

Université du Québec

Institut National de la Recherche Scientifique

Centre Eau Terre Environnement

**Estimation de l'aire des refuges thermiques potentiels en rivière par
l'utilisation du Modèle additif généralisé (GAM) et le Modèle de régression
multivariée par splines adaptatifs (MARS): Étude de cas de la rivière Ste-
Marguerite (Québec, Canada)**

Par

Al Mahdi Saadi

Mémoire présenté pour l'obtention du grade de

Maître ès en Sciences (M.Sc.)

Jury d'évaluation

| | |
|---|--|
| Président du jury et examinateur interne | Normand Bergeron INRS-Eau-Terre-Environnement |
| Examinateur externe | Marc Barbet Hydro-Québec |
| Directeur de recherche | André St-Hilaire INRS-Eau-Terre-Environnement |
| Directeur de recherche | Jamal Chao ISHÉDD/INRS-ETE |
| Codirecteur de recherche | Taha B.M.J. Ouarda INRS-Eau-Terre-Environnement |

REMERCIEMENTS

Tout d'abord, j'aimerais remercier l'ensemble des personnes ayant de près ou de loin contribué à l'élaboration, la concrétisation et la réalisation de ce projet. Aux termes de ces deux années de maîtrise, je tiens à exprimer ma profonde gratitude à mon directeur de recherche André St-Hilaire, Jamal Chao et mon Codirecteur de recherche Taha Ouarda pour m'avoir fait confiance et accepté dans leur équipe et pour avoir dirigé minutieusement et rigoureusement mes travaux de recherche. Je tiens aussi à les remercier également pour leur aide inestimable, leur excellent encadrement ainsi que leurs précieux conseils non seulement professionnels, mais aussi personnels le long de mes travaux.

Mes remerciements aux membres du jury d'évaluation à savoir M. Normand Bergeron et M. Marc Barbet pour leurs disponibilités, leurs commentaires et suggestions afin d'améliorer la qualité du présent manuscrit.

Un merci tout particulier au personnel administratif de l'ISHÉDD et à sa tête monsieur Kamal El Haji ainsi que le corps professoral, de m'avoir préparé à atteindre ce stade.

J'exprime toute ma gratitude envers monsieur André Boivin pour sa disponibilité permanente, son soutien, aussi précieux que volontaire le long de la période des travaux sur le terrain. Mes remerciements au Centre Inter-universitaire de Recherche sur le Saumon Atlantique (CIRSA) pour m'avoir fourni la possibilité de faire mes recherches et travaux de terrain de manière adéquate grâce à leurs installations au sein de la zone d'étude. Ma reconnaissance va aussi aux membres de l'équipe monsieur Christian Charron et madame Amina Msilini pour leur soutien et leurs conseils volontaires.

Je remercie spécialement ma mère Fatima Elaydi Saadi et à mon père Omar Madani Saadi pour m'avoir appris à atteindre mes objectifs de vie et soutenu moralement afin de surmonter les nombreuses épreuves le long de ce parcours, mais aussi ma femme ainsi que mes frères et sœurs pour leur soutien moral et leurs encouragements.

RÉSUMÉ

La température élevée de l'eau peut être mortelle pour de nombreuses espèces de poissons et plus particulièrement le saumon atlantique (*Salmo salar*) qui peut rechercher des refuges thermiques pour se protéger lors d'événements thermiques stressants. Dans cette étude, la superficie de chaque refuge thermique a été interpolée au pas de temps journaliers dans quatre stations différentes situées à la confluence de tributaires froids avec la rivière Ste-Marguerite (Québec, Canada). Deux approches de modélisation statistique ont été utilisées dans le but d'estimer la superficie des refuges thermiques potentiels, à savoir le modèle additif généralisé (MAG) et le modèle de régression multivariée par splines adaptatifs (MARS). Les modèles MAG et MARS ont été ajustés indépendamment pour chaque station d'étude. Les performances des deux modèles ont été évaluées à l'aide de la validation croisée (CV) afin de mettre en évidence des problèmes tels que le surajustement ou le biais de sélection et de donner un aperçu de la manière dont le modèle devrait se généraliser à un ensemble de données indépendants. Plusieurs critères ont été utilisées pour comparer les performances du MAG et du MARS, tels que la racine de l'erreur quadratique moyenne (RMSE), le coefficient de NASH, le critère d'information d'Aikaike (AIC) et finalement le biais (BIAS). Les résultats montrent que les aires modélisées sont proches des aires estimées avec un RMSE relatif variant entre 6,93 % et 38,39 %. En outre, les résultats attestent que le modèle MARS offre une meilleure performance dans la prévision et l'estimation de la variabilité de l'aire de refuge thermique que le MAG dans toutes les stations d'études. Cependant, compte tenu de ces résultats, les deux modèles présentent un grand potentiel.

Mots-clés : Saumon atlantique, refuges thermiques, température journalière de l'eau, model additif généralisé (GAM), le modèle de régression multivariée par splines adaptatifs (MARS).

TABLE DES MATIÈRES

| | |
|---|-----------|
| RÉSUMÉ..... | VII |
| LISTE DES FIGURES..... | XI |
| 1. INTRODUCTION..... | 2 |
| 1.1 MISE EN CONTEXTE | 3 |
| 1.2 VARIABLES INFLUENÇANT LA TEMPERATURE DE L'EAU | 6 |
| 1.3 MODELISATION DE LA TEMPERATURE DE L'EAU | 7 |
| 1.4 OBJECTIFS DE L'ETUDE..... | 9 |
| 1.5 STRUCTURE DU MÉMOIRE..... | 10 |
| 2. SYNTHÈSE DE LA MÉTHODOLOGIE ET DES RÉSULTATS | 12 |
| 1.6 METHODOLOGIE ET SITE D'ETUDE | 13 |
| 1.7 APPROCHES STATISTIQUES | 14 |
| 1.7.1 <i>Modèle additif généralisé</i> | 14 |
| 1.7.2 <i>Modèle de régression multivariée par splines adaptatifs (MARS)</i> | 15 |
| 1.7.3 <i>Evaluation des performances des modèles</i> | 16 |
| 1.8 RESULTATS ET DISCUSSIONS..... | 17 |
| 1.8.1 <i>Variabilité spatio-temporelle</i> | 19 |
| 1.8.2 <i>Comparaison des performances entre GAM et MARS</i> : | 20 |
| 1.9 CONCLUSION ET RECOMMANDATIONS | 22 |
| 3. ARTICLE | 24 |
| 4. RÉFÉRENCES..... | 37 |

LISTE DES FIGURES

| | |
|--|----|
| FIGURE 1 : SAINTE-MARGUERITE RIVER DRAINAGE BASIN WITH HYDROMETRIC, WEATHER AND MEASUREMENT STATIONS LOCATION | 8 |
| FIGURE 2 : EXAMPLE OF THE DEPLOYMENT OF THERMOGRAPHS STATION 1, STE-MARGUERITE RIVER... | 10 |
| FIGURE 3 : TIME SERIES OF POTENTIAL PREDICTORS AND REFUGE AREA IN STATION 1, 2018. | 16 |
| FIGURE 4 : TIME SERIES OF POTENTIAL PREDICTORS AND REFUGE AREA IN STATION 2, 2018. | 16 |
| FIGURE 5 : TIME SERIES OF POTENTIAL PREDICTORS AND REFUGE AREA IN STATION 3, 2018. | 17 |
| FIGURE 6 : TIME SERIES OF POTENTIAL PREDICTORS AND REFUGE AREA IN STATION 2, 2018. | 17 |
| FIGURE 7 : ESTIMATED AREA IN STATION 1 (JULY 2, 2018 AT 12:00PM)..... | 18 |
| FIGURE 8 : ESTIMATED AREA IN STATION 1 (JULY 2, 2018 AT 10:00PM)..... | 19 |
| FIGURE 9 : ESTIMATED AREA IN STATION 1 (AUGUST 2, 2018)..... | 20 |
| FIGURE 10 : ESTIMATED AREA IN STATION 1 (SEPTEMBER 5, 2018). | 21 |
| FIGURE 11 : ESTIMATED VS MODELED AREA FOR THE BEST MODEL IN STATION A) 1 IN 2018, B) 2 IN 2018, C) 3 IN 2018 AND D) 4 IN 2019. | 23 |

1. INTRODUCTION

1.1 Mise en contexte

Les régimes thermiques des rivières sont d'une grande importance non seulement pour le taux de croissance, mais aussi pour la distribution des espèces d'eau douce (Ducharne, 2007; Haag & Westrich, 2002; Ozaki et al., 2003). Il est donc important que le régime thermique des rivières soit bien étudié pour une gestion efficace de la pêche ainsi qu'une évaluation de l'impact environnemental de certains projets d'aménagement en rivière. Les variations de la température de l'eau peuvent se produire naturellement ou bien à cause d'impacts anthropiques tels que les changements climatiques, la déforestation, les barrages ou même la pollution thermique. De plus, on s'attend à ce que les changements futurs de la température en été accentuent l'intensité et l'occurrence des extrêmes thermiques (Arismendi et al., 2013), ce qui aura un impact majeur sur la structure de la population des espèces d'eau froide (Mantua et al., 2010). Certaines espèces aquatiques en rivière sont affectées par les températures élevées qui peuvent entraîner un stress thermique. C'est le cas du saumon juvénile (Isaak et al., 2012). Ce dernier tend à chercher des refuges thermiques lorsque la température de la rivière dépasse 21 °C afin d'éviter l'exposition à ces conditions stressantes (Baird & Krueger, 2003; Daigle et al., 2015; Dugdale et al., 2013; Welsh Jr et al., 2001). La dégradation des habitats due à la hausse de température (Ruesch et al., 2012) est une observation extrêmement significative pour les gestionnaires qui cherchent à atténuer l'impact négatif sur des espèces pêchées comme le saumon atlantique (Battin et al., 2007).

D'après l'étude d'imagerie thermique par infrarouge faite par Dugdale et al. (2013), les refuges thermiques ont été regroupés en sept classes. On retrouve le panache de tributaire (tributary confluence plume) (Torgersen et al., 2012), suintement latéral (lateral seep) (Bilby, 1984; Ebersole et al., 2003), ruisseau d'eau de source (springbrook) (Ebersole et al., 2003; Stanford & Ward,

1993), chenal latéral froid (cold side channel) (Ebersole et al., 2003; Stevens & DuPont, 2011), alcôve froide (cold alcove) (Ebersole et al., 2003; Ozaki, 1987), résurgence hyporhéique (hyporheic upwelling) (Brunke & Gonser, 1997; Burkholder et al., 2008; Poole & Berman, 2001) et finalement l'effluent de versant de vallée (wall-base channel) (Peterson & Reid, 1984; Torgersen et al., 2012). Le type de refuge thermique le plus évident et potentiellement le plus fréquent sur lequel se base la présente étude est associée à la confluence d'un tributaire froid et d'une rivière plus chaude où un panache d'eau froide sera alors créé en aval de la confluence.

À ce jour, il n'existe pas de définition générale d'un refuge thermique qui fasse l'objet d'un accord commun, car cette dernière dépend des conditions de températures pour chaque espèce, mais aussi des conditions locales de la zone d'étude. Par exemple, dans l'étude faite par Ebersole et al. (2001) afin de définir un refuge thermique pour la truite arc-en-ciel (*Oncorhynchus mykiss*), une différence de température de 3 °C a été utilisée entre le refuge et le cours d'eau principal. Sutton et al. (2007) ont décrit un refuge thermique comme étant une eau plus froide que 22-23 °C pour plusieurs espèces de salmonidés sur la côte ouest des États-Unis. Wang et al. (2020) ont utilisé une différence de température de 2 °C entre la température du cours d'eau principal et le panache d'eau froide ainsi qu'un seuil de 21 °C afin de définir un refuge pour la truite steelhead (*Oncorhynchus mykiss*). Breau et al. (2007) ont constaté, d'après les observations effectuées, des agrégations de saumons juvéniles dans des refuges qui ont été définis comme ayant des températures inférieures à 23 °C, tandis que Kurylyk et al. (2015) ont expliqué que les tacons sont connus pour utiliser des refuges caractérisés par une différence de température de 2 °C avec le cours d'eau principal. Dans la présente étude, nous déterminerons les limites du panache d'eau froide en utilisant une différence de température de 1 °C avec le cours d'eau principal. Ce critère est subjectif et a été

choisi pour s'assurer que la différence de température est supérieure à la précision des thermographes (≤ 0.5 °C) qui ont été déployés dans la rivière et ses affluents.

Étant donné la rareté relative des stations d'observations de la température de l'eau des rivières à saumon atlantique au Canada, des outils de modélisation sont nécessaires pour fournir des informations thermiques aux gestionnaires. Deux types de modèles sont généralement utilisés pour modéliser et prévoir la température de l'eau sur différentes échelles : les modèles déterministes et statistiques (Benyahya et al., 2007; Caissie, 2006; Cole et al., 2014; Marceau et al., 1986). Les modèles déterministes sont basés sur la représentation physique des flux de chaleur entre la rivière et son environnement. Cela exige de nombreuses variables en intrants (géométrie du cours d'eau, hydrologie et météorologie) qui ne sont pas toujours facilement disponibles (BOARD, 2003; Caissie et al., 2007; Cox & Bolte, 2007; Edinger et al., 1968; Mohseni et al., 1998; St-Hilaire et al., 2000). Les modèles statistiques nécessitent moins de variables que les modèles physiques, mais exigent habituellement de longues séries de mesures dans le temps (Caissie, 2006; DeWeber & Wagner, 2014; Marceau et al., 1986; Mohseni et al., 2003; Moore, 2006; Morrill et al., 2005; Smith, 1981; Webb et al., 2008). Les forces et les limites de plusieurs modèles statistiques ont été résumées par Benyahya et al. (2007). À priori, les modèles de régression paramétrique précisent le type de relation (linéaire ou non-linéaire). Ces derniers sont plus efficaces pour des pas de temps mensuels ou hebdomadaires que pour les plus hautes fréquences (Caissie, 2006; Erickson & Stefan, 2000; Morrill et al., 2005; Pilgrim et al., 1998). Les modèles non paramétriques ne précisent pas à priori le type de relation entre les prédicteurs et la variable dépendante. Ces approches incluent, les k voisins les plus proches (St-Hilaire et al., 2012), le modèle additif généralisé (Frechette et al., 2019) ou encore le modèle de régression multivariée par splines adaptatifs (Leathwick et al., 2005).

Peu d'études ont tenté de modéliser la superficie des refuges thermiques. La majorité des recherches effectuées ont modélisé la température en un point, c'est-à-dire en 0-D (Daigle et al., 2015; Jeong et al., 2013). Torgersen et al. (2012) ont étudié les préférences thermiques et les conditions favorables au saumon en Oregon (États-Unis). D'autres études se sont basées sur les approches hydrodynamiques qui sont la méthode la plus populaire pour modéliser les panaches thermiques (Decker et al., 2013; Lowe et al., 2009). Dzara et al. (2019) ont utilisé le modèle hydraulique 1-D ainsi qu'un modèle de la qualité de l'eau pour simuler des refuges d'eau froide sur la rivière Walker (Nevada, États-Unis). En plus de ces études mentionnées, d'autres se sont basées sur des approches statistiques afin d'étudier la présence et la distribution des poissons dans les refuges thermiques (par exemple, (Wang et al., 2020)). Étonnamment, à notre connaissance, la présente étude est la première qui compare des modèles statistiques pour estimer les superficies de refuges thermiques en fonction de prédicteurs hydroclimatiques.

1.2 Variables influençant la température de l'eau

Le comportement de la température des rivières dépend des facteurs qui influencent significativement sa variation spatiale et temporelle incluant la dynamique des confluences des cours d'eau (Best, 1988; Mosley, 1976; Rhoads & Kenworthy, 1995). Plusieurs facteurs contrôlent directement ou indirectement les conditions thermiques de la rivière. Ils sont classifiés en quatre catégories : i) climats et conditions locales (Lowney, 2000; Mohseni et al., 1998; Sinokrot & Stefan, 1993; Van Vliet et al., 2011), ii) la géographie physique du bassin versant (Caissie, 2006; Chen et al., 1998; Isaak & Hubert, 2001; Mosley, 1983), iii) l'hydrologie et l'hydraulique de surface (par exemple, le débit) (Caissie, 2006; Gardner et al., 2003), iv) les interactions entre l'eau souterraine et l'eau de surface (Ward, 1985; Webb et al., 2008). La géographie physique du bassin

versant est en partie fonction du relief, mais aussi fonction de l'utilisation du territoire. Les facteurs qui composent ce groupe sont la géomorphologie et la topographie du bassin versant, le pourcentage de territoire occupé par la végétation, incluant la présence ou l'absence d'une bande riveraine (Isaak & Hubert, 2001; St-Hilaire et al., 2000), la situation géographique (latitude, longitude, azimut du cours d'eau) ainsi que la superficie du bassin versant. Les facteurs influençant la capacité des cours d'eau à réguler leur régime thermique dépendent de l'apport d'eau au sein du bassin versant sont par exemple la longueur et le degré de ramification du réseau hydrographique. Les variables météorologiques (température de l'air, précipitations, vent, humidité relative, nébulosité, etc.) sont les principaux intrants dans les calculs du bilan thermique lié aux échanges à l'interface eau-air. (Cole et al., 2014). Des variables hydrauliques jouent aussi un rôle important dans la thermie des rivières, comme le degré de turbulence, les dimensions des sections d'écoulement, mais la variable la plus fréquemment utilisée pour inclure indirectement ces facteurs dans un modèle est le débit (Caissie, 2006; Moatar & Gailhard, 2006; Webb et al., 2003). En plus des facteurs mentionnés, la température de l'eau est fortement affectée par les eaux souterraines. Cependant, en raison de la covariance spatio-temporelle entre les deux températures, la température de l'air reste l'une des variables les plus importantes et parmi les plus utilisées dans les modèles (Guillemette, 2009; Li et al., 2014; Ouellet, 2011; Wehrly et al., 2009).

