Use of the Halphen distribution family for mean wind speed estimation with application to Eastern Canada

Olga Tsvetkova[1,*], Taha B.M.J. Ouarda[1]

[1] Canada Research Chair in Statistical Hydro-Climatology, INRS-ETE, 490 rue de la Couronne, Quebec (QC), Canada, G1K 9A9

[*]Corresponding author. Email: *olga.tsvetkova@ete.inrs.ca*. Phone: +1 418 654 2691

***Abstract* — We introduce the family of three-parameter heavy-tailed distributions, the Halphen distribution family (HDF), to model the mean wind speed for the purpose of wind energy estimation. The HDF has a number of properties favorable to model wind speed data, such as lower bound at zero (absence of location parameter), flexibility to cover a large range of shapes, and an explicit form of moment generating functions. We examined 126 stations in Eastern Canada (125 stations were heavy-tailed with positive excess kurtosis) and found that HDF provides fit superior to the most commonly used distribution for this purpose, the two-parameter Weibull distribution, in 100% of the stations according to the Akaike information criterion (AIC). HDF was compared against 4 two-parameter models (Gaussian, Weibull, Gamma, and inverse Gamma) and 3 three-parameter models (generalized extreme value, generalized Gamma and Burr). The most common best-fit distribution for the stations in the Eastern Canada case study is the HDF (46% according to AIC). The results of the case study show that wind observations cannot be fit by the Halphen inverse type B, but can be modelled by Halphen A and especially well by Halphen B (minimum of Kolmogorov-Smirnov statistic among candidate distributions). 87 stations exhibited class D tail behavior. No correlation between tail behavior class and best-fit distribution was observed. We encourage the use of Halphen A and Halphen B as candidate distributions for wind resource estimation studies that are based on mean wind speed estimation.**

Keywords: Halphen A, Halphen B, Halphen distribution family, wind resource assessment, wind speed modeling

Word count: 5470

Nomenclature

$\widehat{C_s}$	Biased estimator of coefficient of skewness $\widehat{C_s} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3 \,/\, \left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}$ [1]
$\widetilde{C_s}$	Classical correction to the biased estimator $\widehat{C_s}$ [1] $\widetilde{C_s} = \widehat{C_s}\frac{\sqrt{n(n-1)}}{n-2}\left(1+\frac{8.5}{n}\right) = = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3 \,/\, \left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2} \frac{\sqrt{n(n-1)}}{n-2}\left(1+\frac{8.5}{n}\right)$
$\mu_r(D)$	The r-th central moment of the distribution D
$\mu'_r(D)$	The r-th non-central moment of the distribution D
C_k	Coefficient of kurtosis μ_4/μ_2^2 [2]
C_s	Coefficient of skewness $\mu_3/\mu_2^{3/2}$ [2]
C_v	Coefficient of variation $\sqrt{\mu_2}/\mu'_1$ [2]
$K_v(.)$	Bessel function, $K_v(2\alpha) = \frac{1}{2m^v}\int_0^\infty x^{v-1}e^{-\alpha\left(\frac{x}{m}+\frac{m}{x}\right)}\,dx$ [2][3]
$ef_v(.)$	Exponential factorial function $ef_v(\alpha) = 2\int_0^\infty x^{2v-1}e^{-\alpha(-x^2+\alpha x)}\,dx$ [2]
n_{par}	Number of parameters of a model
x_M	Mode of a sample
x_i	*i*-th observation of the wind speed time series
β_1	Skewness squared
β_2	Kurtosis
δ_1	Moment ratio $\delta_1 = \ln(A/G)$ [2]
δ_2	Moment ratio $\delta_2 = \ln(G/H)$ [2]
AD	Anderson-Darling statistic
AIC	Akaike information criterion
BIC	Bayesian information criterion
CDF	Cumulative distribution function. CDF describes the probability that a random variable X with a given probability distribution will be found at a value less than or equal to x $(P(X) \le x)$.
EV1	Extreme value type 1 distribution or Gumbel distribution
EV2	Extreme value type 2 distribution or Frechet distribution
EV3	Extreme value type 3 distribution or three-parameter Weibull distribution
EVI	Extreme value indices
EVT	Extreme value theory
G2	The two-parameter Gamma distribution
GEV	Generalized extreme value distribution
GOF	Goodness-of-fit
HA3	Halphen type A distribution
HB3	Halphen type B distribution
HDF	Halphen distribution family
HIB3	Halphen inverse type B distribution
HLY	Hourly
IG2	Inverse Gamma distribution
k	Shape parameter of the Weibull distribution
KS	Kolmogorov-Smirnov statistic
L	Likelihood function

m	Scale parameter of HAlphen distributions [2]
MK	Mann-Kendall test [4][5]
MLE	Maximum likelihood estimation
MMK	Modified Mann-Kendall test [6][5]
N2	Gaussian (normal) distribution
NB	The province of New Brunswick
NL	The province of New Foundland and Labrador
NS	The province of Nova Scotia
PDF	Probability distribution function. PDF is a mathematical way of describing the range of possible values tHAt a random variable can HAve and the respective probability of such an event occurring.
PEI	The province of Prince Edward Island
QC	The province of Quebec
QQ	Quantile-quantile
NRMSE	Normalized root mean squared error
SW	Shapiro-Wilk test [7]
W2	2-paramter Weibull distribution
WRE	Wind resource estimation
x	Wind speed time series
α	Shape parameter of Halphen and Gamma distributions [2]
λ	Scale parameter of the Weibull distribution
v	Shape parameter of Halphen distributions [2]
σ	Standard deviation
A	Arithmetic mean of a sample $A = \sum_{i=1}^{N} x_i$
F	Cumulative distribution function
$F(x)$	Cumulative distribution function
G	Geometric mean of a sample $G = (\prod_{i=1}^{N} x_i)^{1/N}$
H	Harmonic mean of a sample $H = \frac{N}{\sum_{i=1}^{N} 1/x_i}$
$M(a,b,z)$	confluent hypergeometric function $M(a,b,z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 e^{zt}\, t^{a-1}(1-t)^{b-a-1} dt$ [2][3]
Q	Quadratic moment $Q = \sum_{i=1}^{N} x_i^2$ [2]
$f(x)$	Probability distribution function (density)

List of figures

List of tables

Contents

1. Introduction

The Halphen distribution family (HDF) initially designed to specifically model annual and seasonal flows in flood frequency analysis [2] is introduced to model wind speed for the purpose of wind resource estimation (WRE). With this article we present the use of three-parameter HDF to model mean wind speed necessary for assessing sites for the purpose of wind energy development with a case study of 126 sites in Eastern Canada. Whether HDF is useful for this task is the main research question of this paper. More specifically, whether Halphen type A (HA3), Halphen type B (HB3) or Halphen inverse type B (HIB3) provide superior fit to the most commonly used distribution for WRE, i.e. the two-parameter Weibull distribution (W2)? A research question of secondary interest is the how well HDF models the maximum wind speed, hence, some additional attention has been given to the right tail and tail behavior of candidate distributions(Sections 2.2 and 3.1).

