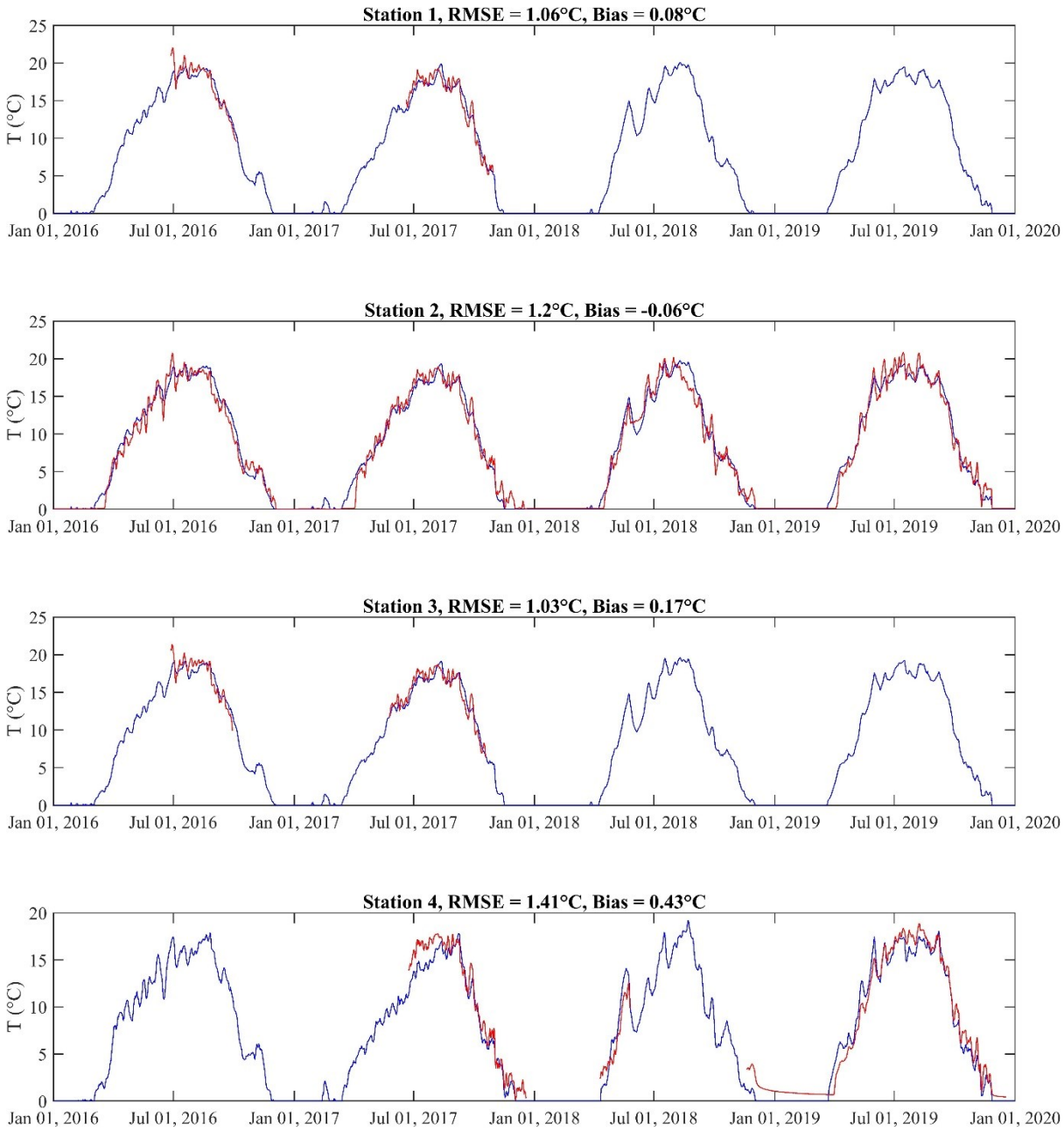

Figure IV-E6: The simulated time series for all stations using the multisite calibration parameters from Model 6 for DLI