1.3 Modélisation de la température de l'eau

Afin de mieux comprendre les variations spatio-temporelles de la température de l'eau, deux grandes catégories de modèles sont utilisées pour l'estimation et la prévision de ces variations (Benyahya et al., 2007; Caissie, 2006; Cole et al., 2014; Marceau et al., 1986): (i) les modèles déterministes (ii) et les modèles statistiques. Les modèles déterministes se basent sur le

calcul des échanges thermiques entre la rivière et son milieu (Caissie et al., 2007; Cox & Bolte, 2007; Edinger et al., 1968; St-Hilaire et al., 2000). Ce type de modèles paramétrise les processus physiques afin de minimiser le plus possible l'erreur entre les valeurs observées et prédictes. La majorité des modèles déterministes sont de types hydrauliques 1-D, c'est-à-dire qu'ils permettent de simuler la température moyennée sur la section transversale et sur toute la profondeur, de l'amont vers l'aval (Caissie et al., 2007; Chen et al., 1998; Morin et al., 1994; St-Hilaire et al., 2003). Cependant, ces modèles nécessitent souvent une grande quantité de données (par exemple, la température de l'air, la vitesse du vent, la radiation solaire, la géométrie du cours d'eau, ..., Etc.) (Mohseni et al., 1998; Risley et al., 2003). Une alternative à ces derniers reste l'utilisation de modèles statistiques ou empiriques (Ahmadi-Nedushan et al., 2006). Ces modèles ont l'avantage d'avoir recours à un minimum de paramètres et souvent un peu de covariables, avec un effort de développement souvent moins important que celui requis pour les modèles déterministes (Eaton & Scheller, 1996; Mohseni & Stefan, 1999). Malgré ce point positif, les modèles statistiques nécessitent en contrepartie de plus longues séries d'observations pour la calibration et la validation (Marceau et al., 1986) que la plupart des modèles déterministes. Ces modèles statistiques sont répartis en deux groupes : (1) les modèles paramétriques qui sont basés sur des structures statistiques prédéfinies. On retrouve plusieurs méthodes dans les modèles non paramétriques tels que (i) les méthodes de régression linéaires basées sur l'hypothèse de forte linéarité, qui peuvent utiliser une seule variable explicative (Caissie et al., 2001; Erickson & Stefan, 2000; Segura et al., 2015; Smith, 1981; Webb et al., 2003) ou plusieurs variables (Ahmadi-Nedushan et al., 2007; Webb et al., 2003; Webb & Nobilis, 2007) pour prédire la température de l'eau. (ii) les méthodes de régression non linéaires, parmi lesquelles on retrouve le modèle logistique qui a été le plus utilisé dans cette sous-catégorie et qui relie la température de l'air à celle de l'eau (Benyahya et

al., 2007; Caissie, 2006; Caissie et al., 2001; Mohseni & Stefan, 1999; Mohseni et al., 1998; Segura et al., 2015) ce qui permet un bon ajustement pour des données journalières, mais aussi fournit une meilleure explication sur l'effet d'hystérèse entre les températures de l'eau et de l'air (Laanaya, 2015). (2) Les modèles non paramétriques ne précisent pas le type de relation entre les prédicteurs et les valeurs prédites. On retrouve notamment plusieurs applications telles que l'approche des K voisins les plus proches (St-Hilaire et al., 2012), le modèle additif généralisé (GAM; Laanaya et al. (2017)), le modèle de régression multivariée par splines adaptatifs (MARS ; Leathwick et al. (2005); (Msilini et al., 2020)) ou encore les réseaux de neurones artificiels (Belanger et al., 2005; Chenard & Caissie, 2008).

1.4 Objectifs de l'étude

La problématique de cette étude consiste à caractériser dans un premier temps, la variation temporelle des superficies de refuge thermique à plusieurs endroits, pour ensuite les estimer à l'aide de modèles statistiques. Deux approches non paramétriques (GAM et MARS) sont testées afin de vérifier si elles peuvent prédire adéquatement et avec un nombre restreint de variables explicatives (voir section 1.8), la variation temporelle des superficies des refuges thermiques à un pas de temps journalier. En deuxième lieu, les deux modèles (GAM ou MARS) sont comparés afin de déterminer lequel a de meilleures performances en tenant compte de la parcimonie. Finalement, on cherche à déterminer si l'inclusion d'une variable initialement négligée, le débit du tributaire, améliore les performances des deux modèles statistiques.

L'approche générale consiste donc à utiliser deux modèles statistiques qui sont le modèle additif généralisé (GAM) et le modèle de régression multivariée par splines adaptatifs (MARS) qui permettent de modéliser les variations spatio-temporelles de la superficie des refuges thermiques.

Les objectifs spécifiques sont :

- Documenter la variabilité temporelle de la superficie d'un type de refuges thermiques (confluence de tributaires froids et d'une rivière plus chaude) au niveau de multiples sites sur une rivière à saumon (rivière Ste-Marguerite).
- Développer deux modèles statistiques qui sont GAM et MARS, afin de modéliser l'aire des refuges thermiques et la variation des superficies de ces refuges thermiques le long de la saison estivale.
- Valider les modèles statistiques et comparer les performances des deux modèles.

1.5 Structure du mémoire

Dans le cadre de cette maîtrise, les travaux réalisés sont présentés sous forme d'un mémoire par article. Le 2^e chapitre représente une synthèse de l'article en français, suivie par la version anglaise de l'article qui s'intitule « *estimation of the area of potential thermal refugia using the generalized additive model and the multivariate adaptive régression splines: case study of the Ste-Marguerite river* » et qui a été soumis à la revue *River Research and Applications*.

L'encadrement, l'orientation et la validation des analyses et résultats obtenus durant ces travaux de recherche ont été assurés et effectués par le directeur de recherche en collaboration avec le codirecteur de recherche. Durant la réalisation de l'article, l'étudiant avait pour rôle de participer aux

travaux de terrain, d'extraire les données échantillonnées, de les compiler afin de procéder à la calibration et la validation des modèles statistiques. L'étudiant a complété l'analyse comparative des résultats et fut le principal rédacteur de l'article.

2. SYNTHÈSE DE LA MÉTHODOLOGIE ET DES RÉSULTATS

1.6 Méthodologie et site d'étude

La rivière Sainte-Marguerite est située sur la rive nord du Saint-Laurent, dans la province de Québec (Canada). C'est une rivière avec une population de saumon atlantique (*Salmo salar*) (Guay et al., 2000). Divisé en deux branches, la rivière principale fait 90 km de long avec une superficie de drainage de 1120 km², tandis que l'autre tronçon appelé branche nord-est fait 85 km de long avec une superficie de drainage de 980 km². La jonction de ces deux branches est située à 2,5 km en amont de l'exutoire dans le Fjord du Saguenay (figure 3-1). Plus de 15 affluents se déversent dans le cours d'eau principal ou dans la branche nord-est, dont neuf ont déjà été étudiés par Daigle et al. (2015). Dans la présente étude, quatre sites ont été sélectionnés dans le cadre d'un projet en cours sur l'inclusion de la température dans l'habitat du saumon atlantique juvénile (figure 3-1).

Les quatre sites (stations) d'études ont été sélectionnés sur la base que chacun d'entre eux représente un refuge thermique potentiel. Les données de température ont été acquises à l'aide de thermographes déployés (53 thermographes) sur la zone d'étude (l'ensemble des quatre sites) et qui ont été positionnés à partir de la confluence de l'affluent et du cours d'eau principal sous forme de transects qui s'étendent de la rive du cours d'eau principal vers le thalweg de ce dernier (figure 3-2). Les stations 1 à 3 ont été suivies en 2018, tandis que la station 4 a été instrumentée en 2019 (tableau 3-1). Sur tous les sites, les transects ont été répartis indépendamment en parallèle. La distance entre chaque thermographe sur un transect et entre les transects varie en fonction de la taille de l'affluent et de la bathymétrie du cours d'eau principal. Les transects ayant la plus forte densité de thermographes étaient en amont près de la confluence, alors que le nombre de thermographes par transect diminue vers l'aval, avec la diminution de la largeur du panache d'eau

froide causée par la dilution avec l'eau plus chaude de la rivière principale. La température a été enregistrée toutes les 15 minutes pour les stations 1, 2 et 3, tandis que pour la station 4 l'enregistrement s'est fait toutes les 30 minutes. Une comparaison des moyennes journalières des températures de l'eau a démontré que celles-ci sont presque identiques lorsqu'elles sont calculées avec des mesures prises toutes les 15 minutes et toutes les 30 minutes. Finalement, même si nous avons modélisé les aires estimées sur une échelle journalière nous avons ajouté, à titre d'exemple, une interpolation au pas de temps horaire sur un seul site afin de montrer la variation diurne des superficies.

La température de l'air a été obtenue à la station météorologique du site expérimental de l'INRS-ETE, située à moins de 30 kilomètres des quatre stations. Ces données ont été utilisées comme prédicteurs pour les stations 1, 2 et 3, alors que pour la station 4, les données ont été collectées par l'installation d'une jauge à pression barométrique et température de l'air sur le site d'étude. Les données du débit de la branche nord-est ont été obtenues auprès du Ministère de l'environnement et de la lutte contre les changements climatiques (www.cehq.gouv.qc.ca). Le débit d'un affluent a été mesuré en 2019 à la station 4 pour tester son mérite en tant que prédicteur. Une jauge à niveau (Hobo Level Logger) a été déployée dans l'affluent et un total de six mesures ponctuelles du débit ont été effectuées afin de construire une courbe de tarage pour cette station.

1.7 Approches statistiques

1.7.1 Modèle additif généralisé

Le modèle additif généralisé introduit par Hastie and Tibshirani (1990) est un modèle non paramétrique qui utilise des fonctions de lissage (Splines) afin de réduire les irrégularités et

singularités (autrement dit les fonctions de lissage tendent à atténuer les perturbations ou les bruits de mesures) et peuvent prendre diverses formes. En effet, la relation entre la variable dépendante et les prédicteurs ne prend pas de forme prédéterminée et se construit sur la base des informations provenant des données. Ces fonctions sont construites pour toutes les variables indépendantes. De plus, une fonction de liaison adaptée à la distribution de la variable de réponse doit aussi être définie. Le modèle GAM a été calibré à l'aide des splines cubiques comme fonctions de lissage et la fonction de lien est une distribution gaussienne (Marra & Wood, 2011; Wood, 2006). Le GAM est exprimé par la formule suivante :

$$g(E(y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) + \varepsilon \quad (1)$$

Les variables prédictives sont sélectionnées via une approche pas à pas (*stepwise*). g est une fonction de lien qui permet de lier la moyenne de la variable dépendante à un ensemble de variables explicatives. $E(y)$ est l'espérance mathématique de la variable de réponse prédite. β_0 est un coefficient ou l'intersection de l'axe des y et f_i ($i=1\dots n$) sont les fonctions de lissage de la $i^{\text{ème}}$ variable explicative. ε est l'erreur estimée.

1.7.2 Modèle de régression multivariée par splines adaptatifs (MARS)

Introduit par Friedman (1991), le modèle MARS est une approche non paramétrique qui est utilisée pour construire un modèle non linéaire entre un ensemble de variables d'entrée (prédicteurs) et variables de réponse. Associé à la catégorie de régression linéaire par morceaux, MARS peut déterminer le nombre de fonctions de base ainsi que les paramètres associés à chacune d'entre elles qui sont automatiquement identifiées par les données d'entraînement. Pour chacune d'entre elles, ce modèle non paramétrique s'apparente à une régression linéaire en divisant les

variables en sous-domaines. Les intervalles réguliers sont les fonctions de base et les valeurs de rupture sont appelées nœuds. Le modèle MARS est exprimé par l'équation suivante :

$$y = f(x_{n1}, \dots, x_{np}) + e = f(x_n) + e \quad (2)$$

y est la variable de réponse, e est l'erreur estimée, $f(x_n)$ est la somme pondérée de fonctions de bases alors que $x_n = (x_{n1}, \dots, x_{np})$ représente la $n^{ième}$ donnée d'entrée avec p le nombre de variables indépendantes. MARS a été appliqué en utilisant l'approche progressive du *Forward and Backward selection*. La première étape qui est le *Forward selection*, sélectionne un ensemble de variables explicatives appropriées mais qui peuvent conduire à un modèle très complexe et surdimensionné (Friedman, 1991; Hastie & Tibshirani, 1996). Alors que la seconde étape consiste à l'utilisation du *Backward selection* qui est basée sur une validation croisée généralisée. La précision des prévisions sera améliorée et les variables moins importantes ou inutiles seront éliminées parmi l'ensemble des variables explicatives qui ont été sélectionnées lors du *Forward selection*.

1.7.3 Evaluation des performances des modèles

Dans cette étude, nous avons utilisé différents critères pour évaluer les performances des deux modèles GAM et MARS. Les critères d'évaluations utilisés sont le RMSE (*Root Mean Square Error*) et RRMSE (*relative Root Mean Square Error*), le coefficient de NASH (coefficient de détermination de la qualité de prédiction/R²), le critère d'Akaike (AIC ou *Akaike Information Criterion*), le biais (BIAS) et le biais relatif (rBIAS). Les quatre critères sont exprimés par les équations suivantes :

$$\text{BIAS} = \frac{1}{n} \sum_{i=1}^n (O_i - P_i) \quad (3)$$

Ou O_i et P_i sont les valeurs observées et les valeurs projetées à la station i respectivement. n est le nombre de point de données.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2} \quad (4)$$

Ou O_i et P_i sont les valeurs observées et les valeurs projetées à la station i respectivement. n est le nombre de point de données.

Le RMSE relatif (RRMSE) permet une comparaison inter-sites des performances :

$$\text{RRMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{(P_i - O_i)^2}{O_i}} * 100 \quad (5)$$

$$\text{NASH} = R^2 = 1 - \left[\frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - O_i^{mean})^2} \right] \quad (6)$$

Ou O_i et P_i sont les valeurs observées et les valeurs projetées, et O_i^{mean} est la moyenne des observations à la station i respectivement.

$$\text{AIC} = 2k - 2\ln(L) \quad (6)$$

Ou k est le nombre de paramètres et L est la fonction de maximum de vraisemblance du modèle.

1.8 Résultats et discussions

Dans le cadre de cette étude, l'estimation des erreurs associée aux instruments de mesure utilisés à l'interpolation reste un point important qui mériterait une analyse plus approfondie. En

ce qui concerne les thermographies utilisés l'incertitude est de $\pm 0,5$ °C. Comme constaté dans la figure 2, plus on s'éloigne de l'embouchure du tributaire, plus le nombre de thermographies diminue, ce qui peut résulter en une incertitude plus élevée en aval lors de l'interpolation des superficies. Cette incertitude n'a pas été évaluée. Les travaux qui ont suivi mes travaux (par ILIAS HANI, 2020. Communication personnelle) ont démontré que l'incertitude sur l'interpolation de ses superficies est < 1 °C en moyenne. Afin d'estimer la surface des refuges thermiques, six prédicteurs ont été testés : débit du cours d'eau principal (QP), débit du tributaire (Qt), différence de température moyenne ($DT/\Delta T$) entre le cours d'eau principal et le tributaire, température de l'air (AT), température du cours d'eau principal (Tm) et température du tributaire (Tt).