A concise literature review of the HDF, its application, popularization and advances are to follow. The HDF is a flexible and complete distribution family to fit independent and identically distributed observations that was designed specifically for river peak flow datasets by Etiene Halphen during World War II [8]. The creation of HDF was driven by "*Halphen's desire to construct his distributions so as to meet some specific requirements of fitting hydrological variables*" [8]. Perrault et al. [9] discussed the statistical properties of the distributions in [9] and later - the parameter estimation with maximum likelihood estimation (MLE) and quantile estimation in [10]. A significant body of research on the significance of the HDF in the field of flood frequency analysis was contributed by El Adlouni & Bobee [2]. Chebana, El Adlouni & Bobee discussed the estimation of HDF parameters with the method of moments in [11] and mixed estimation method in [1]. El Adlouni & Bobee [12] studied the sampling techniques for HDF and developed a discriminatory test in [13]. Zhang & Singh [14] revisited the application of HDF in flood frequency analysis. They used 198 peak flow datasets in their case study, estimated Halphen distributions parameters with the MLE method and used Kolmogorov-Smirnov statistic (KS) as a goodness-of-fit (GOF) metric and root mean square error (RMSE) – as a performance metric. Zhang & Singh [14] found that HA3/HB3 was the preferred distribution for 190 datasets, also the moment ratio diagram was found to be a reliable tool for distribution selection along with the value of kurtosis. Most importantly, the study validated the use of HDF in flood frequency analysis. Zhang & Singh [14] compared the performance of HDF with log-Pearson type 3 (LPT3) distribution. El Adlouni et al. [15] conducted an exploratory study comparing generalized extreme value distribution (GEV3) vs. HDF. The simulation-based comparisons point that HA3 distribution is more flexible than the GEV that leads to "*reasonable bias and RMSE even when samples are generated from GEV distribution*" [15]. Deholme et al. [16] introduced the HDF as a toolbox for modelling travel time variability.

Moving on to the field of WRE, extensive reviews of statistical distributions used in the wind industry have been conducted in a number of review studies [17][18][19][20][21][22]. Jung et al. [17] showed that "*115 parametric distributions were applied 2010–2018*" [17] and that W2 was "*the most frequently (44 times) evaluated distribution*" [17]. They also concluded that W2 was "*the most commonly evaluated and recommended distribution*" [17]. Aries et al. [23] used the W2, the two-parameter Gamma distribution (G2), two-parameter Wald distribution, two-parameter lognormal (LN2), the extreme value type 1 (EV1) or Gumbel, GEV, Nakagami, and generalized logistic distribution to model wind speed for deep assessment of wind speed distribution models for four sites in Algeria. Ouarda et. al. [19] enhanced the list with the three-parameter lognormal (LN3), LPT3, and normal (N). The distributions used in practice might not be limited to the above mentioned distributions, which are merely the most frequently encountered distributions in specialized literature. A comprehensive evaluation of wind speed distribution models with a case study for North Dakota by Zhou et al. [20] included W2, G2, LN2, and inverse Gaussian. Usta et al. [24] propose using skewed generalized error distribution and skewed t distribution.

All distributions that were applied at least in four different studies or recommended for use in at least two distinct studies for the purpose of WRE according to Jung et al. [17] are given in Table 1. Although a number of mixture models have been applied to this task [25], mixture distributions are outside the scope of this study and therefore omitted. Jung et al. [17], conclude that is impossible to "*identify a single distribution that can be fitted universally*" [17].

Table 1: Distributions used for modelling mean wind speed according to [17]

Group of distributions	**List of distributions**
One-parameter distributions	*Exponential, Rayleigh*
Two-parameter distributions	*Birnbaum-Sanders, Erlang, Gamma, Gumbel, Inverse Gamma, Inverse Gaussian, inverse Weibull, logistic, log-Logistic, lognormal, Nakagami, Normal, Weibull*
Three-parameter distributions	*Beta, Burr, Dagum, Generalized extreme value, generalized Pareto, lognormal, log-Pearson type 3, Weibull, generalized Gamma*
Four-parameter distributions	*Kappa*
Five-parameter distributions	*Wakeby*

1.1 Review of the Halphen distribution family (HDF)

The family is flexible and is used to fit a large variety of data [2]. The family is composed of three three-parameter distributions (scale and two shape parameters): Halphen type A (HA3), Halphen type B (HB3) and Halphen inverse type B (HIB3) [2]. HA3 is a generalization of the two-parameter harmonic distribution, HB3 with gamma distribution as a limiting case and HIB3 based on exponential factorial function with inverse Gamma (IG2) distribution as a limiting case [2]. The HDF has a number of properties that are favorable for wind data modeling, such as: lower bound at zero (absence of location parameter), flexibility to cover a large range of shapes, explicit form of moment generating functions [2]. The properties of the HDF ensure optimality and consistency of maximum likelihood estimator (MLE) used for parameter estimation [2]. HDF "*belong to the exponential family of distributions and their sufficient statistics can be obtained from their exponential forms*" [26], it makes the MLE "*asymptotically unbiased and their corresponding variances minimal regardless of the sample size*" [26]. This characteristic of the HDF is obviously of importance in cases when the sample size of interest is limited.

The HA3 distribution was introduced in 1941 as a generalization of the two-parameter Harmonic distribution with the following probability distribution function (PDF):

$$f_{HA}(x,m,\alpha,v) = \frac{1}{2m^{\nu}K_{\nu}(2\alpha)}x^{\nu-1}e^{-\alpha\left(\frac{x}{m}+\frac{m}{x}\right)}, \tag{1}$$

Where *m*>0 is the scale parameter, α>0 and v∈R are shape parameters. $K_\nu(.)$ is the modified Bessel function of second kind [2][3] (Eq.(2), can be computed by calling *bessely* function in Matlab [27] or *BesselK* function of *GPBayes* [28] package in R [29]).

$$K_\nu(2\alpha) = \frac{1}{2}\int_0^\infty e^{-\alpha[t+\frac{1}{t}]}\, t^{\nu-1}dt = \frac{1}{2m^\nu}\int_0^\infty e^{-\alpha[w/m+m/w]}\, w^{\nu-1}dw \tag{2}$$

So,

$$f_{HA}(x,m,\alpha,v) = \frac{x^{\nu-1}e^{-\alpha\left(\frac{x}{m}+\frac{m}{x}\right)}}{\int_0^\infty e^{-\alpha[x/m+m/x]}\, x^{\nu-1}dx}, \tag{3}$$

The HB3 distribution exhibits a different kind of asymptotic behaviour near zero compared to HA3 distribution [2]. The PDF of the HB3 distribution is:

$$f_{HB}(x,m,\alpha,v) = \frac{2}{m^{2\nu}ef_\nu(\alpha)}x^{2\nu-1}e^{-\alpha[-\left(\frac{x}{m}\right)^2+\alpha\frac{x}{m}]} \tag{4}$$

Where *m*>0 is the scale parameter, *α*∈R and *v*> 0 are shape parameters and $ef_\nu(\alpha)$ is the exponential factorial function [2] given in Eq. (5).

$$ef_\nu(\alpha) = 2\int_0^\infty x^{2\nu-1}e^{-\alpha(-x^2+\alpha x)}\,dx \tag{5}$$

So,

$$f_{HB}(x,m,\alpha,v) = \frac{x^{2\nu-1}e^{-\alpha[-\left(\frac{x}{m}\right)^2+\alpha\frac{x}{m}]}}{m^{2\nu}\int_0^\infty x^{2\nu-1}e^{-\alpha(-x^2+\alpha x)}\,dx} \tag{6}$$

For easier computation, the relationship between the exponential factorial function $ef_\nu(\alpha)$ and confluent hypergeometric function $M(a,b,z)$ (Eq. (7)) defined by Abromovitz & Stegan [3] is given in Eq.(8) [2]. The confluent hypergeometric function can be computed with *hypergeom* function in Matlab [27] and *hypergeo* function of package *hypergeo* [30] in R [29].

$$M(a,b,z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_0^1 e^{zt}\, t^{a-1}(1-t)^{b-a-1}dt \tag{7}$$

$$ef_\nu(\alpha) = \Gamma(\nu)M\left(\nu,\frac{1}{2},\frac{\alpha^2}{4}\right) + \alpha\Gamma\left(\nu+\frac{1}{2}\right)M\left(\nu+1/2,\frac{3}{2},\frac{\alpha^2}{4}\right) \tag{8}$$

Finally, Morlat [31] introduced the Halphen inverse type B (HIB3) distribution to complete the distribution family. The PDF of the HIB3 distribution is given in Eq. (9).

$$f_{HIB}(x,m,\alpha,v) = \frac{2}{m^{-2v} e f_v(\alpha)} x^{-2v-1} e^{-\alpha\left[-\left(\frac{x}{m}\right)^2 + \alpha\frac{x}{m}\right]} = \frac{x^{-2v-1} e^{-\alpha\left[-\left(\frac{x}{m}\right)^2 + \alpha\frac{x}{m}\right]}}{m^{-2v}\int_0^\infty x^{2v-1} e^{-\alpha(-x^2+\alpha x)}\, dx} \tag{9}$$

All distributions of the HDF possess an important quality, i.e. they all have triplets of so-called sufficient statistics [2], the summary of which is given in Table 2. The concept of a sufficient statistic, first introduced by Fisher in 1922 and refined by Neyman in 1935, carries as much information as the full dataset [2]. Sufficient statistics also enable efficient parameter estimators, which was the driving force behind Halphen's research pursuit [2].