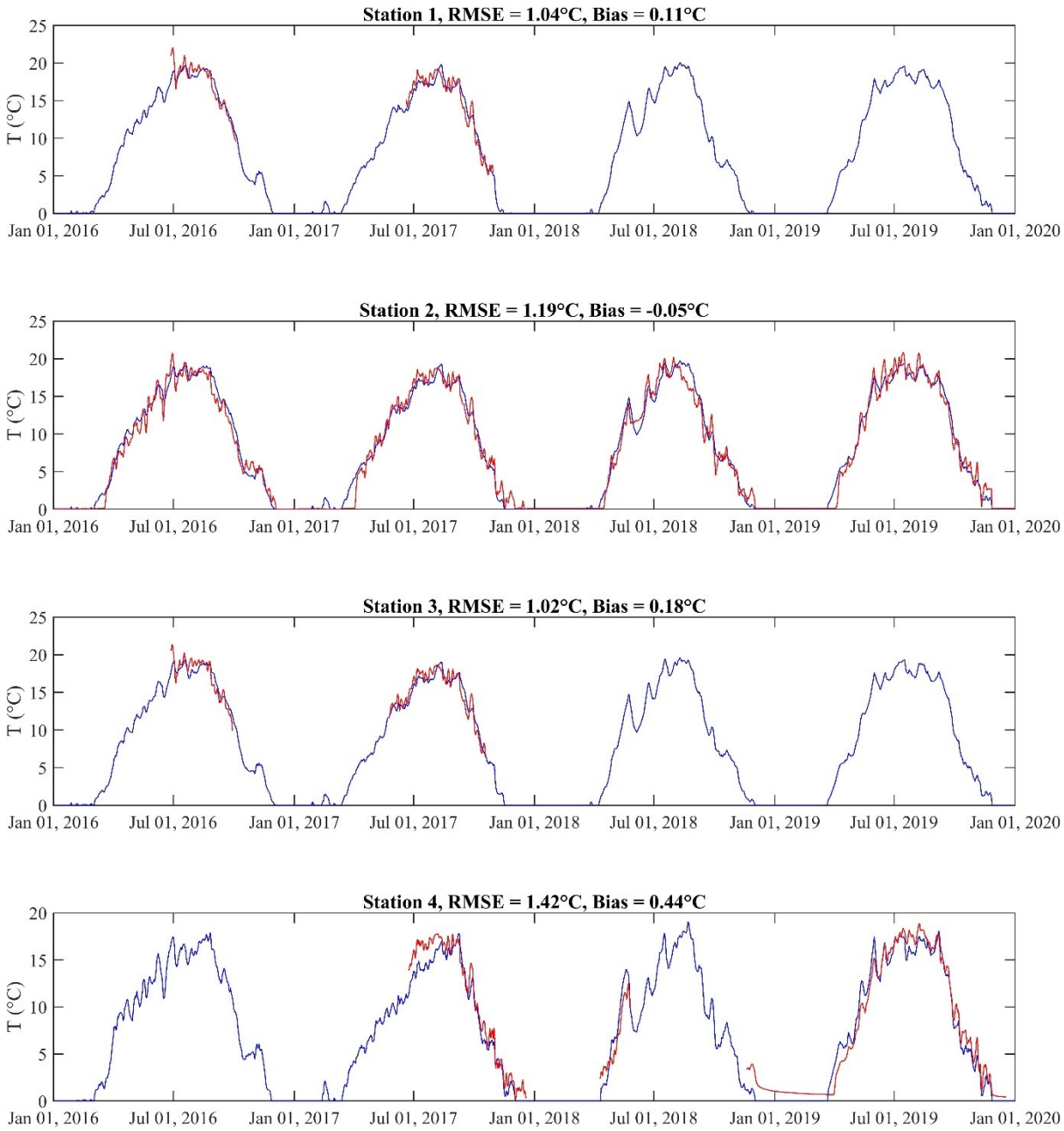

Figure IV-E7: The simulated time series for all stations using the multisite calibration parameters from Model 7 for DLI