Cette approche statistique est transférable sur d'autres cours d'eau dans la mesure ou sur les autres affluents il y a au moins des mesures de la température de l'eau, de débit et de la température de l'air (ce qui constitue un nombre de variables d'entrée moindre que bien des modèles déterministes). Cependant, la méthode nécessite de longues séries chronologiques (au moins deux années de données pour une incertitude $< 30\%$ en moyenne). Bien que cinq mesures de performances aient été calculées, l'accent a été mis sur l'AIC afin de déterminer quel modèle est le mieux adapté tout en pénalisant les différents modèles sur la base de leur nombre de variables utilisées (parcimonie). Toutes les combinaisons des variables explicatives ont été testées et les résultats ont été déterminés sur la base d'une différence de température de $\Delta t > 1$ °C pour définir la frontière entre le refuge thermique et l'eau plus chaude du cours d'eau principal. Les résultats sont présentés en trois groupes : les modèles univariés afin de déterminer la variable expliquant le plus de variabilité des superficies des refuges thermiques pour chaque site, les modèles multivariés en utilisant toutes les variables et les modèles multivariés par combinaisons.

1.8.1 Variabilité spatio-temporelle

Durant la période d'échantillonnage et après avoir modélisé les aires estimées nous avons constaté que les superficies totales des refuges thermiques sur les quatre sites ne sont pas similaires et varient considérablement. Au premier site (station 1), la largeur moyenne du tributaire est de 2 m avec une profondeur moyenne de 30 cm. C'est le tributaire le plus froid des trois. À ce site, nous avons obtenu une surface maximale de 280 m² (figure 3-9). Alors que les superficies des trois autres sites sont respectivement de 46 m² (station 2, dont la largeur du tributaire et de 4 m avec une profondeur moyenne de 50 cm), 183 m² (station 3, dont la largeur du tributaire et de 2.5 m avec une profondeur moyenne de 35 cm) et 28 m² pour le dernier site (station 4, dont la largeur du tributaire et de 1.5 m avec une profondeur moyenne de 20 cm). Dans la station 1, on constate la présence continue d'un panache puisque l'approvisionnement en eau semble d'origine souterraine (la température de l'eau dans le tributaire est toujours < 10 °C) en plus du couvert végétal avec une canopée dense qui contribue à garder l'eau de l'affluent toujours froide. De plus, le chenal de la rivière principale est éloigné de la rive et le panache peut s'étendre en largeur dans une zone peu profonde. À d'autres sites, le panache d'eau froide disparaît occasionnellement en raison d'une augmentation du débit du cours d'eau principal. Ces autres stations ont un couvert végétal au-dessus de l'affluent moins dense qu'à la station 1 et les étiages des tributaires ont semblé plus sévères qu'à la station 1 (observation visuelle). Il convient de noter que dans certain cas, la variabilité diurne peut être importante et la superficie du panache est réduite de plus que la moitié en moins de 24 h (voir figures 3-7 et 3-8, par exemple).

1.8.2 Comparaison des performances entre GAM et MARS :

Les résultats des modèles GAM et MARS ont été comparés afin de déterminer le modèle le plus performant en matière d'estimation de la surface d'un refuge thermique pour chaque station sur la base des critères d'évaluation qui sont l'AIC, RMSE, NASH et le BIAS. Une comparaison initiale a été faite pour différentes définitions du panache, toutes basées sur la différence de la température moyenne dans le panache et celle de la rivière principale (Δt). Cette analyse a démontré que (1) $\Delta t > 1^{\circ}\text{C}$ représente la meilleure définition d'un refuge thermique pour cette étude de cas contrairement aux autres définitions testées qui sont $\Delta t > 2^{\circ}\text{C}$ et $\Delta t > 3^{\circ}\text{C}$. Bien que pour des raisons de préférences thermiques et de comportement du poisson cherchant un refuge, d'autres critères auraient pu être sélectionnés, le $\Delta t > 1^{\circ}\text{C}$ permet d'obtenir les résultats les plus probants sur la base de l'AIC (mis comme critère d'évaluation principal), avec des erreurs d'estimations (RRMSE) pour les deux modèles $< 39\%$.

L'évaluation des deux modèles démontre que les performances du modèle MARS sont légèrement supérieures aux performances du modèle GAM pour les quatre stations. Pour la station 1 le modèle le plus performant est le modèle MARS multivariée incluant toutes les variables avec un AIC = 929,97, RMSE = 14,21 m², NASH = 0,95 et BIAS = 0,21 m² en contraste avec le meilleur modèle GAM qui est le multivarié incluant quatre variables (QP, DT, AT, Tt) avec un AIC = 1251,03, RMSE = 16,64 m², NASH = 0,94 et BIAS = 0,15 m². La même conclusion s'applique pour la station 2 où le meilleur modèle est le MARS incluant toutes les variables avec un AIC = 345,93, RMSE = 1,99 m², NASH = 0,94 et BIAS = 0,05 m² contrairement au meilleur modèle GAM qui est le multivarié incluant quatre variables (QP, DT, AT, Tt) avec un AIC = 420,89, RMSE = 2,91 m², NASH = 0,93 et BIAS = 0,12 m². Nous avons constaté à la station 3 un léger changement pour

le meilleur modèle MARS et GAM. Ce n'est plus celui avec toutes les variables ou quatre, mais celui qui comprend que trois variables (QP, DT, Tm) avec un AIC = 405,53, RMSE = 8,32 m², NASH = 0,91 et BIAS = -0,19 m² et un AIC = 476,46, RMSE = 11,73 m², NASH = 0,91 et BIAS = -0,27 m² respectivement. Cela peut être dû au fait que la série chronologique des observations est plus courte (61 jours) contrairement aux autres stations en plus du fait que la différence de température entre le cours d'eau principal et le tributaire est fréquemment inférieure à la condition que nous avions imposée ($\Delta T = 1 \text{ } ^\circ\text{C}$).

Dans la 4^e station, l'ajout d'une nouvelle variable qui est le débit de l'affluent est significatif pour les deux modèles. Le modèle le plus performant pour MARS ainsi que GAM est celui incluant toutes les variables avec un AIC = 575,02, RMSE = 3,34 m², NASH = 0,83 et BIAS = 0,13 m² et un AIC = 708,46, RMSE = 4,49 m², NASH = 0,80 et BIAS = 0,14 m² respectivement.

Nos résultats montrent que MARS surpassé GAM en termes de RRMSE sur toutes les stations (par exemple, on trouve que pour les stations 2 et 4 le meilleur modèle du GAM a pour RRMSE = 38.4 % et un RRMSE = 34.8 % respectivement, alors que pour le MARS nous avons pour les stations 2 et 4 un RRMSE = 26.3 % et un RRMSE = 26.6 % respectivement), mais aussi que le modèle comprenant quatre ou cinq prédicteurs offre une meilleure performance que celui incluant moins de variables explicatives. Malgré ces résultats, les deux modèles montrent un fort potentiel. En plus de ces méthodes, les études futures devraient envisager de coupler ces techniques à d'autres, tels que la télédétection pour capturer l'étendue du panache ou avoir recours à des capteurs de température répartis selon la taille du panache.

1.9 Conclusion et recommandations

Comme il a déjà été mentionné, l'objectif principal de la présente étude consiste à déterminer la surface des refuges thermiques par l'utilisation de deux modèles statistiques qui sont le GAM et MARS. La mise en œuvre de ces deux approches statistiques est basée sur les données collectées au sein de la rivière Sainte-Marguerite (Québec, Canada) sur la période de juin jusqu'au début d'octobre 2018 et 2019. Nous avons pu conclure que le modèle MARS est plus performant et plus précis que le GAM. Les deux modèles restent performants dans le but de modéliser les variations des superficies de panache d'eau froide à l'échelle journalière. En effet, le MARS était plus apte avec des erreurs relatives inférieures à 30 % ($RRMSE < 30\%$) sur toutes les stations. En dépit de ces performances un peu moins bonnes que le modèle MARS, le GAM reste un outil relativement simple et efficace pour la gestion des refuges. Afin d'améliorer et de déterminer au maximum les performances des deux modèles, il serait préférable d'approfondir encore plus les études futures en évaluant le mérite de l'ajout d'autres variables explicatives associées aux flux de chaleur en surface (radiation solaire, la conductivité, la vitesse du vent, le couvert végétal, ..., Etc.), mais aussi des variables physiographiques (orientation et pentes de tributaires, par exemple). D'autres approches statistiques, ou déterministes pourraient être utilisées et comparées aux modèles développés dans la présente étude. Par exemple, les forêts aléatoires (Random Forest) ont été utilisées pour caractériser les îlots de chaleur en milieu urbain, ce qui est une problématique inverse à celle de refuges thermiques (Ho et al., 2016). Une comparaison de l'approche statistique avec une modélisation hydrodynamique 2-D serait aussi souhaitable. En attendant, les deux modèles appliqués au sein de cette étude sont relativement performants sans toutefois nécessiter un grand effort d'échantillonnage et de temps de calcul.

3. ARTICLE

**ESTIMATION OF THE AREA OF POTENTIAL THERMAL REFUGIA
USING THE GENERALIZED ADDITIVE MODEL AND MULTIVARIATE
ADAPTIVE REGRESSION SPLINES: CASE STUDY OF THE STE-
MARGUERITE RIVER**

Al Mahdi Saadi, Amina Msilini, Christian Charron, André St-Hilaire, Taha B. M. J. Ouarda

1 **Abstract:**

2 Thermal refugia in rivers are becoming critical habitat for ectotherm fish,
3 including Atlantic salmon (*Salmo salar*). In this study, two statistical modelling
4 approaches were used to estimate the areas of potential thermal refugia: the
5 generalized additive models (GAM) and multivariate adaptive regression splines
6 (MARS). GAM and MARS models were fitted independently for four sites on the
7 Ste-Marguerite River, (Quebec, Canada). Model performances were evaluated using
8 the leave-one-out cross validation (LOOCV) and the following criteria: the Akaike
9 information criterion (AIC), root-mean-square error (RMSE), relative root-mean-
10 square error (rRMSE), Nash-Sutcliffe efficiency coefficient (NASH), and finally the
11 bias (BIAS). Using an array of thermographs deployed at the confluence of a cold
12 tributary and the warmer main river stem, refugia were delineated at a daily time
13 step. Model results indicate that the estimated areas are close to the surfaces
14 interpolated using measurements, with both models and for all sites. Results suggest
15 also that MARS provides a better performance than GAM in terms of forecasting
16 and estimating the variability of the area of thermal refugia in all study-stations.

17

18 **Key words:** Thermal refugia, Daily water temperature, Generalized Additive Model
19 (GAM), Multivariate Adaptive Regression Splines (MARS).

20

21 **1. Introduction**

22

23 River thermal regimes are of great importance for the growth and distribution
24 of freshwater organisms (Ducharme, 2007; Haag & Westrich, 2002; Ozaki et al.,
25 2003). It is therefore essential that the thermal regime of rivers be well studied for
26 effective fisheries management and environmental impact assessment. Fluctuations
27 in water temperature can occur naturally or because of anthropogenic disturbances
28 such as thermal pollution, deforestation, and climate change. The major potential
29 impacts of climate change on water availability are floods and droughts. In addition
30 to these quantitative impacts, climate change also affects water temperature. Future
31 summer temperature changes are expected to increase the frequency and the
32 intensity of thermal extremes (Arismendi et al., 2013) and change the cold-water fish
33 population structure (Mantua et al., 2010).

34 High river temperatures can cause thermal stress for certain fish species (Isaak
35 et al., 2012) who may seek thermal refugia to avoid exposure to high-temperature
36 conditions (Baird & Krueger, 2003; Daigle et al., 2015; Dugdale et al., 2013). For
37 instance, juvenile salmonids may seek thermal refugia in river networks, as ambient
38 temperatures will contribute and change their distribution across large spatial ranges
39 (Welsh Jr et al., 2001). The estimated loss of habitat due to rising temperatures
40 (Ruesch et al., 2012) is an important information for managers who are attempting
41 to mitigate this potential negative impact on cold water dwellers such as salmonids
42 (Battin et al., 2007). Juvenile Atlantic salmon ceases feeding when temperatures
43 exceed 21-23 °C (Elliott, 1991) and they start to show stress-related behavioral
44 changes when temperature reaches this threshold (Cunjak et al., 1993).

45 Dugdale et al. (2013) reported that thermal refugia are classified in seven
46 types. This classification was based on Airborne Thermal Infrared (TIR) imagery.
47 The seven types are: Tributary confluence plume, lateral seep, spring brooks, cold
48 side channel, cold alcove, hyporheic upwelling and wall-base channel. In this paper
49 we focus on studying one type of thermal refuge which is the tributary confluence
50 plume (Torgersen et al., 2012).

51 For this type of refuge (and others), there is no general definition that is
52 commonly agreed upon. This is because a refuge definition depends on the
53 temperature preference of the fish species that are present in the river. Ebersole et
54 al. (2001) used a difference of 3 °C between the main river and the colder water
55 patch as their definition of a refuge for steelhead trout (*Oncorhynchus mykiss*).
56 Sutton et al. (2007) described thermal refugia as water cooler than 22-23 °C for
57 different U.S. West Coast species of salmonids. Wang et al. (2020) used a threshold
58 of 21°C and a 2°C difference between the cold plume and main stem temperature to
59 define a refuge for steelhead trout. For Atlantic salmon, Breau et al. (2007) observed
60 aggregations of juveniles in refugia that were defined as having water temperatures
61 colder than 23 °C. While Kurylyk et al. (2015) explained that Atlantic salmon parr
62 are known to use refugia characterized by temperature differences with the main
63 stem < 2 °C. In the present study, we determine the boundaries of the cold-water
64 plume using a temperature difference of 1 °C with the river main stem. This criterion
65 is subjective and was selected to ensure that the temperature difference is greater
66 than the precision of the thermographs (≤ 0.5 °C) that were deployed in the river and
67 tributaries.

68 Given the relative paucity of water temperature monitoring stations in Atlantic
69 salmon rivers in Canada, modelling tools are required to provide thermal information

70 to managers. Two types of models are used conventionally for modeling and
71 forecasting water temperature under different spatial or temporal scales:
72 deterministic and statistical (Benyahya et al., 2007; Caissie, 2006; Cole et al., 2014;
73 Marceau et al., 1986). First, deterministic models are based on the physical
74 representation of heat fluxes between the river and its common environment (Caissie
75 et al., 2007; Cox & Bolte, 2007; Edinger et al., 1968; St-Hilaire et al., 2000). Many
76 input data such as stream geometry, hydrology and meteorology are required for this
77 type of models. These data are not always readily available, particularly at regional
78 scales (BOARD, 2003; Mohseni et al., 1998).

79 Very few studies have attempted to physically model the size of thermal
80 refugia. Some attempts at modelling were done in coastal areas using hydrodynamic
81 models (Decker et al., 2013), which is the most popular method to model thermal
82 plumes in general (Lowe et al., 2009). Dzara et al. (2019) used the 1-D
83 hydrodynamic and water quality River System Model to simulate cold-water refugia
84 on the Walker River (Nevada, U.S.).

85

86 Statistical models remain a solution that typically requires fewer variables
87 than physical models (DeWeber & Wagner, 2014; Mohseni et al., 2003; Moore,
88 2006; Morrill et al., 2005; Smith, 1981). However, they need a long time series of
89 measurements (Marceau et al., 1986). They may inherently represent the physical
90 principle, by considering the correlations between the relatively small number of
91 environmental input variables and the water temperature (Caissie, 2006; Webb et al.,
92 2008). The strengths and limitations of several statistical models for water
93 temperature were presented by Benyahya et al. (2007).

94 Parametric models are a particular type of statistical models (family of
95 probability distributions) that has a finite number of parameters, that is to say a priori
96 specify the relationship between the predictors and the response variables (i.e., linear
97 or non-linear). They seem to be more efficient on a monthly or weekly basis than for
98 high frequency measurements and outputs. (Caissie, 2006; Erickson & Stefan, 2000;
99 Morrill et al., 2005; Pilgrim et al., 1998). Other models are used to overcome the
100 limitations of parametric or semi-parametric models that are commonly utilized.
101 Statistical models also includes non-parametric regression models that do not a priori
102 specify the type of relationship between predictors and predictands such as the K-
103 nearest neighbors (St-Hilaire et al., 2012), the Generalized Additive Model
104 (Frechette et al., 2019) and the Multivariate Adaptive Regression Splines (Leathwick
105 et al., 2005).

106 Very few studies have used statistical models to investigate the characteristics
107 of thermal refugia in rivers. Torgersen et al. (1999) have studied thermal habitat
108 preferences in Oregon rivers for Chinook salmon. Jeong et al. (2013) have modelled
109 temperature at a single point in thermal refugia of the Ouelle River (Canada) using
110 climate change scenarios in order to investigate their possible evolution with
111 predicted air temperature warming. Wilbur et al. (2020) compared thermal refugia
112 usage by Brook trout (*Salvelinus fontinalis*) and Atlantic salmon in New Brunswick
113 river systems (Canada). Statistical approaches have also been used to investigate the
114 presence or distribution of other fish species in refugia (e.g. Wang et al. 2020).