Table 2: Sufficient statistics for the HDF (adapted from [2])

Distribtion	**HA**	**HB3**	**HIB3**
Sufficient statistics	H^{-1}	$ln(G)$	IQ^{-1}
	$ln(G)$	A	H^{-1}
	A	Q	$ln(G)$

Definitions of arithmetic, geometric, harmonic and quadratic means are given in Eq. (10) - Eq. (13) respectively:

$$A = \sum_{i=1}^{N} x_i \tag{10}$$

$$G = \left(\prod_{i=1}^{N} x_i\right)^{1/N} \tag{11}$$

$$H = \frac{N}{\sum_{i=1}^{N} 1/x_i} \tag{12}$$

$$Q = \sum_{i=1}^{N} x_i^2 \tag{13}$$

Although Dvorak et al. [8] reported that "*The Halphen type B distribution has also been identified as Toranzos system*" [8], he studied the HDF in an attempt to classify and find links to known distributions and recognized that HDF form "*their own system of frequency functions"* [8].

To sum up, the main properties of the HDF are the presence of sufficient statistics, non-existence of location parameter and the power to represent different kinds of shapes of frequency distributions [2]. A summary of the HDF is provided for reference in Table 3. For more in details information on the HDF, refer to Perreault et al. [32].

Table 3: Moments and moment ratios for the HDF (adapted from [9])

	HA3	HB3	HIB3
f	$\frac{1}{2m^v K_v(2\alpha)} x^{v-1} e^{-\alpha\left(\frac{x}{m}+\frac{m}{x}\right)}$ $x > 0$, $m > 0$, $\alpha > 0$ and $v \in R$	$\frac{2}{m^{2v} ef_v(\alpha)} x^{2v-1} e^{\left(-\frac{x^2}{m}+\alpha\frac{m}{x}\right)}$ $x > 0$, $m > 0$, $\alpha \in R$, $v > 0$	$\frac{2}{m^{-2v} ef_v(\alpha)} x^{-2v-1} e^{\left(-\frac{m^2}{x}+\alpha\frac{m}{x}\right)}$ $x > 0$, $m > 0$, $\alpha \in R$, $v > 0$
x_M	$m\left[\frac{v-1}{2\alpha}+\sqrt{\left(\frac{v-1}{2\alpha}\right)^2+1}\right]$	Can either have no mode, have a unique mode or possess one mode and one anti-mode [32]	$m\left[-\frac{\alpha}{4v+2}+\sqrt{\left(\frac{\alpha}{4v+2}\right)^2+\frac{2}{2v+1}}\right]$
μ'_r	$\frac{m^r K_{v+r}}{K_v}$	$\frac{m^r ef_{v+r/2}(\alpha)}{ef_v(\alpha)}$	$\frac{m^r ef_{v-r/2}(\alpha)}{ef_v(\alpha)}$
μ'_1	$\frac{m K_{v+1}}{K_v}$	$\frac{m ef_{v+1/2}(\alpha)}{ef_v(\alpha)}$	$\frac{m ef_{v-1/2}(\alpha)}{ef_v(\alpha)}$
μ'_2	$m^2/K_v^2(K_v K_{v+2} - K_{v+1}^2)$	$\frac{m^2(ef_v ef_{v+1} - ef_{v+1/2}^2)}{ef_v^2}$	$\frac{m^2(ef_v ef_{v-1} - ef_{v-1/2}^2)}{ef_v^2}$
μ'_3	$\frac{m^3}{K_v^3}(K_v^2 K_{v+3} - 3K_v K_{v+1} K_{v+3} + 2K_{v+1}^3)$	$\frac{m^3}{ef_v^3}(ef_v^2 ef_{v+1/2} - 3ef_v ef_{v+1/2} ef_{v+1} + 2ef_{v+1/2}^3)$	$\frac{m^3}{ef_v^3}(ef_v^2 ef_{v-1/2} - 3ef_v ef_{v-1/2} ef_{v-1/2} + 2ef_{v-1/2}^3)$
μ'_4	$\frac{m^4}{K_v^4}(K_v^3 K_{v+4} - 4K_{v+1} K_v^2 K_{v+3} + 6K_v K_{v+1}^2 K_{v+2} - 3K_{v+1}^4)$	$\frac{m^4}{ef_v^4}(ef_v^3 ef_{v+2} - 4ef_{v+1/2} ef_v^2 ef_{v+3/2} + 6ef_v ef_{v+1/2}^2 ef_{v+1} - 3ef_{v+1/2}^4)$	$\frac{m^4}{ef_v^4}(ef_v^3 ef_{v-2} - 4ef_{v-1/2} ef_v^2 ef_{v-3/2} + 6ef_v ef_{v-1/2}^2 ef_{v-1} - 3ef_{v-1/2}^4)$
C_v	$\frac{\sqrt{K_v K_{v+2} - K_{v+1}^2}}{K_{v+1}}$	$\frac{\sqrt{ef_v ef_{v+1} - ef_{v+1/2}^2}}{ef_{v+1/2}}$	$\frac{\sqrt{ef_{v-1} ef_v - ef_{v-1/2}^2}}{ef_{v-1/2}}$
C_s	$\frac{K_v^2 K_{v+3} - 3K_v K_{v+1} K_{v+2} + 2K_{v+1}^3}{(K_v K_{v+2} - K_{v+1}^2)^{3/2}}$	$\frac{ef_v^2 ef_{v+3/2} - 3ef_v ef_{v+1/2} ef_{v+1} + 2ef_{v+1/2}^3}{(ef_v ef_{v+1} - ef_{v+1/2}^2)^{3/2}}$	$\frac{ef_v^2 ef_{v-3/2} - 3ef_{v-1} ef_{v-1/2} ef_v + 2ef_{v-1/2}^3}{(ef_v ef_{v-1} - ef_{v-1/2}^2)^{3/2}}$

The article is organized in the following manner: 1 Introduction, 2 Methods, 3 Results and discussion, and 4 Conclusions and future work. The 2 Methods section includes HDF in the context of wind modelling (2.1), classes of heavy-tailed distributions in extreme value theory in Section 2.2, goodness-of-fit (GOF) metrics in Section 2.3, and finally, the description of Eastern Canada case study in Section 2.4.

2 Methods

This section describes the methods used, the data chosen and the experiment designed – the Eastern Canada case study (Section 2.4). Section 2.4. contains the description of the wind speed data, models and GOFs used in the case study.

2.1 Haphen distribution family in the context of wind speed modelling

The analysis of wind speed data often reveals that the following are typical for such data:
(1) Asymmetry and positive skewness are often present
(2) The empirical distribution has a lower bound (0 m/s)
(3) The empirical distribution typically starts at 0, reaches the maximum and then approaches 0 again at a varying rate

All of the three characteristics are also typical for hydrological data [41]. These characteristics are not limited to the field of hydrology and wind energy but are also common in other fields.

According to Bobee & Ashkar [41], such limitations imposed by the real world data lead statisticians like Karl Pearson to search for a way of constructing a family of distributions that could model such data.