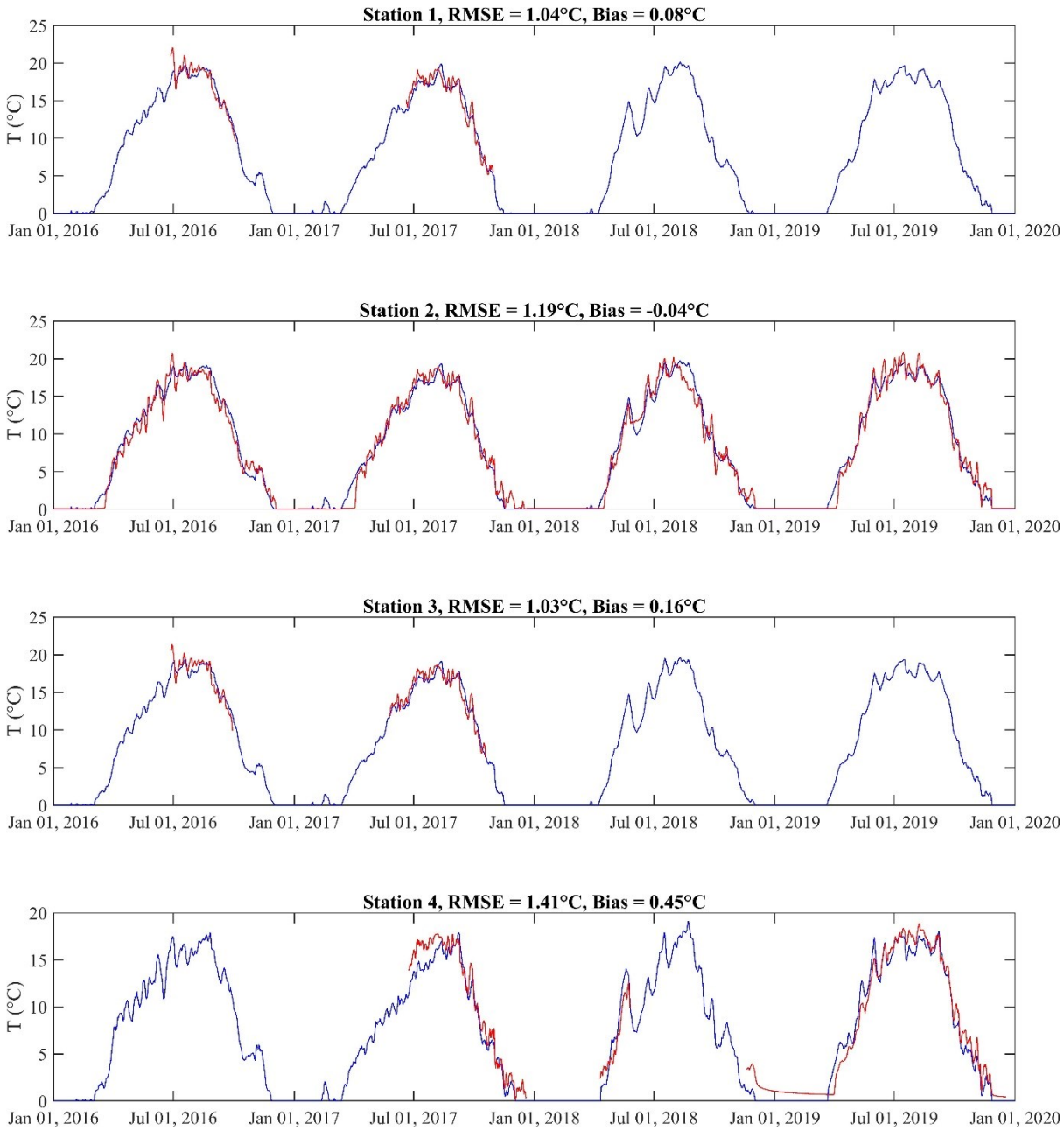

Figure IV-E8: The simulated time series for all stations using the multisite calibration parameters from Model 8 for DLI