115

116 The aim of this study is to estimate the area of thermal refugia using two
117 different statistical models (GAM and MARS) for refugia monitored on the Ste-
118 Marguerite River (Québec, Canada). The specific objectives are (1) to develop for

119 the first time a reliable statistical model that uses a few relevant explanatory
120 variables to estimate tributary plume thermal refugia surface, (2) to evaluate and
121 compare the models performances at multiple sites with different characteristics.

122

123 **2. Materials and Methods**

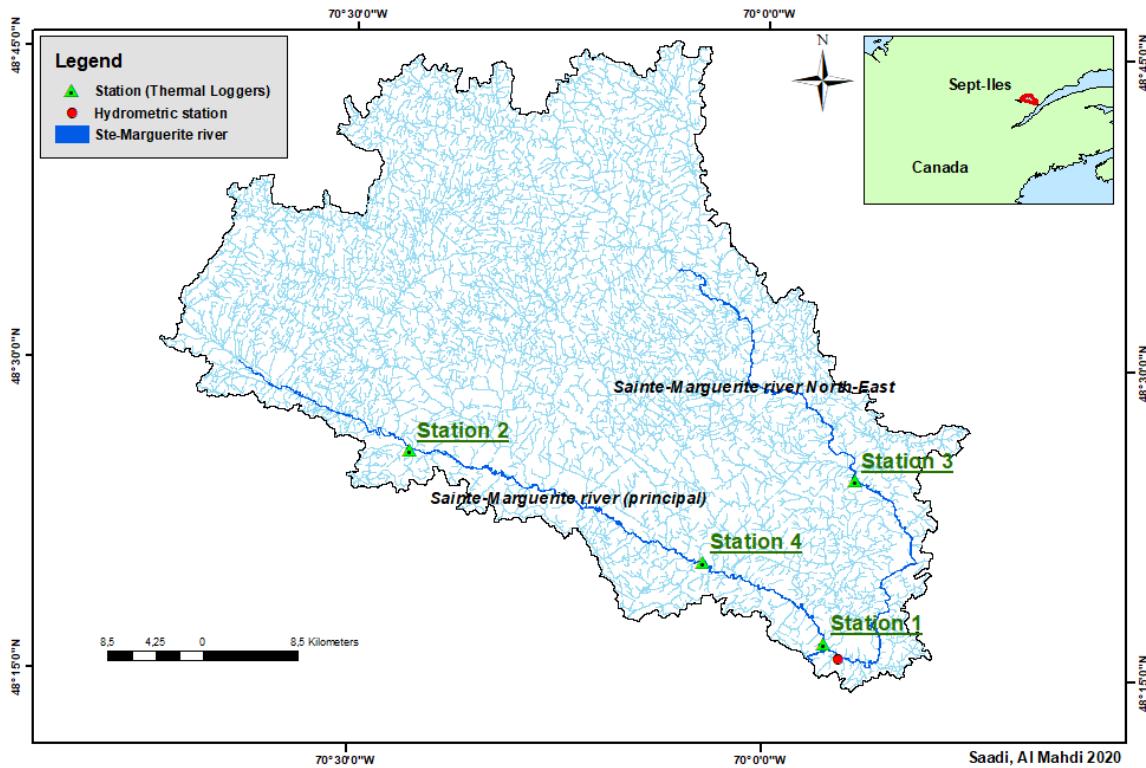
124

125 *2.1. Study Area*

126

127 The Ste-Marguerite River is located on the St. Lawrence North shore, in
128 the province of Quebec (Canada) and runs through two main segments. The first
129 one (called the Main Branch) is 90 Km long, (1120 Km² drainage area) while the
130 second one (called the Northeast Branch) is 85 Km (980 Km² drainage area).
131 Their confluence is located 2.5 Km before discharging a total of 2100 Km²
132 drainage area in the Saguenay Fjord (Figure 1). Over 15 small tributaries
133 discharge into the Main or the Northeast branches, nine of which were previously
134 studied by Daigle et al. (2015). In the present study four stations with cold water
135 tributaries emptying into the main river were selected as part as an ongoing study
136 on the inclusion of temperature in juvenile Atlantic salmon habitat (Figure 1).

Watershed of Sainte-Marguerite River



137

138 Figure 1 : Sainte-Marguerite River drainage basin with hydrometric, weather and measurement
139 stations location.

140 *2.2. Data measurement*

141

142 With the objective of developing relatively simple models, five candidate
143 predictors were tested at each site: 1) discharge of the mainstream (Qm), 2) air
144 temperature (AT), 3) water temperature in the main stem (Tm), 4) water
145 temperature in the tributary (Tt), 5) difference in temperature between the main
146 stream and the tributary (DT) and 6) discharge of the tributary (Qt).

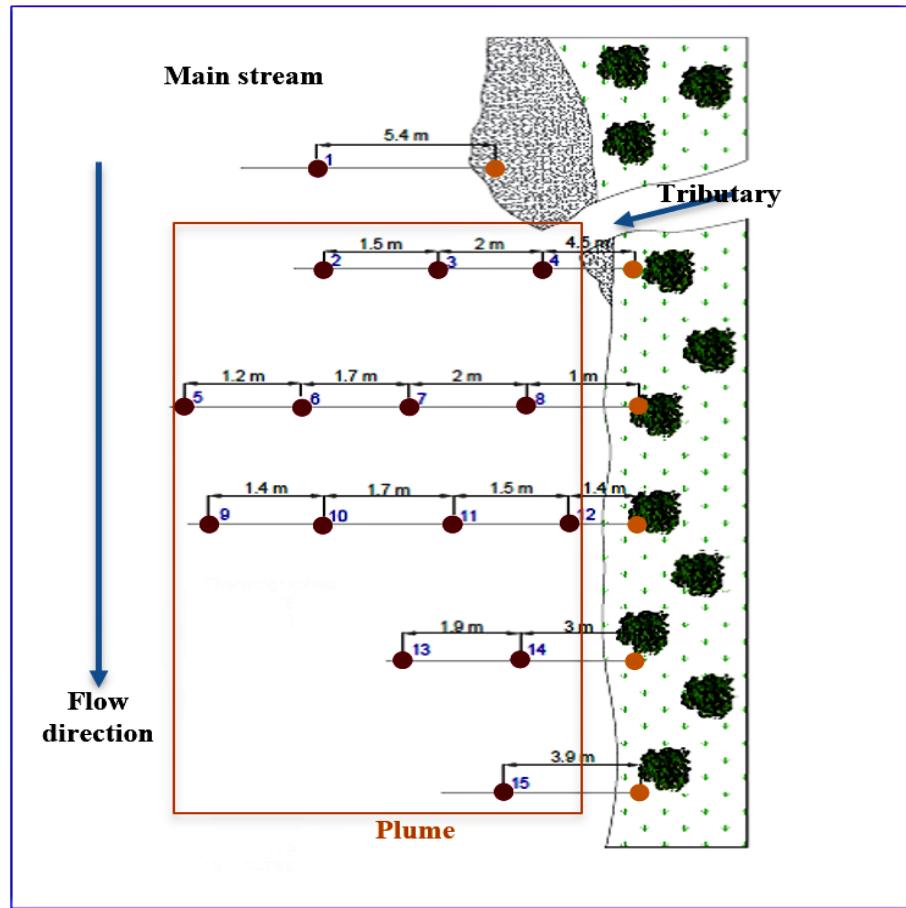
147

148 2.2.1. Water Temperature

149
150 Each site was selected because it is a potential thermal refuge at the
151 confluence of the main river and a tributary. Thermographs were deployed at
152 these four sites on the main branch of the Ste-Marguerite River. Stations 1 to
153 3 were monitored in 2018, while station 4 was monitored in 2019.
154 Temperature data from the river and plume were gathered using an array of
155 thermographs. Cold-water plume thermographs were positioned from the
156 confluence of the tributary and the main stem along transects that extend from
157 the river bank towards the main stem thalweg. At all sites, transects were
158 deployed in parallel with a distance between thermographs that vary
159 according to the size of the tributary and the main stem bathymetry (e.g.
160 Figure 2).

161
162
163 Transects with the highest thermograph density were upstream, nearest
164 the confluence and transect lengths decreased going downstream (e.g. Figure
165 2). One thermograph was placed upstream of the confluence in the main river
166 and one was installed in the tributary.

167 Air temperature data for the Ste-Marguerite River were obtained from
168 the meteorological station located less than 25 km from station 4 and less than
169 15 km from the others. They were used as predictors for stations 1, 2, and 3 in
170 2018. However, in 2019, at-site air temperature data were gathered at Station
171 4.



172

173 **Figure 2 :** Example of the deployment of thermographs Station 1, Ste-Marguerite
174 River.

175 Temperatures were recorded every 15 minutes for stations 1, 2 and 3,
176 while for station 4 in 2019, they were recorded every 30 minutes. A total of
177 53 thermographs were deployed on the four sites. Table 1 provides the number
178 of thermographs deployed, deployment dates and days of observations for
179 each site:

180 **Table 1 :** Number of deployed thermographs and their recording period for each
181 station.

| Station | Number of deployed thermographs | Data gathering period | Observations days |
|-----------|---------------------------------|--|-------------------|
| Station 1 | 15 | 28 June 2018 to 25 Septembre 2018 | 118 |
| Station 2 | 11 | 29 June 2018 to 25 Septembre 2018 | 89 |
| Station 3 | 12 | 28 July 2018 to 26 Septembre 2018 | 61 |
| Station 4 | 15 | 28 June 2019 to 13 October 2019 | 108 |

182

183

2.2.2. Flow

184

185

Ste-Marguerite River flow data were obtained from the provincial department of the ministry of Environment and Climate Change (www.cehq.gouv.qc.ca). One hydrometric station is located on the Northeast Branch and daily mean discharge records were used as a potential predictor. In addition, tributary discharge was included in 2019 to test its merit as a predictor for Station 4. A water level gauge (Hobo Level logger) was moored

191 in the tributary. A total of six spot discharge measurements were made to
192 construct a rating curve for this station.

193

194 *2.3. Statistical analysis*

195

196 2.3.1. Generalized Additive Model (GAM)

197

198 Introduced by Hastie and Tibshirani (1990), the Generalized Additive
199 Model is a nonparametric model that uses a smoothing function by regression
200 splines that can take on a variety of forms. It includes a flexible nonparametric
201 framework to build the interaction between the explanatory variables and the
202 response variable. The response variable can be explained as the sum of the
203 separate individual functions of any of the independent variables with a link
204 function suitable for the response variable distribution. In the model fitting,
205 GAM models were calibrated with a smoothing spline function and a
206 Gaussian distribution (Marra & Wood, 2011; Wood, 2006). GAM is
207 expressed by the following equation:

208

$$209 g(E(y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_n(x_n) + \varepsilon \quad (1)$$

210

211 Predictors were selected manually by a stepwise procedure. The link
212 function g is a parametric function that links the dependent variable mean to
213 a set of explanatory variables; $E(y)$ is the expectation of the predicted response
214 variable. β_0 is the coefficient or intercept of the y -axis, f_j is the j -th explanatory
215 variable's smooth function. ε is an error term that is normally distributed with

216 variance σ_ε . GAM is gaining popularity in the general field of water resources
217 (Chebana et al., 2014; Ouarda et al., 2018; Rahman et al., 2018).

219 2.3.2. Multivariate Adaptive Regression Splines (MARS)

221 Introduced by Friedman (1991), the Multivariate Adaptive Regression
222 Splines (MARS) is a nonparametric model that is used to construct a nonlinear
223 relationship between a set of input variables (predictors) and dependent
224 variables. Affiliated to the stepwise linear regression category, MARS
225 determines a number of basic functions as well as the parameters associated
226 with each one that are determined by calibration with the training data. To
227 each one of them, this nonparametric model fits a linear regression by splitting
228 the input variables into domains. The regular intervals are “Basis Functions”,
229 and break values are called “knots”. MARS is expressed by the following
230 equation :

232
$$y = f(x_{i1}, \dots, x_{ip}) + e = f(x_i) + e \quad (2)$$

234 Where y_i is the response variable, e is the error, and $x_i = (x_{i1}, \dots, x_{ip})$
235 represents the i th input data, with p representing the number of independent
236 variables. MARS was applied using a forward and backward stepwise
237 approach. The first step, which is the forward stepwise selects a set of
238 appropriate input variables (Friedman, 1991; Hastie & Tibshirani, 1996) that
239 builds a model that fits well the data, but that can be very complex and
240 overfitted (Hastie & Tibshirani, 1996). The second step is the backward
241 stepwise method where the Generalized Cross-Validation is used. This

242 additional step can increase the prediction precision, and the unnecessary
243 variables among the set of input variables that were selected in the forward
244 step are eliminated.

246 2.3.3. Assessment of model performances

248 In this study, we used different criteria to evaluate both of GAM and
249 MARS performances. These criteria are the bias (BIAS), root-mean-square
250 error (RMSE), Nash-Sutcliffe model efficiency coefficient (NASH), and
251 finally the Akaike information criterion (AIC). These four criteria are
252 expressed by :

254

$$\text{BIAS} = \frac{1}{n} \sum_{i=1}^n (O_i - P_i) \quad (3)$$

255

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2} \quad (4)$$

257
258 Where O_i and P_i are the observed values and the projected values
259 at station i respectively and n is the number of data points.

260

$$\text{NASH} = R^2 = 1 - \left[\frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - O_i^{mean})^2} \right] \quad (5)$$

262 Where O_i are the observed values, P_i are the predicted values,
263 and O_i^{mean} is the mean of the observations at station i .

264

$$265 \quad \text{AIC} = 2k - 2\ln(L) \quad (6)$$

266

267 Where k is the number of parameters, and L is the value of the
268 maximum likelihood function of the model.

269

270 3. Results

271

272 3.1. Temporal variability

273

274 The total areas of refugia at the four sites are not similar and vary
275 significantly over the sampling period. The first site has a mean refuge area of
276 280 m^2 (Figures 3-6). The Mean areas in the three other sites are 46 m^2 (station
277 2), 183 m^2 (station 3) and 28 m^2 (station 4) respectively. It can further be seen
278 that the refuge area reaches 0 at numerous occasions for all sampled stations. In
279 some cases, the cold-water plume disappearance is associated with an increase in
280 the main river discharge (e.g. Site 1). In other cases, it may be associated with
281 higher tributary temperatures or low flows. In Station 4, where tributary
282 discharge was monitored, the cold-water plume remained present, albeit small,
283 during very low flow conditions, due to relatively constant high main stem
284 temperature.

285

286
287

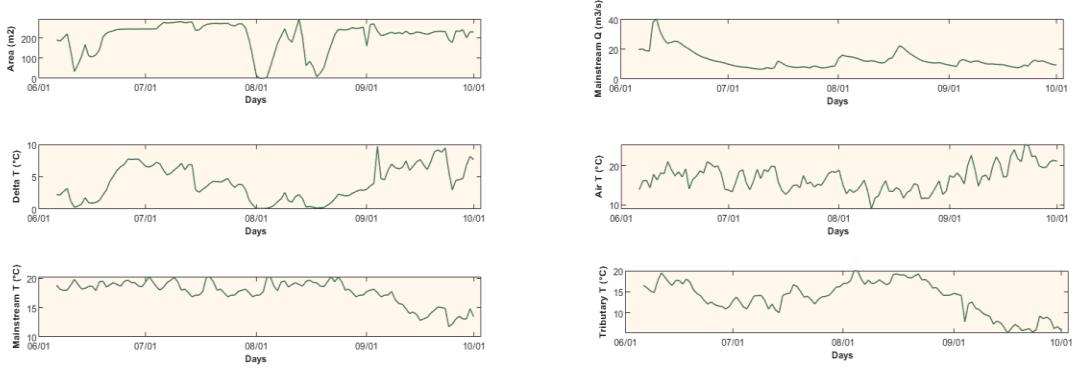


Figure 3 : Time series of potential predictors and refuge area in station 1, 2018.