2.2 Classes of heavy-tailed distributions in extreme value theory (EVT)

The 4th moment, kurtosis, of a normal distribution equals to 3. If *X* is a random variable and μ_x and σ_x are the mean and the standard deviation of that variable, then the distribution is called *heavy-tailed* (alternatively, fat-tailed or long-tailed), given Eq. (14) holds [2][33].

$$C_k = E\left[(\frac{X - \mu_x}{\sigma_x})^4\right] > 3, \tag{14}$$

Different classes of heavy-tailed distributions exist that are nested into one another [33][34] (See Figure 1). In Figure 1, **class E** represents distributions with non-existence of exponential moments, **class D** - sub-exponential distributions, **class C** stands for regularly varying distributions, **class B** – for Pareto-type tail distributions, and **class A**- for α-stable (non-normal) distributions [2][34][35]. Note, that HA3 and HB3 belong to **class D\C** (i.e., a class of sub-exponential distributions that are not regularly varying [35]), and HIB3 – to class **C\B** (Figure 2).

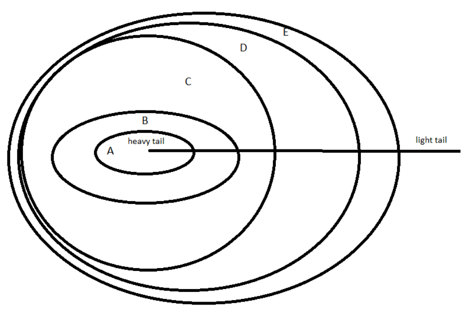

Figure 1: Nested classes of heavy-tailed distributions (adapted from [15][35])

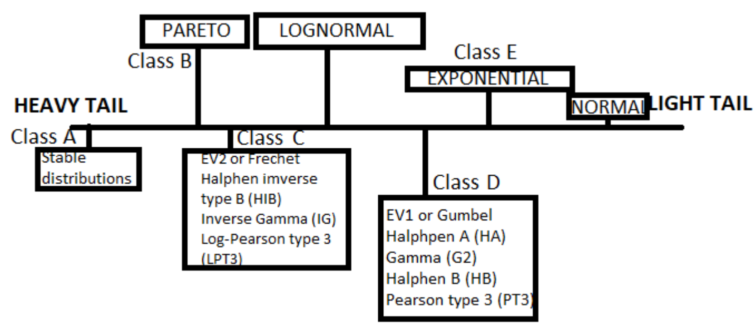

Figure 2: Distributions commonly used in hydrology, classed according to their tail behavior (adapted from [2][35])

El Adlouni & Bobee [2] proposed a decision support system to discriminate between the classes of heavy-tailed distributions. Log-log plot [36] helps discriminate between class C and class D, the plot of mean excess function [37] - between classes D and E, while Jackson's statistic [38] and Hill's ratio [39] are used to confirm the conclusions drawn previously using the aforementioned plots. Summary of the decision support system can be found in Figure 3 (adapted from [40]).

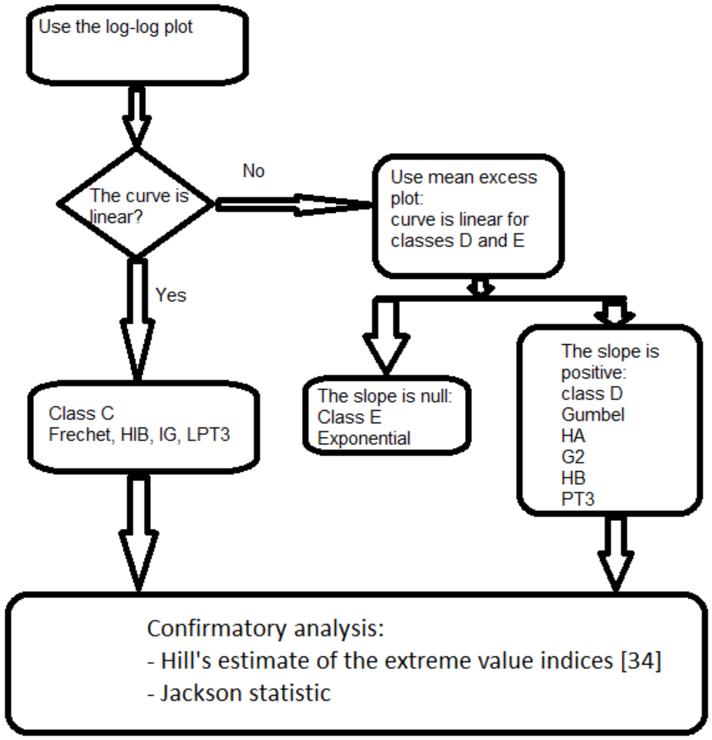

Figure 3: Discrimination between classes C, D and E (adapted from [2][40])

2.3 Goodness-of-fit (GOF) metrics

Jung et al. [17] found that KS is the most common GOF used in WRE studies. Other GOF metrics in decreasing popularity order were R^2_{pp}, χ^2, $RMSE_{PP}$, $RMSE_{PDF}$, Akaike information criterion (AIC), log-likelihood (LL), Bayesian information criterion (BIC), R^2_{PDF}, R^2_{QQ}, $RMSE_{QQ}$, Anderson-Darling statistic (AD), $RMSE_{power}$, MAE_{QQ}, and R^2_{power}.

A number of GOF metrics are based on the calculation of the likelihood function. The log-likelihood function is given in Eq. (15) [42]. The higher the value, the better the fit.

$$ln(L) = ln\left(\prod_{i=1}^{N} f_{\hat{\vartheta}}(x_i)\right) \tag{15}$$

Where $f_{\hat{\vartheta}}(x_i)$ is the PDF of the model with the estimated parameters $\hat{\vartheta}$.

One of the widely used GOF criteria that involves the log-likelihood is AIC developed by Hirotsugu Akaike [43]. AIC can be used as the means for model (distribution) selection:

$$AIC = -2\ln(L) + 2n_{par} \tag{16}$$

Where $ln(L)$ is the log likelihood function calculated according to Eq. (15), n_{par} is the number of estimated parameters in the model. A lower value indicates better model fit. The AIC is a parsimonious criterion because it takes the number of estimated parameters into account following the concept of Ockham's razor or also known as the law of parsimony.

Another GOF metric that also penalizes the number of parameters of a model (but stronger than AIC for samples with $N \geq 8$ [42]) is the Bayesian information criterion (BIC) defined as given in Eq. (17) [44]. Again, the lower the value, the better the model fit is.

$$BIC = -2 \ln L + k \ln N \tag{17}$$

Ouarda et al. [42] mention using (β_1, β_2) moment ratio diagram for model selection (See Eq. (18) for definition of order *r, r>1*), where $\beta_1$is the squared coefficient of skewness (Eq. (19)) and β_2 equals the coefficient of kurtosis (Eq. (20)).

$$C_r = \frac{\mu_r}{\mu_2^{r/2}} = \frac{E(X - \mu)^r}{\mu_2^{r/2}} \tag{18}$$

$$\beta_1 = {C_s}^2 = {C_3}^2 = \frac{\mu_3}{\mu_2^{3/2}} \tag{19}$$

$$\beta_2 = C_k = C_4 = \frac{\mu_4}{\mu_2^2} \tag{20}$$

Moments can be estimated from the data sample in the following manner, as given in Eq. (21):

$$m_r = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^r \tag{21}$$

The coefficient of skewness can hence be estimated with Eq. (22) and coefficient of kurtosis – Eq. (22):

$$\widehat{C_s} = \frac{m_3}{m_2^{3/2}} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^3 \Big/ \left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^{3/2} \tag{22}$$

$$\widehat{C_k} = \frac{m_4}{m_2^2} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^4 \Big/ \left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^{2} \tag{23}$$