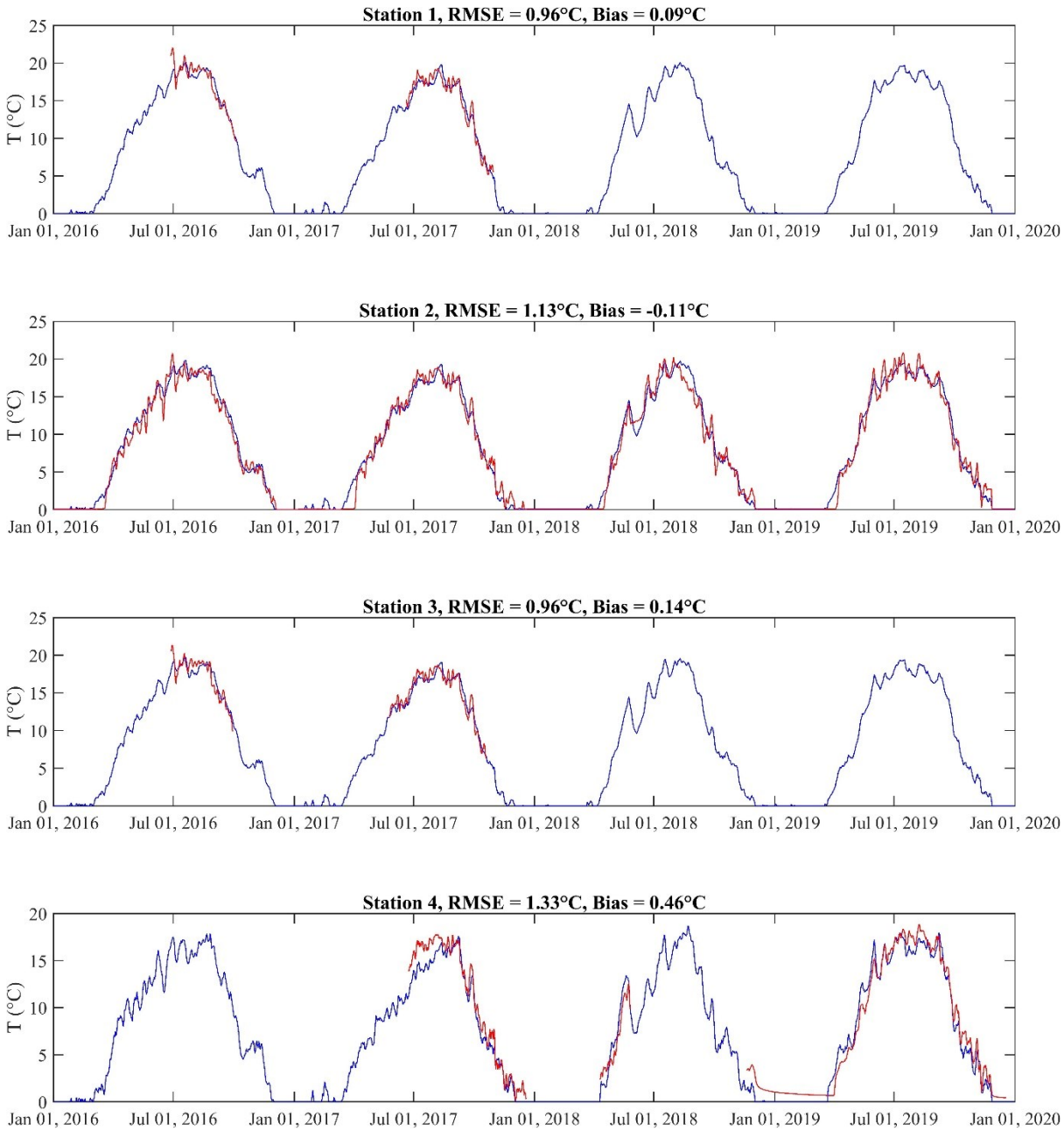

Figure IV-E9: The simulated time series for all stations using the multisite calibration parameters from Model 9 for DLI