288

289
290

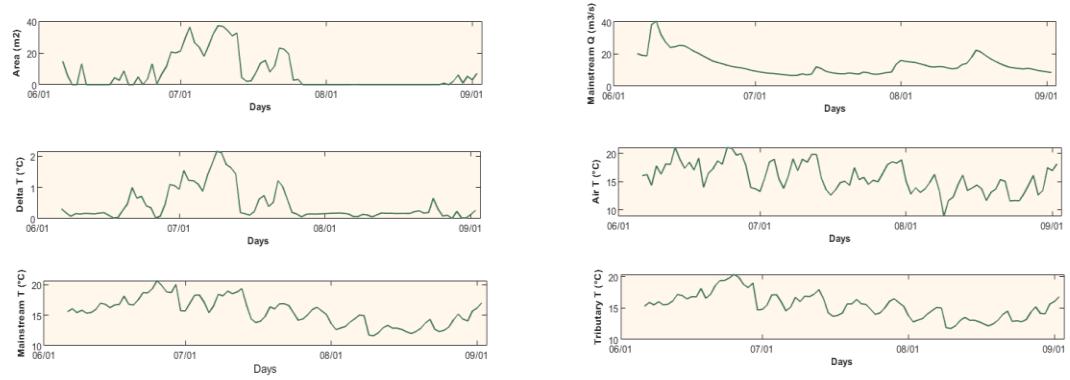
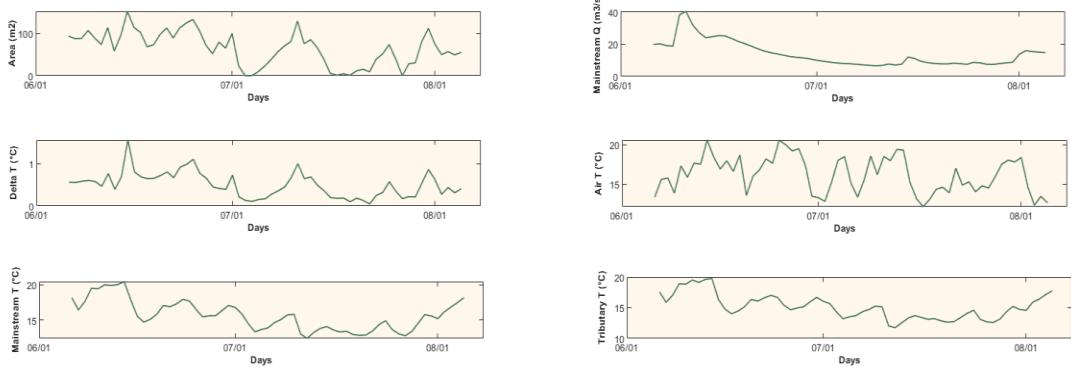
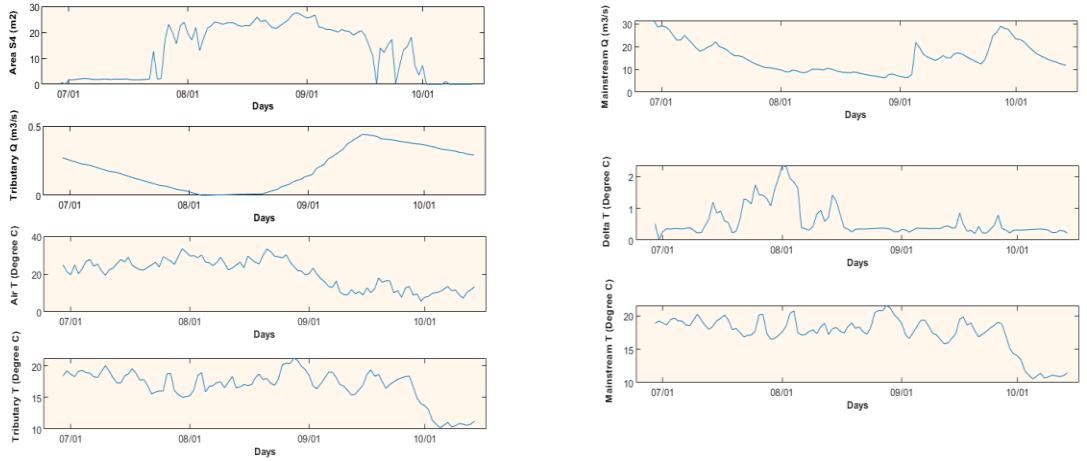


Figure 4 : Time series of potential predictors and refuge area in station 2, 2018.



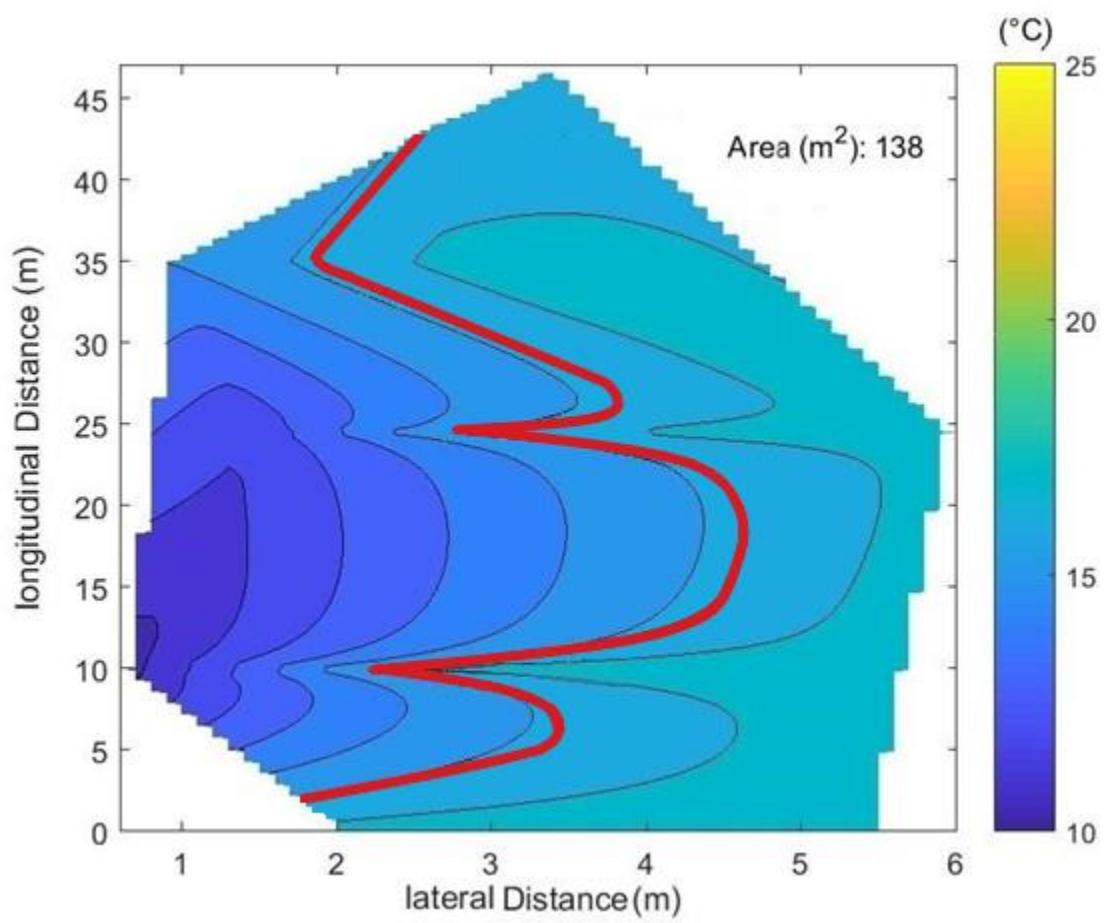
291

292
293**Figure 5 :** Time series of potential predictors and refuge area in station 3, 2018.

294

295
296**Figure 6 :** Time series of potential predictors and refuge area in station 2, 2018.297
298
299
300
301

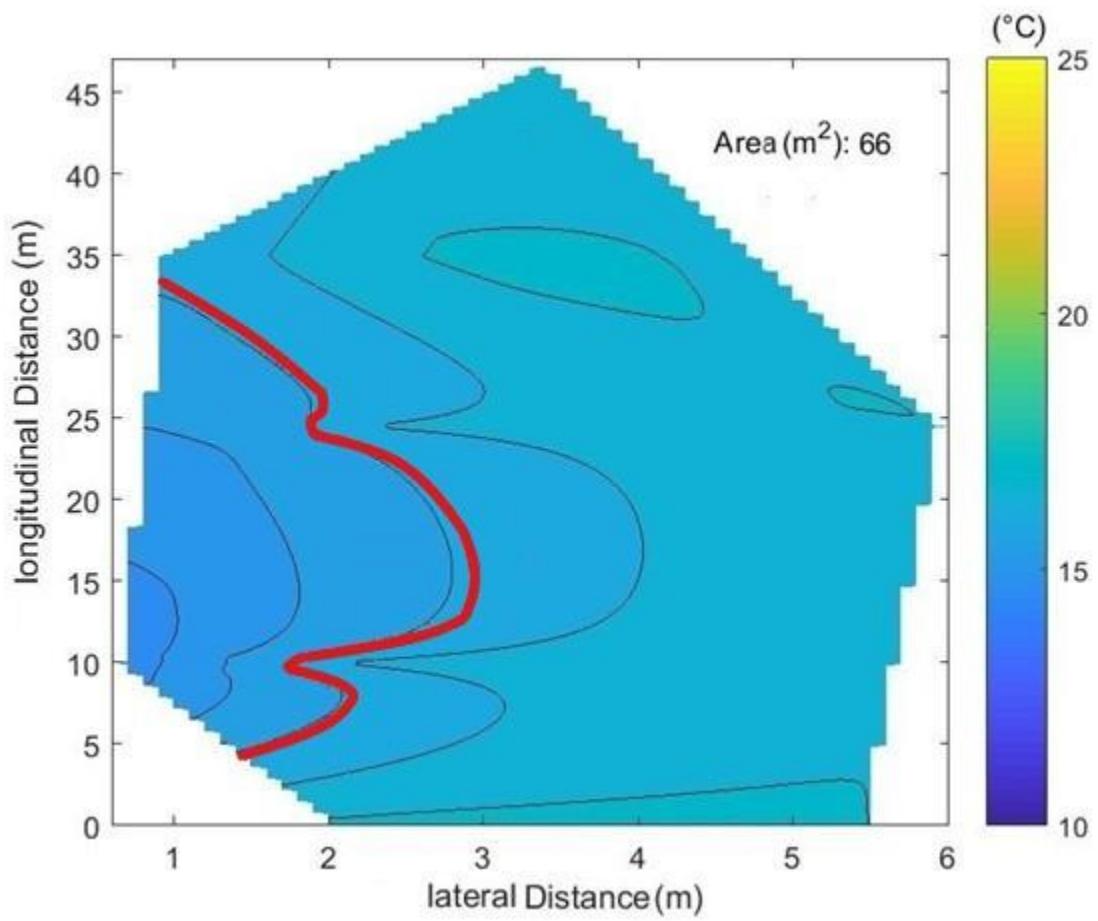
Figures 7 to 10 show examples of the spatial delineation (redline) of thermal refugia. The daily mean area was used in the present study. However, it should be noted that diel variability can be important in some cases. In the cases shown in figures 7 and 8 for instance, the area has decreased by more than half between 12:00pm to 10:00pm.



302

303

Figure 7 : Estimated area in station 1 (July 2, 2018 at 12:00pm).

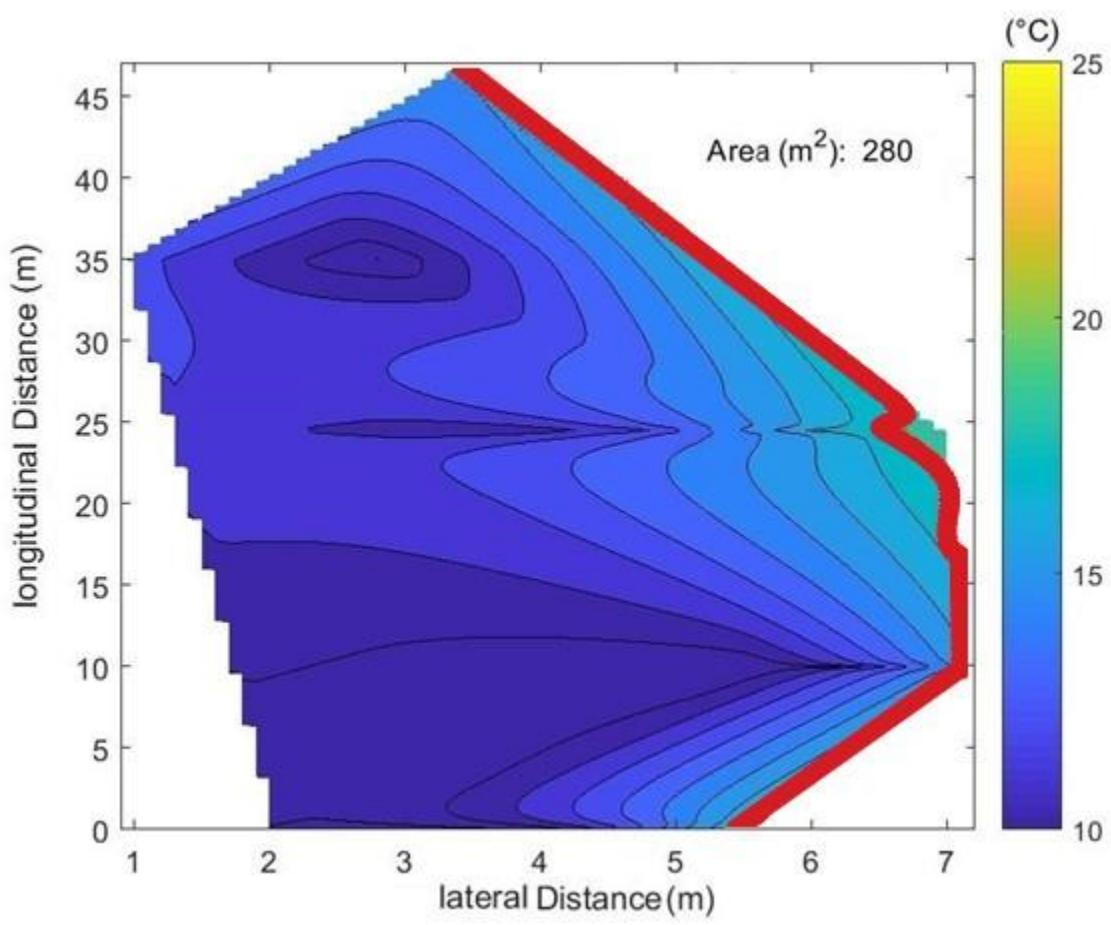


304

305

306

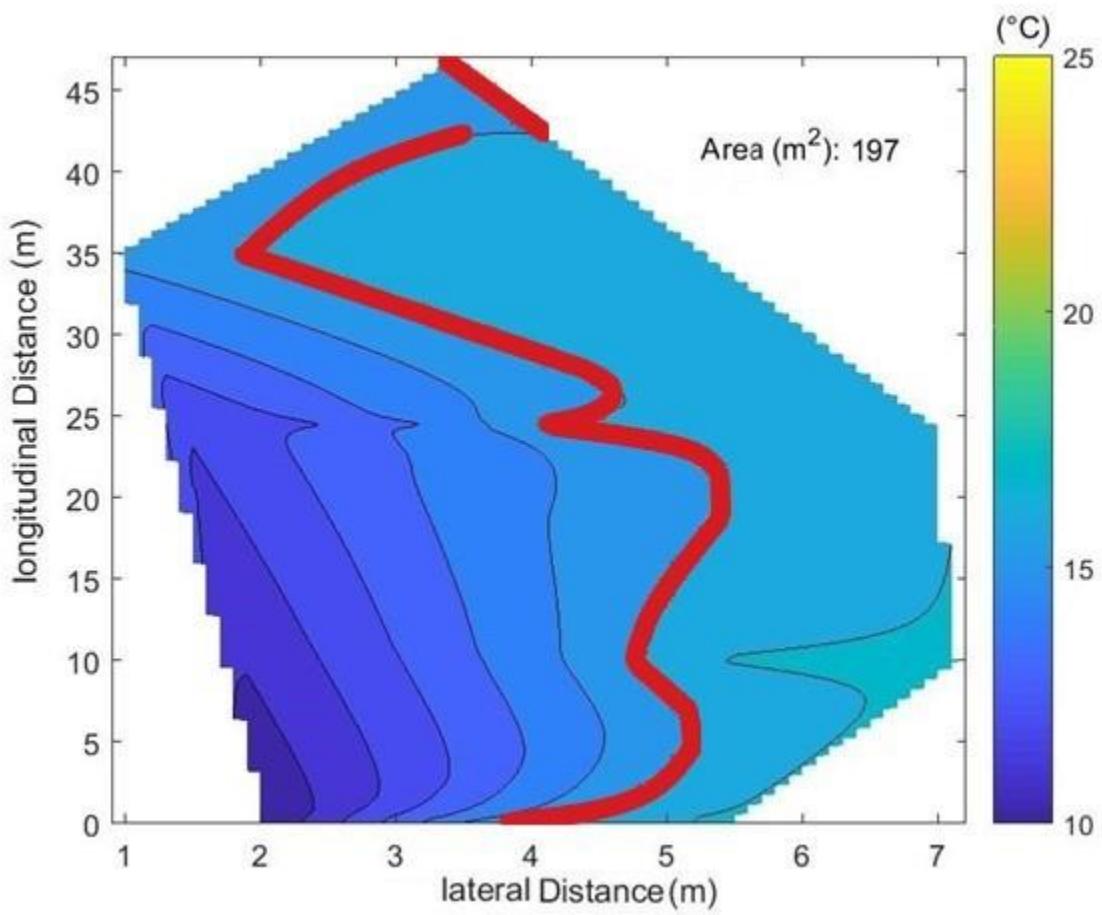
Figure 8 : Estimated area in station 1 (July 2, 2018 at 10:00pm).