For the HDF, two moment ratio diagrams are common - the (C_v, C_s) and the (δ_1, δ_2) moment ratio diagrams [2]. C_v is the coefficient of variation, defined by Eq. (24) and estimated from sample by Eq. (25). C_s is the coefficient of skewness and was already discussed above. Chebana et al. [1] showed that Eq. (22) is a biased estimator of C_s and gave a correction with Eq. (26).

$$C_v = \frac{\sigma}{\mu} \tag{24}$$

$$\widehat{C_v} = \frac{s}{\bar{x}} \tag{25}$$

$$\widetilde{C_s} = \widehat{C_s}\frac{\sqrt{N(N-1)}}{N-2}\left(1+\frac{8.5}{n}\right) = \frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^3 \Big/ \left(\frac{1}{n}\sum_{i=1}^{N}(x_i-\bar{x})^2\right)^{3/2} \frac{\sqrt{N(N-1)}}{N-2}\left(1+\frac{8.5}{N}\right) \tag{26}$$

According to El Adlouni & Bobee [2], δ_1, δ_2 are defined by Eq. (27) and Eq. (28) respectively.

$$\delta_1 = \ln\left(\frac{A}{G}\right) = \ln\left(\frac{\sum_{i=1}^{N} x_i}{\left(\prod_{i=1}^{N} x_i\right)^{1/N}}\right) \tag{27}$$

Where *A* and *G* are the arithmetic and geometric means respectively.

$$\delta_2 = \ln\left(\frac{G}{H}\right) = \ln\left(\frac{\left(\prod_{i=1}^{N} x_i\right)^{1/N}}{\frac{N}{\sum_{i=1}^{N} 1/x_i}}\right) \tag{28}$$

Where *G* and *H* are the geometric and harmonic means respectively.

2.4 The Eastern Canada case study

This subsection presents the description of the Eastern Canada case study conducted in the practical part of this work. The case study was performed with the use of Matlab [27] and R [29].

The case study included 126 stations across the Eastern part of Canada (see Figure 4): 13 stations in the province of New Brunswick (NB), 35 – in New Foundland and Labrador (NL), 23 – in Nova Scotia (NS), 5 – in Prince Edward Island (PE), and 50 in the province of Quebec (QC). The dataset used in the case study contains hourly observations of wind speed at 126 locations in Eastern Canada and can be obtained for reproducibility in [45], the initial meteorological data source is [46].

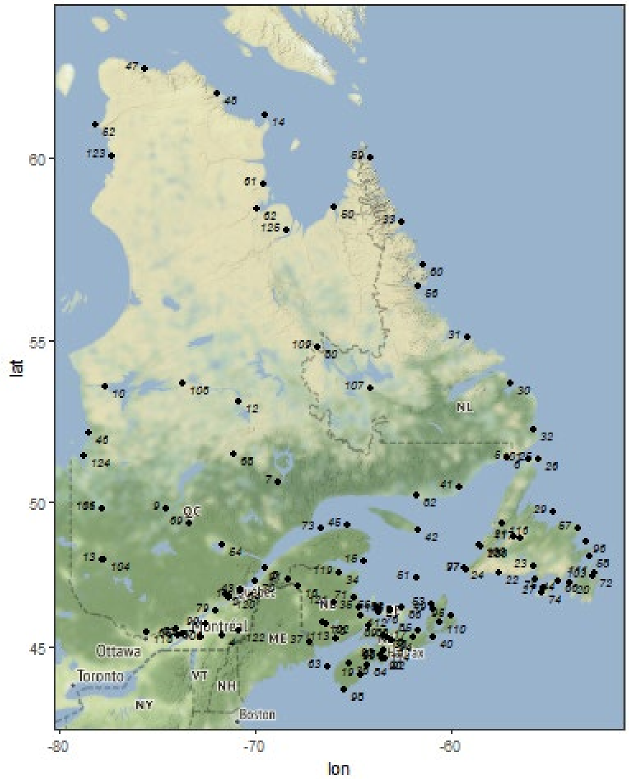

Figure 4: 126 stations in Eastern Canada that passed the modified Mann-Kendall test [4]

All stations passed the modified Mann-Kendall (MMK) test for serially correlated data [5] using Hamed & Rao variance correction approach [6], so the samples can be considered stationary (i.e. there is no monotonic trend in the data). Moreover, the stations passed the Bayesian change point detection procedure of Seidou et al. [47] without detection of shifts. Details about the stations can be found in Table 12 (Appendix I). The summary of wind speed data sample statistics, i.e. minimum, maximum value, arithmetic, geometric, harmonic means, standard deviation, skewness, kurtosis, excess kurtosis, coefficient of variation and skewness are given in Figure 5. The values of modified Mann-Kendall test (MMK) test statistic and associated p-value can be found in Table 13 (Appendix II).

The GOF metrics used in the case study are summarized in Table 4.

Table 4: Goodness-of-fit metrics used in the Eastern Canada case study

GOFM	Formula	Comment
Likelihood	$L = \prod_{i=1}^{N} f_{\hat{\vartheta}}(x_i)$	$f_{\hat{\vartheta}}(x_i)$ is the PDF of the model with the estimated parameters $\hat{\vartheta}$
Akaike information criterion	$AIC = -2\ln\sum_{i=0}^{N} f(x_i) + 2n_{par}$	n_{par} is the number of parameters of a distribution
Bayesian information criterion	$BIC = -2\ln\sum_{i=0}^{N} f(x_i) + n_{par}\ln(N)$	*N* is the wind speed sample size, n_{par} is the number of parameters of a distribution

Kolmogorov-Smirnov statistic	$KS = \max_{0 \le i \le N} \lvert F_i - \widehat{F_i} \rvert$	F_i is the empirical cumulative distribution function (CDF) and $\widehat{F_i}$ is the theoretical CDF
Anderson-Darling statistic	$AD = N \int_{-\infty}^{\infty} \frac{\left(F(x) - \widehat{F(x)}\right)^2}{\widehat{F(x)}(1 - \widehat{F(x)})} dF(x)$	F is the empirical CDF and $\hat{F}$ is the theoretical CDF. AD places more weight on the tails of the distribution.
Normalized root mean square error	$NRMSE = \sqrt{\frac{\sum_{i=0}^{n}(x_i - \widehat{x_i})^2}{n}} \Big/ \bar{x}$	x_i is the observed wind speed, $\widehat{x_i}$ is the predicted wind speed, $\bar{x}$ is the average observation.

The distributions used in Eastern Canada case study with their respective acronyms, PDFs and references of use in similar studies are given in Table 5.