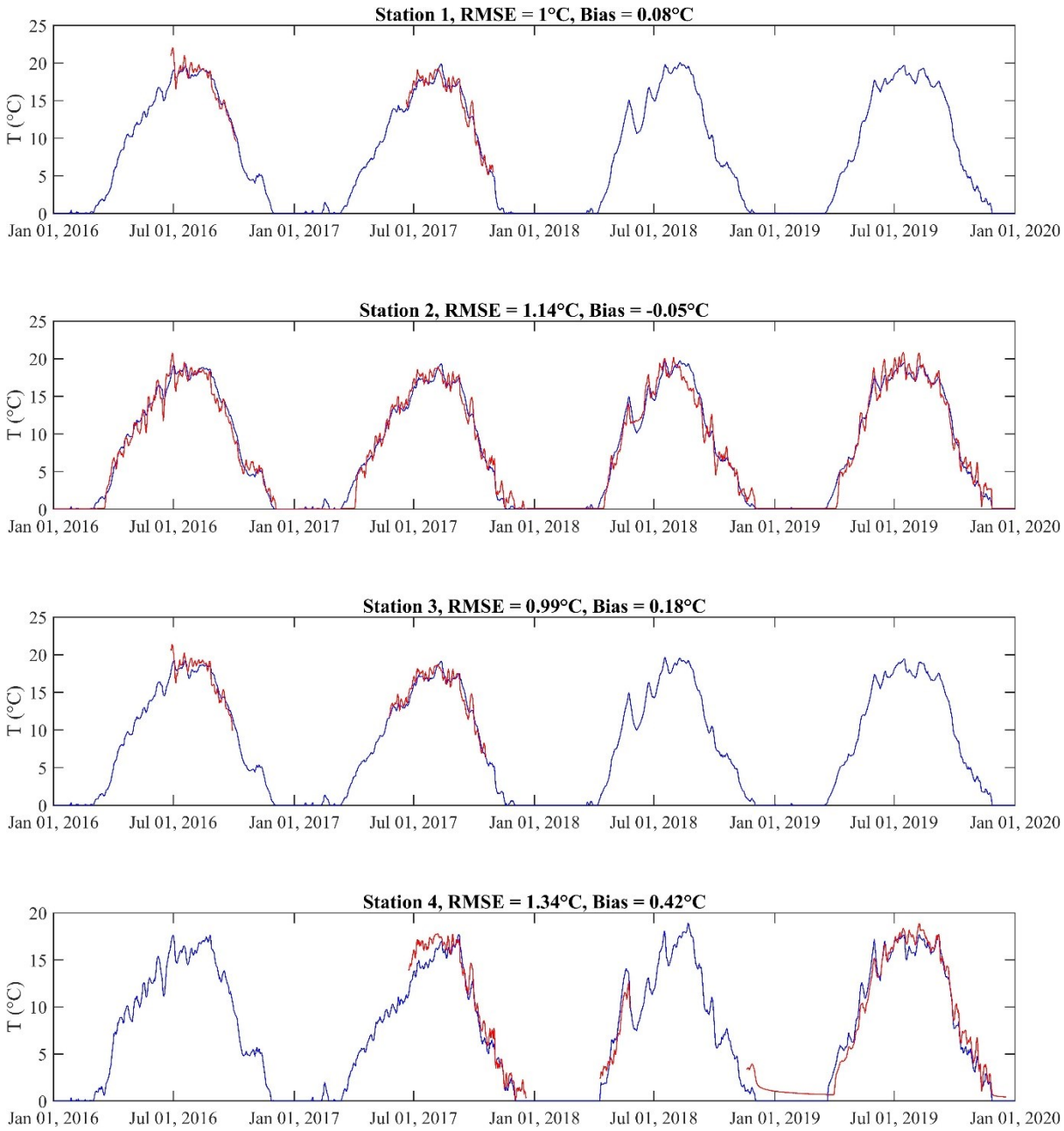

Figure IV-E9: The simulated time series for all stations using the multisite calibration parameters from the ERA5 dataset for DLI

Appendix IV-F: The heat budget contributions for Vanderhoof station (CP = 23)