307

308

Figure 9 : Estimated area in station 1 (August 2, 2018).



309

310

Figure 10 : Estimated area in station 1 (September 5, 2018).

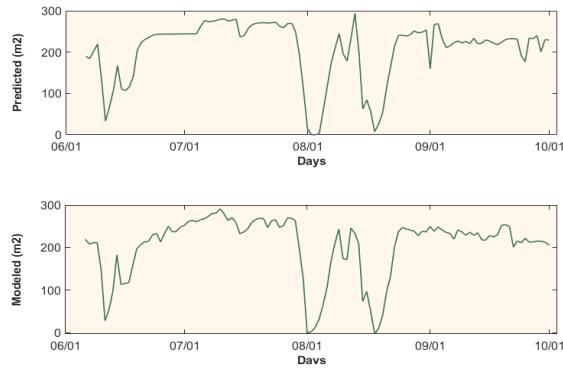
311

312

Figures 11 compares modelled areas to those interpolated from measurements using the most performant model in each site.

313

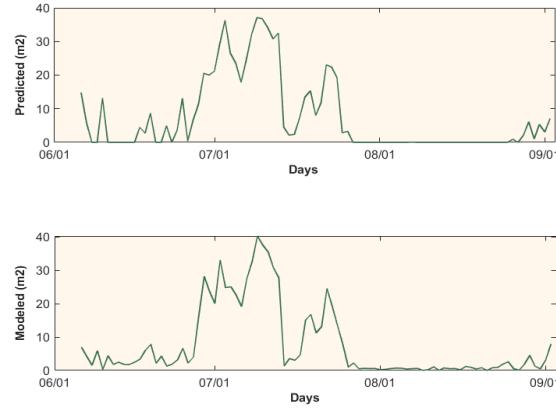
a)



314

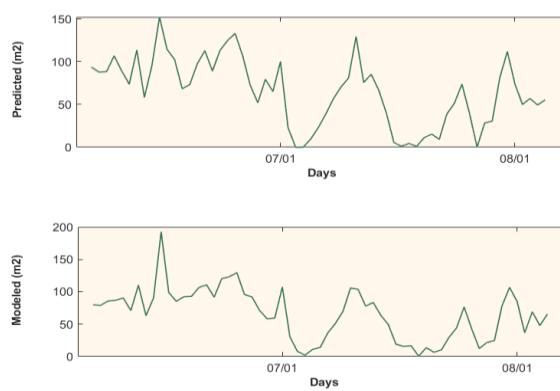
315

b)

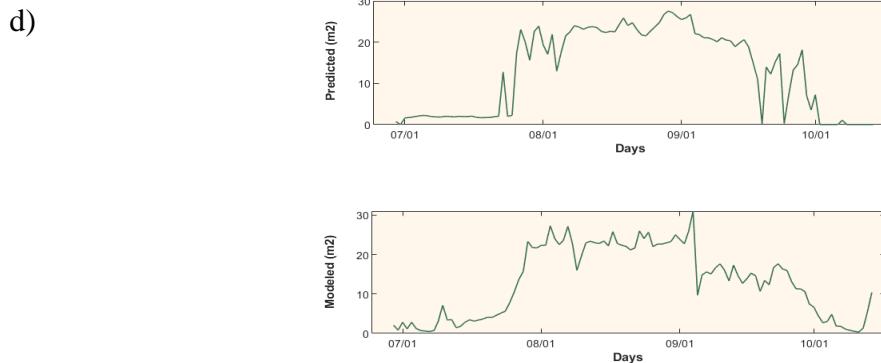


316

c)



317



318

319 **Figure 11 :** Estimated vs modeled area for the best model in station a) 1 in 2018,
 320 b) 2 in 2018, c) 3 in 2018 and d) 4 in 2019.

321

322 *3.2. Comparison of GAM and MARS performances*

323

324 Six potential predictors were tested (QM, QT, DT, AT, Tm, Tt) to estimate
 325 thermal refugia area as defined by the area with temperatures at least 1 °C colder
 326 than the river main stem.

327

328 Although four performance metrics were calculated, the emphasis was put
 329 on the results obtained by the AIC as a main criterion to compare which model
 330 is best because it accounts for parsimony. All combinations of the explanatory
 331 variables were tested, and results are presented for the best univariate models, the
 332 models with all predictors and the multivariate model that includes the best subset
 of predictors (Tables 2-5).

333

334 Results clearly indicate that MARS outperforms GAM in terms of rRMSE
 335 at all stations. In addition, the tabulated AIC values show that including four to
 five predictors leads to a better performance than fewer independent variables.

336 2019 was the only year where the discharge data from a tributary were available.
337 Its inclusion as a predictor improved the model performance.

338 **Table 2 :** Best generalized additive models and multivariate adaptive regression splines models performances evaluation using the AIC,
 339 RMSE, NASH and BIAS in station 1.

| <u><i>Station 1 (2018)</i></u> | | | | | | | | | | | | | |
|-------------------------------------|--------------|--------------|---------------|----------------|---------------|--------------|--------------------|--------------|--------------|---------------|---------------|---------------|--------------|
| <u><i>GAM</i></u> | | | | | | | <u><i>MARS</i></u> | | | | | | |
| Variable | RMSE | rRMSE | Nash | AIC | BIAS | rBIAS | Variable | RMSE | rRMSE | Nash | AIC | BIAS | rBIAS |
| <i>Univariate</i> | | | | | | | | | | | | | |
| DT | 25,97 | 12,67% | 0,8719 | 1285,86 | 0,0944 | 0,05% | DT | 22,48 | 10,96% | 0,8726 | 1022,36 | 0,0779 | -0,04% |
| <i>Multivariate (All Variables)</i> | | | | | | | | | | | | | |
| Q, DT, | | | | | | | Q, DT, | | | | | | |
| AT, | 16,83 | 8,61% | 0,9436 | 1251,14 | 0,3375 | 0,17% | AT, | 14,21 | 6,93% | 0,9455 | 929,97 | 0,2141 | 0,11% |
| Tm, Tt | | | | | | | Tm, Tt | | | | | | |
| <i>Multivariate (4 variables)</i> | | | | | | | | | | | | | |
| Q, DT, | | | | | | | Q, DT, | | | | | | |
| AT, Tt | 16,64 | 8,51% | 0,9428 | 1251,03 | 0,1458 | 0,07% | AT, Tt | 14,06 | 6,86% | 0,9429 | 930,96 | 0,1169 | 0,06% |

340 *Q: discharge of the mainstream; DT: difference in temperature; AT: air temperature; Tm: mainstream temperature; Tt: tributary temperature.

341 **Table 3 :** Best generalized additive models and multivariate adaptive regression splines models performances evaluation using the AIC,
 342 RMSE, NASH and BIAS in station 2.

| <i>Station 2 (2018)</i> | | | | | | | | | | | | | |
|--|-------------|---------------|-------------|---------------|--------------|--------------|--|-------------|---------------|-------------|---------------|--------------|--------------|
| <i>GAM</i> | | | | | | | <i>MARS</i> | | | | | | |
| <u>Variable</u> | <u>RMSE</u> | <u>rRMSE</u> | <u>Nash</u> | <u>AIC</u> | <u>BIAS</u> | <u>rBIAS</u> | <u>Variable</u> | <u>RMSE</u> | <u>rRMSE</u> | <u>Nash</u> | <u>AIC</u> | <u>BIAS</u> | <u>rBIAS</u> |
| <i>Univariate</i> | | | | | | | <i>Univariate</i> | | | | | | |
| DT | 4.32 | 64.73% | 0.89 | 526.27 | 0.11 | 1.51% | DT | 2.07 | 27.31% | 0.94 | 358.4 | 0.094 | 1.24% |
| <i>Multivariate (All Variables)</i> | | | | | | | <i>Multivariate (All Variables)</i> | | | | | | |
| Q, DT, | | | | | | | Q, DT, | | | | | | |
| AT, | 3.13 | 41.29% | 0.92 | 428.8 | 0.058 | 0.77% | AT, | 1.99 | 26.25% | 0.94 | 345.93 | 0.050 | 0.65% |
| Tm, Tt | | | | | | | Tm, Tt | | | | | | |
| <i>Multivariate (4 variables)</i> | | | | | | | <i>Multivariate (4 variables)</i> | | | | | | |
| Q, DT, | 2.91 | 38.39% | 0.93 | 420.89 | 0.120 | 1.58% | Q, DT, | | | | | | |
| AT, Tt | | | | | | | AT, Tt | 2.01 | 26.51% | 0.95 | 345.93 | 0.118 | 1.56% |

343 *Q: discharge of the mainstream; DT: difference in temperature; AT: Air Temperature; Tm: mainstream temperature; Tt: tributary temperature.

344

345 **Table 4 :** Best generalized additive models and multivariate adaptive regression splines models performances evaluation using the AIC,
 346 RMSE, NASH and BIAS in station 3.

| <u>Station 3 (2018)</u> | | | | | | | | | | | | | |
|-------------------------------------|--------------|---------------|--------------|---------------|---------------|---------------|-------------------------------------|-------------|---------------|--------------|---------------|---------------|---------------|
| <u>GAM</u> | | | | | | | <u>MARS</u> | | | | | | |
| <u>Variable</u> | <u>RMSE</u> | <u>rRMSE</u> | <u>Nash</u> | <u>AIC</u> | <u>BIAS</u> | <u>rBIAS</u> | <u>Variable</u> | <u>RMSE</u> | <u>rRMSE</u> | <u>Nash</u> | <u>AIC</u> | <u>BIAS</u> | <u>rBIAS</u> |
| <i>Univariate</i> | | | | | | | <i>Univariate</i> | | | | | | |
| DT | 11.07 | 17.07% | 0.2195 | 611.27 | 0.068 | 0.11% | DT | 9.72 | 14.98% | 0.920 | 417.68 | 0.055 | 0.08% |
| <i>Multivariate (All Variables)</i> | | | | | | | <i>Multivariate (All Variables)</i> | | | | | | |
| Q, DT, | | | | | | | Q, DT, | | | | | | |
| AT | 12.45 | 19.20% | 0.90 | 477.9 | -0.140 | 0.22% | AT | 8.31 | 12.81% | 0.902 | 406.42 | -0.111 | -0.17% |
| Tm, Tt | | | | | | | Tm, Tt | | | | | | |
| <i>Multivariate (3 variables)</i> | | | | | | | <i>Multivariate (3 variables)</i> | | | | | | |
| Q, DT, Tm | 11.73 | 18.09% | 0.910 | 476.46 | -0.266 | -0.41% | Q, DT, Tm | 8.32 | 12.83% | 0.911 | 405.53 | -0.187 | -0.29% |

347 *Q: discharge of the mainstream; DT: difference in temperature; AT: air temperature; Tm: mainstream temperature; Tt: tributary temperature.

348 **Table 5** : Best generalized additive models and multivariate adaptive regression splines models performances evaluation using the AIC,
 349 RMSE, NASH and BIAS in station 4.

| <u>Station 4 (2019)</u> | | | | | | | | | | | | | |
|-------------------------------------|-------------|---------------|--------------|---------------|--------------|--------------|-------------------------------------|-------------|---------------|---------------|---------------|--------------|--------------|
| <u>GAM</u> | | | | | | | <u>MARS</u> | | | | | | |
| <u>Variable</u> | <u>RMSE</u> | <u>rRMSE</u> | <u>Nash</u> | <u>AIC</u> | <u>BIAS</u> | <u>rBIAS</u> | <u>Variable</u> | <u>RMSE</u> | <u>rRMSE</u> | <u>Nash</u> | <u>AIC</u> | <u>BIAS</u> | <u>rBIAS</u> |
| <i>Univariate</i> | | | | | | | <i>Univariate</i> | | | | | | |
| QP | 7.17 | 55.62% | 0.492 | 745.85 | -0.110 | -0.86% | QP | 6.08 | 47.15% | 0.492 | 678.2 | -0.097 | -0.75% |
| <i>Multivariate (All Variables)</i> | | | | | | | <i>Multivariate (All Variables)</i> | | | | | | |
| Qm, Qt, | | | | | | | Qm, Qt, | | | | | | |
| DT, AT | 4.49 | 34.83% | 0.801 | 708.46 | 0.144 | 1.12% | DT, AT | 3.34 | 26.59% | 0.8346 | 575.02 | 0.128 | 0.99% |
| Tm, Tt | | | | | | | Tm, Tt | | | | | | |
| <i>Multivariate (3 variables)</i> | | | | | | | <i>Multivariate (3 variables)</i> | | | | | | |
| Qm, AT, Tm | 4.09 | 31.73% | 0.835 | 709.45 | 0.052 | 0.40% | Qm, AT, Tm | 3.82 | 29.62% | 0.822 | 587.77 | 0.035 | 0.27% |

350 *Qm: discharge of the mainstream; Qt: discharge of the tributary; DT: difference in temperature; AT: air temperature; Tm: mainstream temperature; Tt: tributary temperature.

351 4. Discussion and Conclusion

352

353 The present study focused on the first comparison of two different
354 statistical methods (MARS and GAM) to estimate thermal refuge areas using
355 hydro-meteorological predictors. Refuge area were computed from data gathered
356 in four stations along the river branches that were monitored using a total of 53
357 thermographs.

358 GAM and MARS are found to be auspicious tools that can provide an
359 estimation of thermal refugia areas using a relatively small number of predictors.
360 This tool is a relatively simple alternative to coupling hydrodynamic and heat
361 exchange models and can help considerably in the management of aquatic
362 resources. Also, the comparison of the two models has shown that MARS is more
363 accurate than GAM in all the four stations. These results highlight the interest in
364 using nonparametric approaches to describe estimates of thermal refugia areas.

365 Both models have the ability to represent nonlinear relationships between
366 the dependent variable and the predictors. For stations 1 and 2, the best results
367 were obtained when four to five predictors were used. Station 3, which is
368 characterized by a much larger tributary than the other stations, obtained best
369 results with only three predictors, excluding the temperature of the air and the
370 temperature of the tributary. This can be due to the fact that the tributary and main
371 stem temperatures are more similar at this site than the others (as quantified by
372 the Delta T) predictor, which is much smaller at Station 3 than at the other sites
373 (Figure 5). Finally, fewer observations (only 61 days) were available at Station 3
374 to calibrate and validate the models. This may also have an impact on the results
375 and on predictor selection.

376 In the fourth station, we did include a new predictor which is the tributary
377 flow (Qt). The results demonstrate that this additional predictor may be important
378 for both models and should be further investigated.

379 MARS has provided better performances in forecasting and estimating the
380 variability of the area of thermal refugia than GAM in all the stations. However,
381 despite this outcome, both models show strong potential.

382 It is possible that in some cases, the array of thermographs was insufficient
383 to cover the full extent of the cold-water plume. However, the relative variation
384 of the monitored plume area was captured, as shown in Figures 8-11. Future
385 studies should entertain the possibility of using remote sensing techniques to
386 capture the full extent of the cold-water plume or, in some cases, use Distribute
387 Temperature Sensors when the size of the plume allows for their implementation.

388 To conclude, we developed a dependable tributary-river scale statistical
389 model that uses a limited number of hydro-meteorological predictors to estimate
390 the area of the cold-water refugia. We evaluated the model performances using
391 as a major criterion the AIC then followed by the RMSE, rRMSE, NASH and
392 BIAS. The results show that all four stations represent potential thermal refugia
393 in the Ste-Marguerite River and that the area can be predicted with errors
394 ($rRMSE < 30\%$). Having relatively simple tools to predict the size of refugia may
395 be of assistance to improve management strategies to preserve these habitats that
396 are critical for the survival of cold-water dwellers such as salmonids.