Table 5: Distributions used in the Eastern Canada case study

Distribution	**Acronym**	**PDF**	**Reference of use in similar WRE studies**
Generalized extreme value	GEV3	$f_{GEV3}(x,k,\lambda,u) = \frac{1}{\lambda}\left[1 - \frac{k}{\lambda}(x-u)\right]^{\frac{1}{k}-1} e^{-\left[1-\frac{k}{\lambda}(x-u)\right]^{\frac{1}{k}}}$, $\begin{cases} x > u + \alpha/_k \ if\ k < 0 \\ x < u + \alpha/_k \ if\ k > 0 \end{cases}$	[23] [48][49][50][51]
Two-parameter Weibull	W2	$f_{W2}(x,k,\lambda) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}, k > 0, \lambda > 0, x \ge 0$	[50][51][52][53] [54][55][56][57] [58][59][60][61]
Halphen type A	HA3	$f_{HA3}(x,m,\alpha,\nu) = \frac{1}{2m^{\nu}K_{\nu}(2\alpha)} x^{\nu-1} e^{-\alpha\left(\frac{x}{m}+\frac{m}{x}\right)}$, $m,\nu,x > 0$	First time use for wind speed data
Halphen type B	HB3	$f_{HB3}(x,m,\alpha,\nu) = \frac{2}{m^{2\nu} ef_{\nu}(\alpha)} x^{2\nu-1} e^{-\alpha\left[-\left(\frac{x}{m}\right)^2 + \alpha\frac{x}{m}\right]}$, $m,\nu,x > 0$	First time use for wind speed data
Halphen inverse type B	HIB3	$f_{HIB3}(x,m,\alpha,\nu) = \frac{2}{m^{-2\nu} ef_{\nu}(\alpha)} w^{-2\nu-1} e^{-\alpha\left[-\left(\frac{w}{m}\right)^2 + \alpha\frac{w}{m}\right]}$, $m,\nu,x > 0$	First time use for wind speed data
Gaussian (Normal)	N2	$f_{N2}(x,\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	[17][62]
Two-parameter Gamma	G2	$f_{G2}(x,\alpha,\beta) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}, k,\theta,x > 0$	[23] [48][50][51][54] [60][63]
Generalized Gamma	GG3	$f_{GG3}(x,a,d,p) = \frac{p}{\Gamma(d/p)a^d} x^{d-1} e^{-\left(\frac{x}{a}\right)^p}, a,d,p,x > 0$	[48]
Inverse Gamma	IG2	$f_{IG2}(x,\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\left(\frac{1}{x}\right)^{\alpha+1} e^{-\frac{\beta}{x}}, x > 0$	[51][63]
Burr	BUR3	$f_{BUR3}(x,\alpha,\beta) = \frac{\frac{kc}{\alpha}(x/a)^{c-1}}{(1+(x/a)^c)^{k+1}} x > 0, \alpha > 0, c > 0, k > 0$	[64]

The MLE is the most commonly used distribution parameter estimation technique for wind speed data [17]. MLE was used for parameter estimation of all models in the Eastern Canada case study.

3 Results and discussion

The descriptive statistics were calculated for the wind speed hourly data samples (Figure 5). It is worth mentioning that all stations except for Lac Beniot (CS ID 68) are heavy-tailed with positive excess kurtosis. Tail behavior class was determined for each station according to the discrimination procedure outlined in Figure 3. The majority of the stations (87 out of 126) exhibited class D tail behavior (Table 13 in Appendix II). No correlation between tail behavior class and best-fit distribution model was observed.

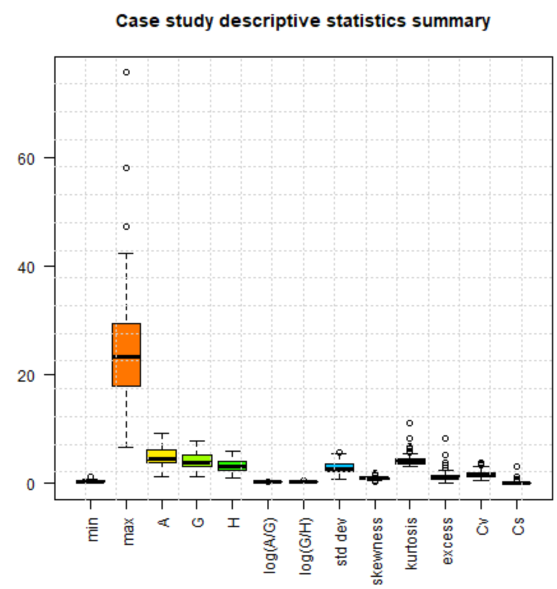

Figure 5: Descriptive statistics of the 126 stations summarized in a box plot

The best distribution fit results according to all metrics used in the case study are summarized in Table 6. The most common best-fit model is the HDF, found best on average in 76% of the stations (26% for HA3 and 38% for HB3). The second to best-fit are the W2 (15% on average) and GG3 (13%). What is significant is that no station data could be fitted to the HIB3 model, hence removed altogether from Figures 6-11. HDF (HA3/HB3) fit was found superior to the most commonly used model for the task of wind resource assessment, i.e. the W2, 73%-100% of the time according to different GOF metrics (see Table 6).

Table 6: Best-fit distribution according to all GOF metrics for all stations of the Eastern Canada case study

	GEV3/ML	**W2/ML**	**HDF/ML**	**N2/ML**	**G2/ML**	**GG3/ML**	**IG2/ML**	**BUR3/ML**	**HDF superior to W2**
LL			46%			10%		44%	95%
AIC			46%			10%		44%	100%
BIC		1%	45%			9%		44%	98%
KS	2%	20%	51%	1%	11%	5%	1%	9%	74%
AD		23%	48%		8%	9%		12%	73%

NRMSE	6%	40%	30%	6%	8%	6%	4%		54%

For simplicity of results assessment, GOF metrics for all stations are summarized in boxplots in Figures 6-11. HB3 corresponds to the minimum among the model candidates in KS boxplot (Figure 9). A better fit of HB3 (minimum of AD in Figure 10) compared to the rest of candidate distributions can be observed for the right tail of the distribution (AD prioritizes the right tail fit).