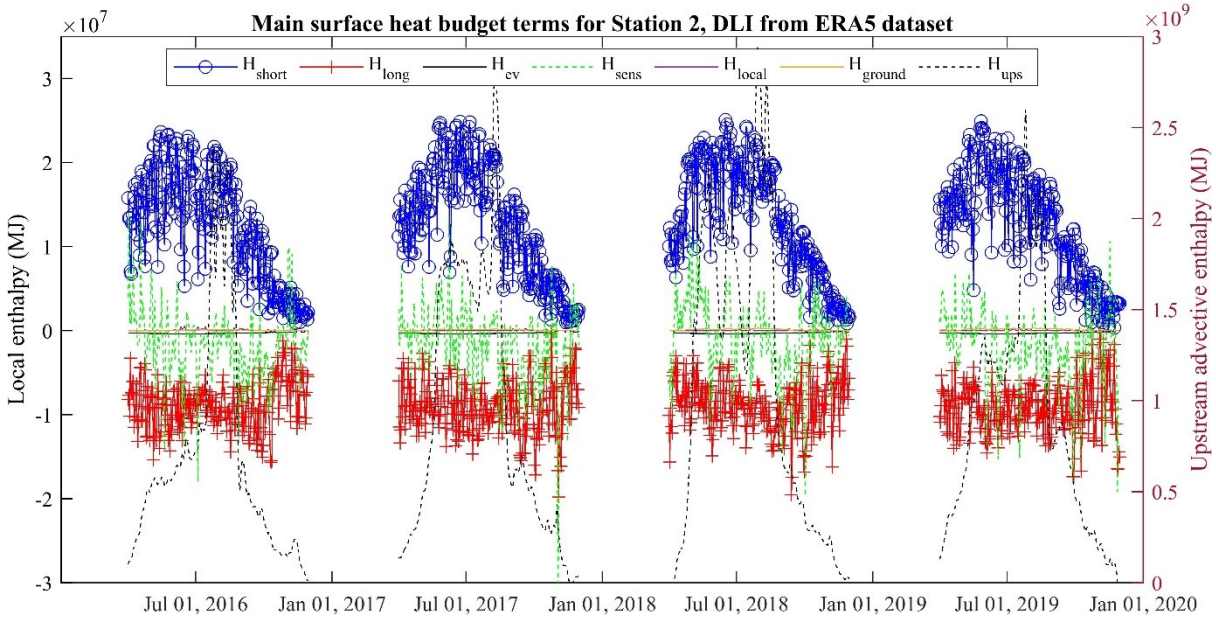

Figure IV-F1: The simulated heat budget time series for Station 2 (Vanderhoof) using the multisite calibration parameters from using the ERA5 dataset for DLI