397

398

399 **References**

- 400 Arismendi, I., Safeeq, M., Johnson, S.L., Dunham, J.B., Haggerty, R. 2013. Increasing synchrony
401 of high temperature and low flow in western North American streams: double trouble for
402 coldwater biota? *Hydrobiologia*, **712**(1), 61-70.
- 403 Baird, O.E., Krueger, C.C. 2003. Behavioral thermoregulation of brook and rainbow trout:
404 comparison of summer habitat use in an Adirondack River, New York. *Transactions of the*
405 *American Fisheries Society*, **132**(6), 1194-1206.
- 406 Battin, J., Wiley, M.W., Ruckelshaus, M.H., Palmer, R.N., Korb, E., Bartz, K.K., Imaki, H. 2007.
407 Projected impacts of climate change on salmon habitat restoration. *Proceedings of the*
408 *national academy of sciences*, **104**(16), 6720-6725.
- 409 Benyahya, L., Caissie, D., St-Hilaire, A., Ouarda, T.B., Bobée, B. 2007. A review of statistical
410 water temperature models. *Canadian Water Resources Journal*, **32**(3), 179-192.
- 411 BOARD, E. 2003. Estimating water temperatures in small streams in western Oregon using neural
412 network models. *Water-Resources Investigations Report*, **2**, 4218.
- 413 Breau, C., Cunjak, R., Bremset, G. 2007. Age-specific aggregation of wild juvenile Atlantic
414 salmon *Salmo salar* at cool water sources during high temperature events. *Journal of Fish*
415 *Biology*, **71**(4), 1179-1191.
- 416 Caissie, D. 2006. The thermal regime of rivers: a review. *Freshwater biology*, **51**(8), 1389-1406.
- 417 Caissie, D., Satish, M.G., El-Jabi, N. 2007. Predicting water temperatures using a deterministic
418 model: Application on Miramichi River catchments (New Brunswick, Canada). *Journal of*
419 *Hydrology*, **336**(3-4), 303-315.
- 420 Chebana, F., Charron, C., Ouarda, T.B., Martel, B. 2014. Regional frequency analysis at ungauged
421 sites with the generalized additive model. *Journal of Hydrometeorology*, **15**(6), 2418-2428.
- 422 Cole, J.C., Maloney, K.O., Schmid, M., McKenna Jr, J.E. 2014. Developing and testing
423 temperature models for regulated systems: A case study on the Upper Delaware River.
424 *Journal of hydrology*, **519**, 588-598.
- 425 Cox, M., Bolte, J. 2007. A spatially explicit network-based model for estimating stream
426 temperature distribution. *Environmental Modelling & Software*, **22**(4), 502-514.
- 427 Cunjak, R.A., Caissie, D., El-Jabi, N., Conlon, J., Pollock, T., Giberson, D.J., Komadina-
428 Douthwright, S. 1993. The Catamaran Brook (New Brunswick) habitat research project:
429 biological, physical and chemical conditions (1990-1992).
- 430 Daigle, A., Jeong, D.I., Lapointe, M.F. 2015. Climate change and resilience of tributary thermal
431 refugia for salmonids in eastern Canadian rivers. *Hydrological Sciences Journal*, **60**(6),
432 1044-1063.

- 433 Decker, M.B., Liu, H., Ciannelli, L., Ladd, C., Cheng, W., Chan, K.-S. 2013. Linking changes in
434 eastern Bering Sea jellyfish populations to environmental factors via nonlinear time series
435 models. *Marine Ecology Progress Series*, **494**, 179-189.
- 436 DeWeber, J.T., Wagner, T. 2014. A regional neural network ensemble for predicting mean daily
437 river water temperature. *Journal of Hydrology*, **517**, 187-200.
- 438 Ducharne, A. 2007. Importance of stream temperature to climate change impact on water quality.
439 *Hydrology and Earth System Sciences Discussions*, **4**(4), 2425-2460.
- 440 Dugdale, S.J., Bergeron, N.E., St-Hilaire, A. 2013. Temporal variability of thermal refuges and
441 water temperature patterns in an Atlantic salmon river. *Remote Sensing of Environment*,
442 **136**, 358-373.
- 443 Dzara, J.R., Neilson, B.T., Null, S.E. 2019. Quantifying thermal refugia connectivity by combining
444 temperature modeling, distributed temperature sensing, and thermal infrared imaging.
445 *Hydrology and Earth System Sciences*, **23**(7), 2965.
- 446 Ebersole, J., Liss, W., Frissell, C. 2001. Relationship between stream temperature, thermal refugia
447 and rainbow trout *Oncorhynchus mykiss* abundance in arid-land streams in the
448 northwestern United States. *Ecology of freshwater fish*, **10**(1), 1-10.
- 449 Edinger, J.E., Duttweiler, D.W., Geyer, J.C. 1968. The response of water temperatures to
450 meteorological conditions. *Water Resources Research*, **4**(5), 1137-1143.
- 451 Elliott, J. 1991. Tolerance and resistance to thermal stress in juvenile Atlantic salmon, *Salmo salar*.
452 *Freshwater Biology*, **25**(1), 61-70.
- 453 Erickson, T.R., Stefan, H.G. 2000. Linear air/water temperature correlations for streams during
454 open water periods. *Journal of Hydrologic Engineering*, **5**(3), 317-321.
- 455 Frechette, D.M., St-Hilaire, A., Bergeron, N. 2019. Statistical analysis of fish ladder attractivity
456 on the Nord-Est Sainte-Marguerite River, INRS, Centre Eau, Terre et Environnement.
- 457 Friedman, J.H. 1991. Multivariate adaptive regression splines. *The annals of statistics*, 1-67.
- 458 Haag, I., Westrich, B. 2002. Processes governing river water quality identified by principal
459 component analysis. *Hydrological processes*, **16**(16), 3113-3130.
- 460 Hastie, T., Tibshirani, R. 1996. Discriminant analysis by Gaussian mixtures. *Journal of the Royal
461 Statistical Society: Series B (Methodological)*, **58**(1), 155-176.
- 462 Hastie, T.J., Tibshirani, R.J. 1990. *Generalized additive models*. CRC press.
- 463 Isaak, D., Wollrab, S., Horan, D., Chandler, G. 2012. Climate change effects on stream and river
464 temperatures across the northwest US from 1980–2009 and implications for salmonid
465 fishes. *Climatic change*, **113**(2), 499-524.

- 466 Jeong, D.I., Daigle, A., St-Hilaire, A. 2013. Development of a stochastic water temperature model
467 and projection of future water temperature and extreme events in the Ouelle River basin in
468 Québec, Canada. *River Research and Applications*, **29**(7), 805-821.
- 469 Kurylyk, B.L., MacQuarrie, K.T., Linnansaari, T., Cunjak, R.A., Curry, R.A. 2015. Preserving,
470 augmenting, and creating cold-water thermal refugia in rivers: Concepts derived from
471 research on the Miramichi River, New Brunswick (Canada). *Ecohydrology*, **8**(6), 1095-
472 1108.
- 473 Leathwick, J., Rowe, D., Richardson, J., Elith, J., Hastie, T. 2005. Using multivariate adaptive
474 regression splines to predict the distributions of New Zealand's freshwater diadromous fish.
475 *Freshwater Biology*, **50**(12), 2034-2052.
- 476 Lowe, S.A., Schuepfer, F., Dunning, D.J. 2009. Case study: three-dimensional hydrodynamic
477 model of a power plant thermal discharge. *Journal of Hydraulic Engineering*, **135**(4), 247-
478 256.
- 479 Mantua, N., Tohver, I., Hamlet, A. 2010. Climate change impacts on streamflow extremes and
480 summertime stream temperature and their possible consequences for freshwater salmon
481 habitat in Washington State. *Climatic Change*, **102**(1-2), 187-223.
- 482 Marceau, P., Cluis, D., Morin, G. 1986. Comparaison des performances relatives à un modèle
483 déterministe et à un modèle stochastique de température de l'eau en rivière. *Canadian
484 Journal of Civil Engineering*, **13**(3), 352-364.
- 485 Marra, G., Wood, S.N. 2011. Practical variable selection for generalized additive models.
486 *Computational Statistics & Data Analysis*, **55**(7), 2372-2387.
- 487 Mohseni, O., Stefan, H.G., Eaton, J.G. 2003. Global warming and potential changes in fish habitat
488 in US streams. *Climatic change*, **59**(3), 389-409.
- 489 Mohseni, O., Stefan, H.G., Erickson, T.R. 1998. A nonlinear regression model for weekly stream
490 temperatures. *Water Resources Research*, **34**(10), 2685-2692.
- 491 Moore, R.D. 2006. Stream temperature patterns in British Columbia, Canada, based on routine
492 spot measurements. *Canadian Water Resources Journal*, **31**(1), 41-56.
- 493 Morrill, J.C., Bales, R.C., Conklin, M.H. 2005. Estimating stream temperature from air
494 temperature: implications for future water quality. *Journal of Environmental Engineering*,
495 **131**(1), 139-146.
- 496 Ouarda, T.B., Charron, C., Hundecha, Y., St-Hilaire, A., Chebana, F. 2018. Introduction of the
497 GAM model for regional low-flow frequency analysis at ungauged basins and comparison
498 with commonly used approaches. *Environmental Modelling & Software*, **109**, 256-271.

- 499 Ozaki, N., Fukushima, T., Harasawa, H., Kojiri, T., Kawashima, K., Ono, M. 2003. Statistical
500 analyses on the effects of air temperature fluctuations on river water qualities. *Hydrological*
501 *Processes*, **17**(14), 2837-2853.
- 502 Pilgrim, J.M., Fang, X., Stefan, H.G. 1998. STREAM TEMPERATURE CORRELATIONS
503 WITH AIR TEMPERATURES IN MINNESOTA: IMPLICATIONS FOR CLIMATE
504 WARMING 1. *JAWRA Journal of the American Water Resources Association*, **34**(5),
505 1109-1121.
- 506 Rahman, A., Charron, C., Ouarda, T.B., Chebana, F. 2018. Development of regional flood
507 frequency analysis techniques using generalized additive models for Australia. *Stochastic*
508 *environmental research and risk assessment*, **32**(1), 123-139.
- 509 Ruesch, A.S., Torgersen, C.E., Lawler, J.J., Olden, J.D., Peterson, E.E., Volk, C.J., Lawrence, D.J.
510 2012. Projected climate-induced habitat loss for salmonids in the John Day River network,
511 Oregon, USA. *Conservation Biology*, **26**(5), 873-882.
- 512 Smith, K. 1981. The prediction of river water temperatures/prédiction des températures des eaux
513 de rivière. *Hydrological Sciences Journal*, **26**(1), 19-32.
- 514 St-Hilaire, A., Morin, G., El-Jabi, N., Caissie, D. 2000. Water temperature modelling in a small
515 forested stream: implication of forest canopy and soil temperature. *Canadian Journal of*
516 *Civil Engineering*, **27**(6), 1095-1108.
- 517 St-Hilaire, A., Ouarda, T.B., Bargaoui, Z., Daigle, A., Bilodeau, L. 2012. Daily river water
518 temperature forecast model with ak-nearest neighbour approach. *Hydrological Processes*,
519 **26**(9), 1302-1310.
- 520 Sutton, R.J., Deas, M.L., Tanaka, S.K., Soto, T., Corum, R.A. 2007. Salmonid observations at a
521 Klamath River thermal refuge under various hydrological and meteorological conditions.
522 *River Research and Applications*, **23**(7), 775-785.
- 523 Torgersen, C.E., Ebersole, J.L., Keenan, D.M. 2012. Primer for identifying cold-water refuges to
524 protect and restore thermal diversity in riverine landscapes. *Environmental Protection*
525 *Agency EPA*.
- 526 Torgersen, C.E., Price, D.M., Li, H.W., McIntosh, B.A. 1999. Multiscale thermal refugia and
527 stream habitat associations of chinook salmon in northeastern Oregon. *Ecological*
528 *Applications*, **9**(1), 301-319.
- 529 Wang, T., Kelson, S.J., Greer, G., Thompson, S.E., Carlson, S.M. 2020. Tributary confluences are
530 dynamic thermal refuges for a juvenile salmonid in a warming river network. *River*
531 *Research and Applications*.

- 532 Webb, B.W., Hannah, D.M., Moore, R.D., Brown, L.E., Nobilis, F. 2008. Recent advances in
533 stream and river temperature research. *Hydrological Processes: An International Journal*,
534 22(7), 902-918.
- 535 Welsh Jr, H.H., Hodgson, G.R., Harvey, B.C., Roche, M.F. 2001. Distribution of juvenile coho
536 salmon in relation to water temperatures in tributaries of the Mattole River, California.
537 *North American Journal of Fisheries Management*, 21(3), 464-470.
- 538 Wilbur, N.M., O'Sullivan, A.M., MacQuarrie, K.T., Linnansaari, T., Curry, R.A. 2020.
539 Characterizing physical habitat preferences and thermal refuge occupancy of brook trout
540 (*Salvelinus fontinalis*) and Atlantic salmon (*Salmo salar*) at high river temperatures. *River*
541 *Research and Applications*.
- 542 Wood, S.N. 2006. On confidence intervals for generalized additive models based on penalized
543 regression splines. *Australian & New Zealand Journal of Statistics*, 48(4), 445-464.
- 544
- 545

4. RÉFÉRENCES

- Ahmadi-Nedushan, B., St-Hilaire, A., Bérubé, M., Robichaud, É., Thiémonge, N., Bobée, B. 2006. A review of statistical methods for the evaluation of aquatic habitat suitability for instream flow assessment. *River Research and Applications*, 22(5), 503-523.
- Ahmadi-Nedushan, B., St-Hilaire, A., Ouarda, T.B., Bilodeau, L., Robichaud, E., Thiémonge, N., Bobée, B. 2007. Predicting river water temperatures using stochastic models: case study of the Moisie River (Québec, Canada). *Hydrological Processes: An International Journal*, 21(1), 21-34.
- Arismendi, I., Safeeq, M., Johnson, S.L., Dunham, J.B., Haggerty, R. 2013. Increasing synchrony of high temperature and low flow in western North American streams: double trouble for coldwater biota? *Hydrobiologia*, 712(1), 61-70.
- Baird, O.E., Krueger, C.C. 2003. Behavioral thermoregulation of brook and rainbow trout: comparison of summer habitat use in an Adirondack River, New York. *Transactions of the American Fisheries Society*, 132(6), 1194-1206.
- Battin, J., Wiley, M.W., Ruckelshaus, M.H., Palmer, R.N., Korb, E., Bartz, K.K., Imaki, H. 2007. Projected impacts of climate change on salmon habitat restoration. *Proceedings of the national academy of sciences*, 104(16), 6720-6725.
- Belanger, M., El-Jabi, N., Caissie, D., Ashkar, F., Ribi, J. 2005. Water temperature prediction using neural networks and multiple linear regression. *Revue des Sciences de l'Eau/Journal of Water Science*, 18(3), 403-421.
- Benyahya, L., Caissie, D., St-Hilaire, A., Ouarda, T.B., Bobée, B. 2007. A review of statistical water temperature models. *Canadian Water Resources Journal*, 32(3), 179-192.
- Best, J.L. 1988. Sediment transport and bed morphology at river channel confluences. *Sedimentology*, 35(3), 481-498.
- Bilby, R.E. 1984. Characteristics and frequency of cool-water areas in a western Washington stream. *Journal of Freshwater Ecology*, 2(6), 593-602.
- BOARD, E. 2003. Estimating water temperatures in small streams in western Oregon using neural network models. *Water-Resources Investigations Report*, 2, 4218.
- Breau, C., Cunjak, R., Bremset, G. 2007. Age-specific aggregation of wild juvenile Atlantic salmon *Salmo salar* at cool water sources during high temperature events. *Journal of Fish Biology*, 71(4), 1179-1191.
- Brunke, M., Gonser, T. 1997. The ecological significance of exchange processes between rivers and groundwater. *Freshwater biology*, 37(1), 1-33.

- Burkholder, B.K., Grant, G.E., Haggerty, R., Khangaonkar, T., Wampler, P.J. 2008. Influence of hyporheic flow and geomorphology on temperature of a large, gravel-bed river, Clackamas River, Oregon, USA. *Hydrological Processes: An International Journal*, 22(7), 941-953.
- Caissie, D. 2006. The thermal regime of rivers: a review. *Freshwater biology*, 51(8), 1389-1406.
- Caissie, D., El-Jabi, N., Satish, M.G. 2001. Modelling of maximum daily water temperatures in a small stream using air temperatures. *Journal of Hydrology*, 251(1-2), 14-28.
- Caissie, D., Satish, M.G., El-Jabi, N. 2007. Predicting water temperatures using a deterministic model: Application on Miramichi River catchments (New Brunswick, Canada). *Journal of Hydrology*, 336(3-4), 303-315.
- Chebana, F., Charron, C., Ouarda, T.B., Martel, B. 2014. Regional frequency analysis at ungauged sites with the generalized additive model. *Journal of Hydrometeorology*, 15(6), 2418-2428.
- Chen, Y.D., Carsel, R.F., McCutcheon, S.C., Nutter, W.L. 1998. Stream temperature simulation of forested riparian areas: I. Watershed-scale model development. *Journal of Environmental Engineering*, 124(4), 304-315.
- Chenard, J.F., Caissie, D. 2008. Stream temperature modelling using artificial neural networks: application on Catamaran Brook, New Brunswick, Canada. *Hydrological Processes: An International Journal*, 22(17), 3361-3372.
- Cole, J.C., Maloney, K.O., Schmid, M., McKenna Jr, J.E. 2014. Developing and testing temperature models for regulated systems: A case study on the Upper Delaware River. *Journal of hydrology*, 519, 588-598.
- Cox, M., Bolte, J. 2007. A spatially explicit network-based model for estimating stream temperature distribution. *Environmental Modelling & Software*, 22(4), 502-514.
- Cunjak, R.A., Caissie, D., El-Jabi, N., Conlon, J., Pollock, T., Giberson, D.J., Komadina-Douthwright, S. 1993. The Catamaran Brook (New Brunswick) habitat research project: biological, physical and chemical conditions (1990-1992).
- Daigle, A., Jeong, D.I., Lapointe, M.F. 2015. Climate change and resilience of tributary thermal refugia for salmonids in eastern Canadian rivers. *Hydrological Sciences Journal*, 60(6), 1044-1063.
- Decker, M.B., Liu, H., Ciannelli, L., Ladd, C., Cheng, W., Chan, K.-S. 2013. Linking changes in eastern Bering Sea jellyfish populations to environmental factors via nonlinear time series models. *Marine Ecology Progress Series*, 494, 179-189.
- DeWeber, J.T., Wagner, T. 2014. A regional neural network ensemble for predicting mean daily river water temperature. *Journal of Hydrology*, 517, 187-200.