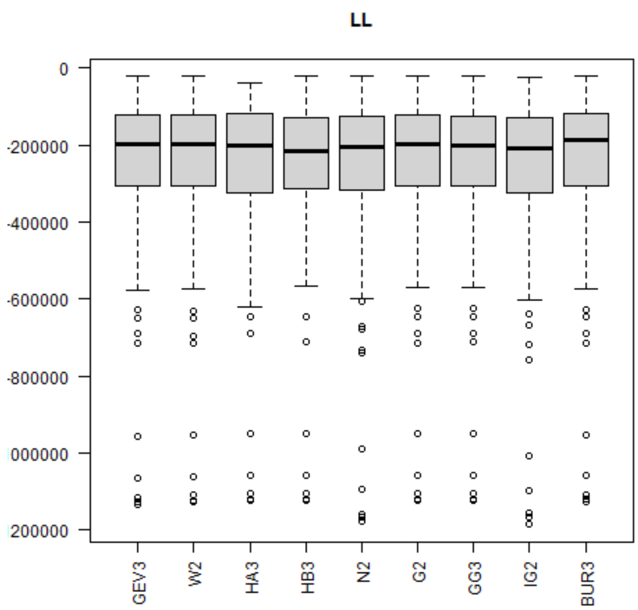

Figure 6: LL box plot

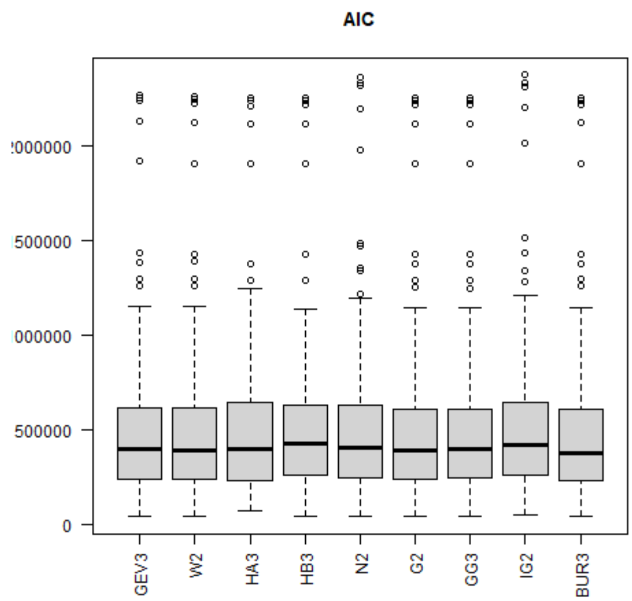

Figure 7: AIC box plot

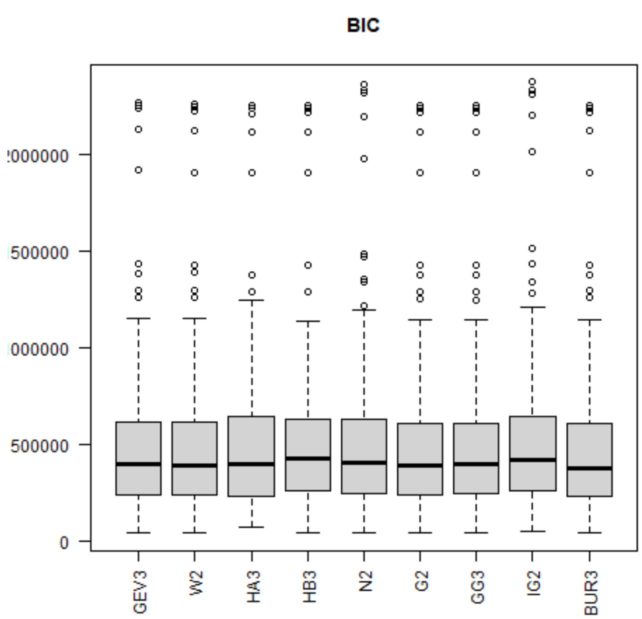

Figure 8: BIC box plot

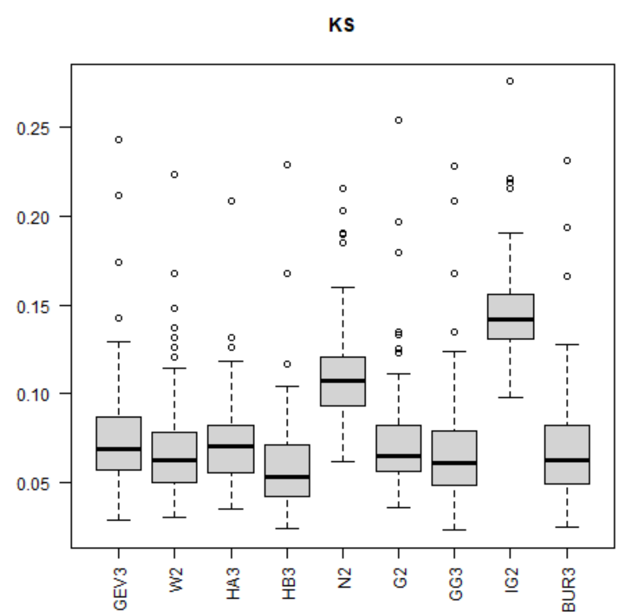

Figure 9: KS box plot

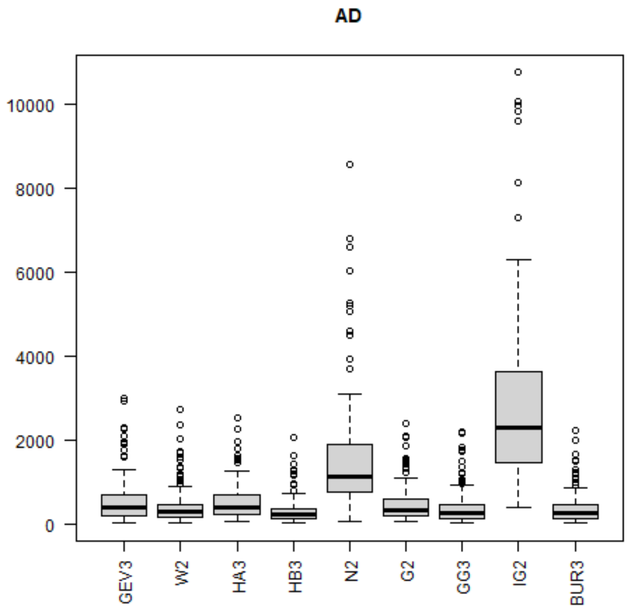

Figure 10: AD box plot

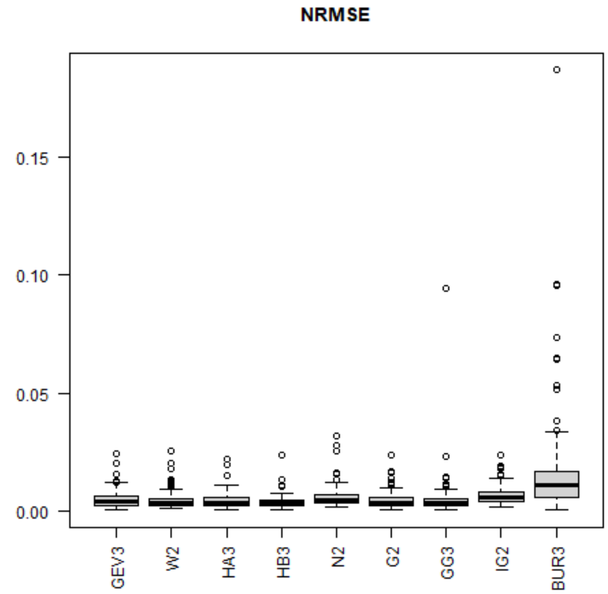

Figure 11: NRMSE box plot

No specific best-fit distributional pattern for onshore/offshore stations or short-term/long-term observation record was observed.

Perreault, et al. [9] showed that "*the Halphen type A distribution has the gamma (G) and the inverse gamma (IG) distributions as limiting forms*" [9]. This can be observed in the moment ratio diagrams of Figures 12-13 plotted with the stations data. The (C_v, C_s) moment ratio diagram given in Figure 12 is not consistent with best-fit results according to any GOF measure. All stations, except for the station with CS ID = 23 (Station ID = 6708) fell into the area of the diagram that relates to HB3.

The (δ_1, δ_2) moment ratio diagram is illustrated in Figure 13. According to the (δ_1, δ_2) moment ratio diagram, all stations are best characterized with a HA3 distribution with a positive v scale parameter. While all the scale v parameters, when their estimation was possible, were indeed found to be positive (see Table 14 in Appendix III), for some stations fitting to HA3 was not at all possible (the procedure of parameter estimation did not converge). This result (all stations fall into HA3 region in Figure 13) is neither consistent with the (C_v, C_s) moment ratio diagram given in Figure 13, nor with the best-fit according to any of the GOF metrics. This finding does not support the finding of Zhang & Singh [14] that moment ratio diagram was a reliable tool for determining the preferred distribution.

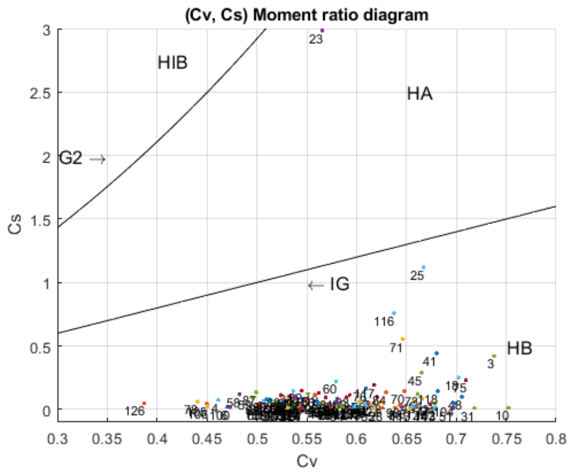

Figure 12: (C_v, C_s) Moment ratio diagram

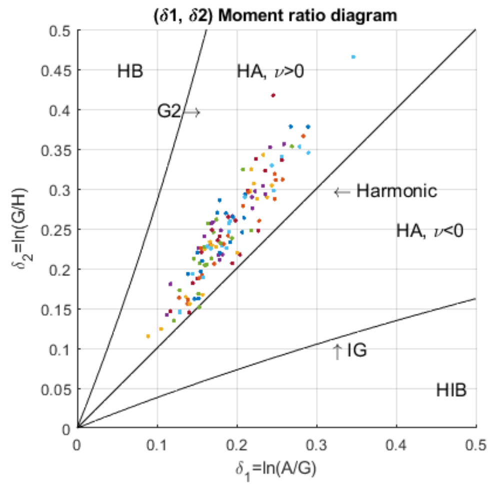

Figure 13: (δ_1, δ_2) Moment ratio diagram

3.1 Results for selected stations and tail behavior

We now look into the details of the cases of wind modelling for two stations for which W2 (station ID = 5251) and HA3 (station ID = 6501) have been determined to be best-fit according to KS. It's worth mentioning that according to log-likelihood related metrics (LL, AIC, and BIC) the top candidate for both 5251 and 6501 is HB3. Both stations, 5251 and 6501, have a long observation record (60 and 59 years respectively).