Appendix IV-G: The C++ codes implemented in Article IV

```
#include "stdafx.h"
#include "mex.h"
#include "DLI.h"
// Constructeur.
DLI::DLI() {
    FILE_LOG(logDEBUG) << "DLI(int moduleDLI, const mxArray* structDonneesMeteo)";
    moduleDLI_ = 1;
    params_.u = 0;
    params_.v = 0;
    params_.a = 0;
    params_.b = 0;
    params_.c = 0;
    params_.d = 0;
}

// Get
int DLI::moduleDLI() {
    return moduleDLI_;
}

// Sélectionnez la fonction de calcul correcte
float DLI::calculeRadiation(float sigma, float ENNUAG, float TKR, float PREVAP, float TKO,
float LAI, float DLI) {
    float B = ENNUAG;
    float ea = PREVAP;
    float RADIN = 0;
    float ATMOS = 0;

    switch (moduleDLI_) {
        case 0:
            ATMOS = calculeM4(B, ea);
            break;
        case 1:
            ATMOS = calculeM1(B);
            break;
        case 2:
            ATMOS = calculeM2(B, TKR);
            break;
        case 3:
            ATMOS = calculeM3(B, TKR);
            break;
        case 4:
            ATMOS = calculeM4(B, ea);
            break;
        case 5:
            ATMOS = calculeM5(B, TKR, ea);
            break;
        case 6:
            ATMOS = calculeM6(B, TKR, ea);
            break;
        case 7:
            ATMOS = calculeM7(B, TKR, ea);
            break;
        case 8:
            ATMOS = calculeM8(B, TKR, ea);
            break;
        case 9:
            ATMOS = calculeM9(B, ea, LAI);
            break;
        case 10:
            RADIN = DLI - 0.97f * sigma * pow(TKO, 4.0f);
            return RADIN;
            break;
        default:
            break;
    }

    RADIN = sigma * ATMOS * pow(TKR, 4.0f) - 0.97f * sigma * pow(TKO, 4.0f);
    return RADIN;
}

float DLI::calculeM1(float B) {
    return (1.0f + params_.u * pow(B, params_.v)) * params_.a;
}
```

```
float DLI::calculeM2(float B, float TKR) {
    return (1.f + params_.u * pow(B, params_.v)) * (params_.a * pow(TKR, 2.0f));
}

float DLI::calculeM3(float B, float TKR) {
    return (1.0f + params_.u * pow(B, params_.v)) * (1.0f - params_.a * exp(params_.b *
pow(273.0f - TKR, 2.0f)));
}

float DLI::calculeM4(float B, float ea) {
    return (1.0f + params_.u * pow(B, params_.v)) * 0.97f * (params_.a + params_.b * ea);
}

float DLI::calculeM5(float B, float TKR, float ea) {
    return (1.0f + params_.u * pow(B, params_.v)) * (params_.a * pow(1.3332239f * ea / TKR,
params_.b));
}

float DLI::calculeM6(float B, float TKR, float ea) {
    return (1.0f + params_.u * pow(B, params_.v)) * (params_.a * (1.0f - exp(-1.0f *
pow(1.3332239f * ea, TKR / params_.b))));
}

float DLI::calculeM7(float B, float TKR, float ea) {
    float temp = (-1.0f) * pow(params_.b + params_.c * params_.a * (1.3332239f * ea / TKR),
params_.d);
    return (1.0f + params_.u * pow(B, params_.v)) * (1.0f - (1.0f + params_.a * (1.3332239f *
ea / TKR)) * exp(temp));
}

float DLI::calculeM8(float B, float TKR, float ea) {
    if (1.3332239f * ea >= 2.0f) {
        return (1.0f + params_.u * pow(B, params_.v)) * (params_.a + params_.b * (1.3332239f *
ea - 2.0f));
    }
    else {
        return (1.0f + params_.u * pow(B, params_.v)) * (params_.a - params_.c * (1.3332239f *
ea - 2.0f));
    }
}

float DLI::calculeM9(float B, float ea, float LAI) {
    return (1.0f + params_.u * pow(B, params_.v)) * exp(-1 * params_.c * LAI) * 0.97f *
(params_.a + params_.b * ea) + 0.98 * (1.0f - exp(-1.0f * params_.c * LAI));
}

// Initialisation
void DLI::initialiser(int moduleDLI, const mxArray* structParamDLI) {
    mxArray* m;

    moduleDLI_ = moduleDLI;

    switch (moduleDLI_) {
    case 0:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m4");
        lireParametres(m, 4);
        break;
    case 1:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m1");
        lireParametres(m, 3);
        break;
    case 2:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m2");
        lireParametres(m, 3);
        break;
    case 3:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m3");
        lireParametres(m, 4);
        break;
    case 4:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m4");
        lireParametres(m, 4);
        break;
    case 5:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m5");
        lireParametres(m, 4);
```

```
        break;
    case 6:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m6");
        lireParametres(m, 4);
        break;
    case 7:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m7");
        lireParametres(m, 6);
        break;
    case 8:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m8");
        lireParametres(m, 5);
        break;
    case 9:
        m = MatlabHelper::mhMxGetField(structParamDLI, 0, "m9");
        lireParametres(m, 5);
    default:
        break;
    }
}

// Lire les parametres pour l'equation
void DLI::lireParametres(const mxArray* paramsDLI, int numParams) {
    FILE_LOG(logINFO) << "numParams " << numParams;

    if (numParams >= 1) {
        lireParametresHelper(paramsDLI, "u", params_.u);
    }
    FILE_LOG(logINFO) << "params_.u " << params_.u;

    if (numParams >= 2) {
        lireParametresHelper(paramsDLI, "v", params_.v);
    }
    FILE_LOG(logINFO) << "params_.v " << params_.v;

    if (numParams >= 3) {
        lireParametresHelper(paramsDLI, "a", params_.a);
    }
    FILE_LOG(logINFO) << "params_.a " << params_.a;

    if (numParams >= 4) {
        lireParametresHelper(paramsDLI, "b", params_.b);
    }
    FILE_LOG(logINFO) << "params_.b " << params_.b;

    if (numParams >= 5) {
        lireParametresHelper(paramsDLI, "c", params_.c);
    }
    FILE_LOG(logINFO) << "params_.c " << params_.c;

    if (numParams >= 6) {
        lireParametresHelper(paramsDLI, "d", params_.d);
    }
    FILE_LOG(logINFO) << "params_.d " << params_.d;;

}
```