- Ducharne, A. 2007. Importance of stream temperature to climate change impact on water quality. *Hydrology and Earth System Sciences Discussions*, 4(4), 2425-2460.
- Dugdale, S.J., Bergeron, N.E., St-Hilaire, A. 2013. Temporal variability of thermal refuges and water temperature patterns in an Atlantic salmon river. *Remote Sensing of Environment*, 136, 358-373.
- Dzara, J.R., Neilson, B.T., Null, S.E. 2019. Quantifying thermal refugia connectivity by combining temperature modeling, distributed temperature sensing, and thermal infrared imaging. *Hydrology and Earth System Sciences*, 23(7), 2965.
- Eaton, J.G., Scheller, R.M. 1996. Effects of climate warming on fish thermal habitat in streams of the United States. *Limnology and oceanography*, 41(5), 1109-1115.
- Ebersole, J., Liss, W., Frissell, C. 2001. Relationship between stream temperature, thermal refugia and rainbow trout *Oncorhynchus mykiss* abundance in arid-land streams in the northwestern United States. *Ecology of freshwater fish*, 10(1), 1-10.
- Ebersole, J.L., Liss, W.J., Frissell, C.A. 2003. COLD WATER PATCHES IN WARM STREAMS: PHYSICOCHEMICAL CHARACTERISTICS AND THE INFLUENCE OF SHADING 1. *JAWRA Journal of the American Water Resources Association*, 39(2), 355-368.
- Edinger, J.E., Duttweiler, D.W., Geyer, J.C. 1968. The response of water temperatures to meteorological conditions. *Water Resources Research*, 4(5), 1137-1143.
- Elliott, J. 1991. Tolerance and resistance to thermal stress in juvenile Atlantic salmon, *Salmo salar*. *Freshwater Biology*, 25(1), 61-70.
- Erickson, T.R., Stefan, H.G. 2000. Linear air/water temperature correlations for streams during open water periods. *Journal of Hydrologic Engineering*, 5(3), 317-321.
- Frechette, D.M., St-Hilaire, A., Bergeron, N. 2019. Statistical analysis of fish ladder attractivity on the Nord-Est Sainte-Marguerite River, INRS, Centre Eau, Terre et Environnement.
- Friedman, J.H. 1991. Multivariate adaptive regression splines. *The annals of statistics*, 1-67.
- Gardner, B., Sullivan, P.J., Lembo, J., Arthur J. 2003. Predicting stream temperatures: geostatistical model comparison using alternative distance metrics. *Canadian Journal of Fisheries and Aquatic Sciences*, 60(3), 344-351.
- Guay, J., Boisclair, D., Rioux, D., Leclerc, M., Lapointe, M., Legendre, P. 2000. Development and validation of numerical habitat models for juveniles of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, 57(10), 2065-2075.
- Guillemette, N. 2009. Modélisation de la température de l'eau basée sur une méthode d'interpolation spatiale combinée à une approche multivariée, Université du Québec, Institut national de la recherche scientifique.

- Haag, I., Westrich, B. 2002. Processes governing river water quality identified by principal component analysis. *Hydrological processes*, 16(16), 3113-3130.
- Hastie, T., Tibshirani, R. 1996. Discriminant analysis by Gaussian mixtures. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 155-176.
- Hastie, T.J., Tibshirani, R.J. 1990. *Generalized additive models*. CRC press.
- Ho, H.C., Knudby, A., Xu, Y., Hodul, M., Aminipouri, M. 2016. A comparison of urban heat islands mapped using skin temperature, air temperature, and apparent temperature (Humidex), for the greater Vancouver area. *Science of the Total Environment*, 544, 929-938.
- Isaak, D., Wollrab, S., Horan, D., Chandler, G. 2012. Climate change effects on stream and river temperatures across the northwest US from 1980–2009 and implications for salmonid fishes. *Climatic change*, 113(2), 499-524.
- Isaak, D.J., Hubert, W.A. 2001. A HYPOTHESIS ABOUT FACTORS THAT AFFECT MAXIMUM SUMMER STREAM TEMPERATURES ACROSS MONTANE LANDSCAPES 1. *JAWRA Journal of the American Water Resources Association*, 37(2), 351-366.
- Jeong, D.I., Daigle, A., St-Hilaire, A. 2013. Development of a stochastic water temperature model and projection of future water temperature and extreme events in the Ouelle River basin in Québec, Canada. *River Research and Applications*, 29(7), 805-821.
- Kurylyk, B.L., MacQuarrie, K.T., Linnansaari, T., Cunjak, R.A., Curry, R.A. 2015. Preserving, augmenting, and creating cold-water thermal refugia in rivers: Concepts derived from research on the Miramichi River, New Brunswick (Canada). *Ecohydrology*, 8(6), 1095-1108.
- Laanaya, F. 2015. Modélisation de la température de l'eau en rivière à l'aide du modèle additif généralisé et comparaison avec d'autres approches statistiques, Université du Québec, Institut national de la recherche scientifique.
- Laanaya, F., St-Hilaire, A., Gloaguen, E. 2017. Water temperature modelling: comparison between the generalized additive model, logistic, residuals regression and linear regression models. *Hydrological sciences journal*, 62(7), 1078-1093.
- Leathwick, J., Rowe, D., Richardson, J., Elith, J., Hastie, T. 2005. Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology*, 50(12), 2034-2052.
- Li, H., Deng, X., Kim, D.Y., Smith, E.P. 2014. Modeling maximum daily temperature using a varying coefficient regression model. *Water Resources Research*, 50(4), 3073-3087.

- Lowe, S.A., Schuepfer, F., Dunning, D.J. 2009. Case study: three-dimensional hydrodynamic model of a power plant thermal discharge. *Journal of Hydraulic Engineering*, 135(4), 247-256.
- Lowney, C.L. 2000. Stream temperature variation in regulated rivers: Evidence for a spatial pattern in daily minimum and maximum magnitudes. *Water Resources Research*, 36(10), 2947-2955.
- Mantua, N., Tohver, I., Hamlet, A. 2010. Climate change impacts on streamflow extremes and summertime stream temperature and their possible consequences for freshwater salmon habitat in Washington State. *Climatic Change*, 102(1-2), 187-223.
- Marceau, P., Cluis, D., Morin, G. 1986. Comparaison des performances relatives à un modèle déterministe et à un modèle stochastique de température de l'eau en rivière. *Canadian Journal of Civil Engineering*, 13(3), 352-364.
- Marra, G., Wood, S.N. 2011. Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis*, 55(7), 2372-2387.
- Moatar, F., Gailhard, J. 2006. Water temperature behaviour in the River Loire since 1976 and 1881. *Comptes Rendus Geoscience*, 338(5), 319-328.
- Mohseni, O., Stefan, H. 1999. Stream temperature/air temperature relationship: a physical interpretation. *Journal of hydrology*, 218(3-4), 128-141.
- Mohseni, O., Stefan, H.G., Eaton, J.G. 2003. Global warming and potential changes in fish habitat in US streams. *Climatic change*, 59(3), 389-409.
- Mohseni, O., Stefan, H.G., Erickson, T.R. 1998. A nonlinear regression model for weekly stream temperatures. *Water Resources Research*, 34(10), 2685-2692.
- Moore, R.D. 2006. Stream temperature patterns in British Columbia, Canada, based on routine spot measurements. *Canadian Water Resources Journal*, 31(1), 41-56.
- Morin, G., Nzakimuena, T.-J., Sochanski, W. 1994. Prévision des températures de l'eau en rivières à l'aide d'un modèle conceptuel: le cas de la rivière Moisie. *Canadian Journal of Civil Engineering*, 21(1), 63-75.
- Morrill, J.C., Bales, R.C., Conklin, M.H. 2005. Estimating stream temperature from air temperature: implications for future water quality. *Journal of Environmental Engineering*, 131(1), 139-146.
- Mosley, M.P. 1976. An experimental study of channel confluences. *The journal of geology*, 84(5), 535-562.
- Mosley, M.P. 1983. Variability of water temperatures in the braided Ashley and Rakaia rivers. *New Zealand journal of marine and freshwater research*, 17(3), 331-342.

- Msilini, A., Masselot, P., Ouarda, T.B. 2020. Regional Frequency Analysis at Ungauged Sites with Multivariate Adaptive Regression Splines. *Journal of Hydrometeorology*, 21(12), 2777-2792.
- Ouarda, T.B., Charron, C., Hundecha, Y., St-Hilaire, A., Chebana, F. 2018. Introduction of the GAM model for regional low-flow frequency analysis at ungauged basins and comparison with commonly used approaches. *Environmental Modelling & Software*, 109, 256-271.
- Ouellet, V. 2011. Température de l'eau du fleuve Saint-Laurent: extrêmes hydroclimatiques, modélisation 2D et application à l'habitat du poisson, Université du Québec, Institut national de la recherche scientifique.
- Ozaki, N., Fukushima, T., Harasawa, H., Kojiri, T., Kawashima, K., Ono, M. 2003. Statistical analyses on the effects of air temperature fluctuations on river water qualities. *Hydrological Processes*, 17(14), 2837-2853.
- Ozaki, V.L. 1987. Geomorphic and hydrologic conditions for cold pool formation on Redwood Creek, California. IN: *Erosion and Sedimentation in the Pacific Rim*. IAHS Publication(165).
- Peterson, N., Reid, L. 1984. Wall-base channels: their evolution, distribution, and use by juvenile coho salmon in the Clearwater River, Washington. *Proceedings of the Olympic Wild Fish Conference. Fisheries Technology Program, Peninsula College, Port Angeles, Washington*. pp. 215-225.
- Pilgrim, J.M., Fang, X., Stefan, H.G. 1998. STREAM TEMPERATURE CORRELATIONS WITH AIR TEMPERATURES IN MINNESOTA: IMPLICATIONS FOR CLIMATE WARMING 1. *JAWRA Journal of the American Water Resources Association*, 34(5), 1109-1121.
- Poole, G.C., Berman, C.H. 2001. An ecological perspective on in-stream temperature: natural heat dynamics and mechanisms of human-caused thermal degradation. *Environmental management*, 27(6), 787-802.
- Rahman, A., Charron, C., Ouarda, T.B., Chebana, F. 2018. Development of regional flood frequency analysis techniques using generalized additive models for Australia. *Stochastic environmental research and risk assessment*, 32(1), 123-139.
- Rhoads, B.L., Kenworthy, S.T. 1995. Flow structure at an asymmetrical stream confluence. *Geomorphology*, 11(4), 273-293.
- Risley, J.C., Roehl, E.A., Conrads, P.A. 2003. *Estimating water temperatures in small streams in western Oregon using neural network models*. US Department of the Interior, US Geological Survey.

- Ruesch, A.S., Torgersen, C.E., Lawler, J.J., Olden, J.D., Peterson, E.E., Volk, C.J., Lawrence, D.J. 2012. Projected climate-induced habitat loss for salmonids in the John Day River network, Oregon, USA. *Conservation Biology*, 26(5), 873-882.
- Segura, C., Caldwell, P., Sun, G., McNulty, S., Zhang, Y. 2015. A model to predict stream water temperature across the conterminous USA. *Hydrological Processes*, 29(9), 2178-2195.
- Sinokrot, B.A., Stefan, H.G. 1993. Stream temperature dynamics: measurements and modeling. *Water resources research*, 29(7), 2299-2312.
- Smith, K. 1981. The prediction of river water temperatures/prédiction des températures des eaux de rivière. *Hydrological Sciences Journal*, 26(1), 19-32.
- St-Hilaire, A., Morin, G., El-Jabi, N., Caissie, D. 2000. Water temperature modelling in a small forested stream: implication of forest canopy and soil temperature. *Canadian Journal of Civil Engineering*, 27(6), 1095-1108.
- St-Hilaire, A., El-Jabi, N., Caissie, D., Morin, G. 2003. Sensitivity analysis of a deterministic water temperature model to forest canopy and soil temperature in Catamaran Brook (New Brunswick, Canada). *Hydrological processes*, 17(10), 2033-2047.
- St-Hilaire, A., Ouarda, T.B., Bargaoui, Z., Daigle, A., Bilodeau, L. 2012. Daily river water temperature forecast model with ak-nearest neighbour approach. *Hydrological Processes*, 26(9), 1302-1310.
- Stanford, J.A., Ward, J. 1993. An ecosystem perspective of alluvial rivers: connectivity and the hyporheic corridor. *Journal of the North American Benthological Society*, 12(1), 48-60.
- Stevens, B.S., DuPont, J.M. 2011. Summer use of side-channel thermal refugia by salmonids in the North Fork Coeur d'Alene River, Idaho. *North American Journal of Fisheries Management*, 31(4), 683-692.
- Sutton, R.J., Deas, M.L., Tanaka, S.K., Soto, T., Corum, R.A. 2007. Salmonid observations at a Klamath River thermal refuge under various hydrological and meteorological conditions. *River Research and Applications*, 23(7), 775-785.
- Torgersen, C.E., Ebersole, J.L., Keenan, D.M. 2012. Primer for identifying cold-water refuges to protect and restore thermal diversity in riverine landscapes. *Environmental Protection Agency EPA*.
- Torgersen, C.E., Price, D.M., Li, H.W., McIntosh, B.A. 1999. Multiscale thermal refugia and stream habitat associations of chinook salmon in northeastern Oregon. *Ecological Applications*, 9(1), 301-319.
- Van Vliet, M., Ludwig, F., Zwolsman, J., Weedon, G., Kabat, P. 2011. Global river temperatures and sensitivity to atmospheric warming and changes in river flow. *Water Resources Research*, 47(2).

- Wang, T., Kelson, S.J., Greer, G., Thompson, S.E., Carlson, S.M. 2020. Tributary confluences are dynamic thermal refuges for a juvenile salmonid in a warming river network. *River Research and Applications*.
- Ward, J. 1985. Thermal characteristics of running waters. in: *Perspectives in southern hemisphere limnology*, Springer, pp. 31-46.
- Webb, B., Clack, P., Walling, D. 2003. Water-air temperature relationships in a Devon river system and the role of flow. *Hydrological processes*, 17(15), 3069-3084.
- Webb, B.W., Hannah, D.M., Moore, R.D., Brown, L.E., Nobilis, F. 2008. Recent advances in stream and river temperature research. *Hydrological Processes: An International Journal*, 22(7), 902-918.
- Webb, B.W., Nobilis, F. 2007. Long-term changes in river temperature and the influence of climatic and hydrological factors. *Hydrological Sciences Journal*, 52(1), 74-85.
- Wehrly, K.E., Brenden, T.O., Wang, L. 2009. A comparison of statistical approaches for predicting stream temperatures across heterogeneous landscapes 1. *JAWRA Journal of the American Water Resources Association*, 45(4), 986-997.
- Welsh Jr, H.H., Hodgson, G.R., Harvey, B.C., Roche, M.F. 2001. Distribution of juvenile coho salmon in relation to water temperatures in tributaries of the Mattole River, California. *North American Journal of Fisheries Management*, 21(3), 464-470.
- Wilbur, N.M., O'Sullivan, A.M., MacQuarrie, K.T., Linnansaari, T., Curry, R.A. 2020. Characterizing physical habitat preferences and thermal refuge occupancy of brook trout (*Salvelinus fontinalis*) and Atlantic salmon (*Salmo salar*) at high river temperatures. *River Research and Applications*.
- Wood, S.N. 2006. On confidence intervals for generalized additive models based on penalized regression splines. *Australian & New Zealand Journal of Statistics*, 48(4), 445-464.