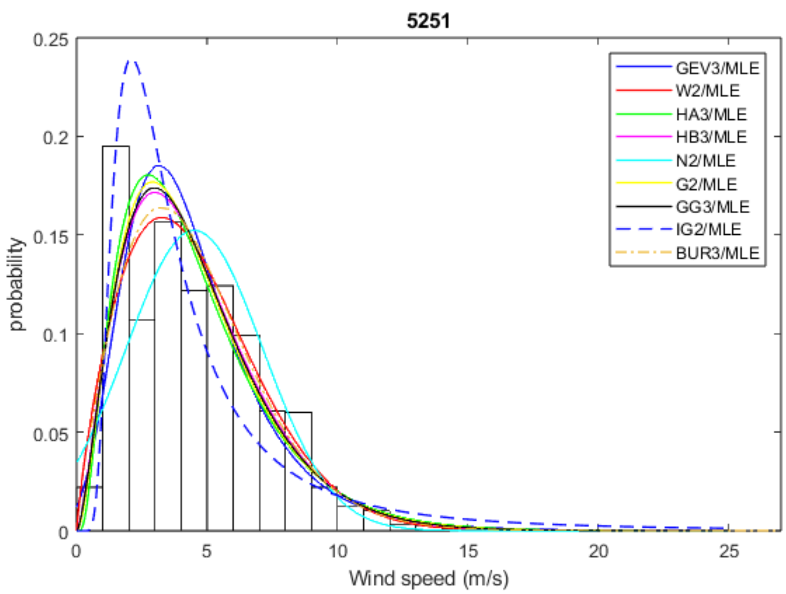

Figure 14: Histogram and fitted models for station ID = 5251 (W2 is best-fit according to KS)

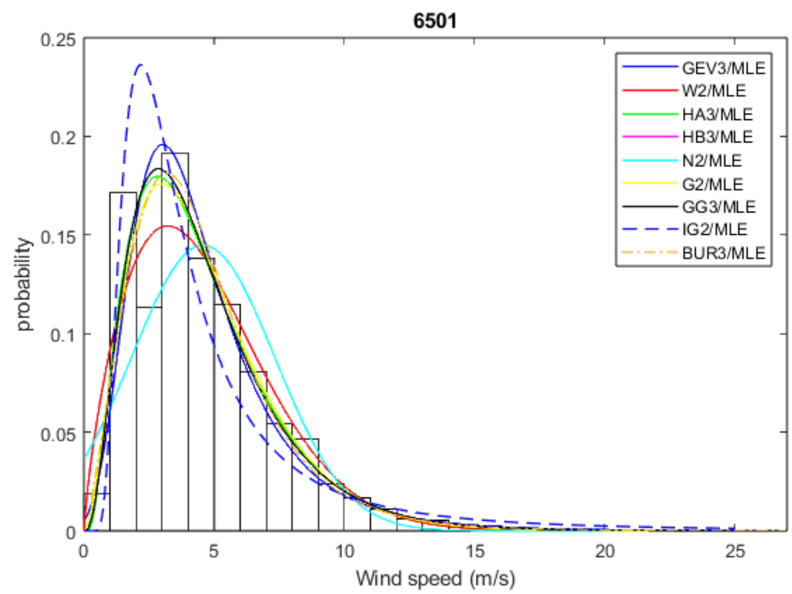

Figure 15: Histogram and fitted models for station ID = 6501 (HA3 is best-fit according to KS)

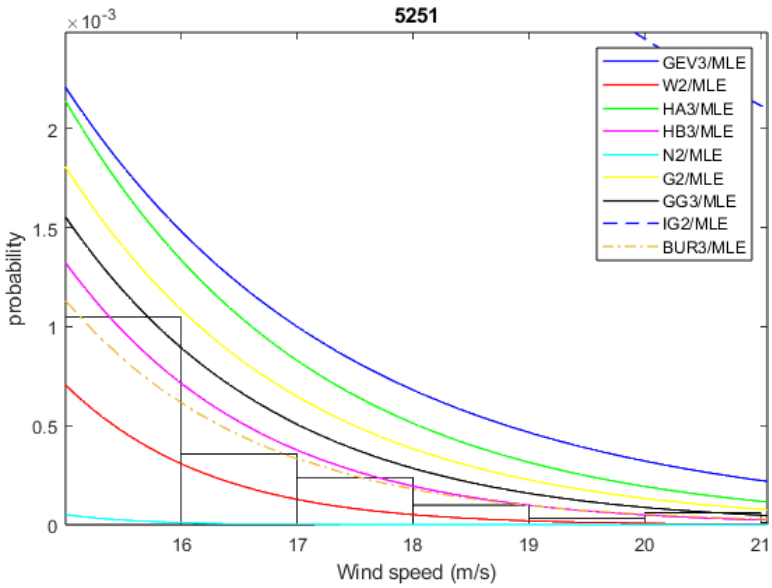

Figure 16: Tail zoom-in of histogram and fitted models for 5251

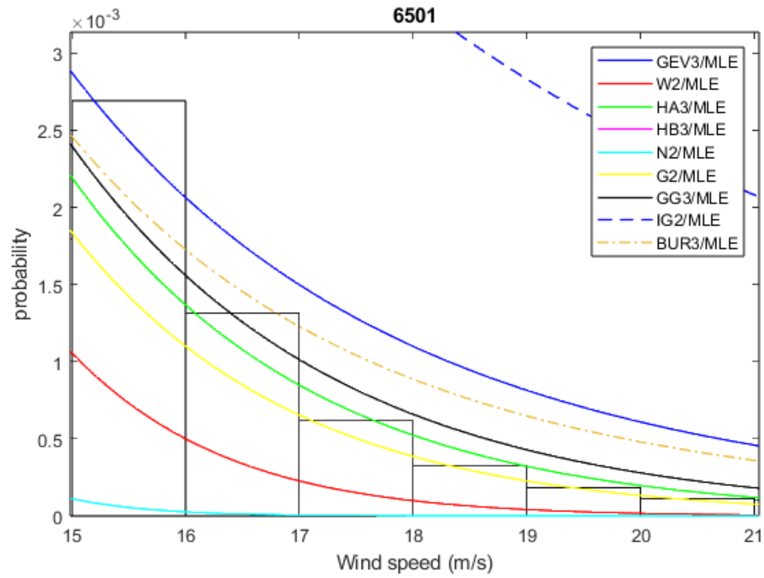

Figure 17: Tail zoom-in of histogram and fitted models for station 6501

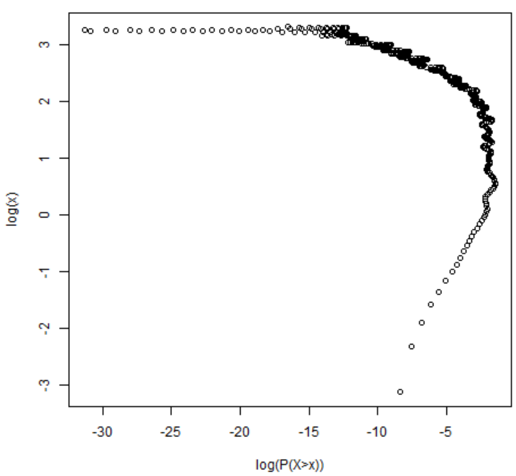

Figure 18: Log-log plot [36] for station ID = 5251 (W2 is best-fit according to KS)

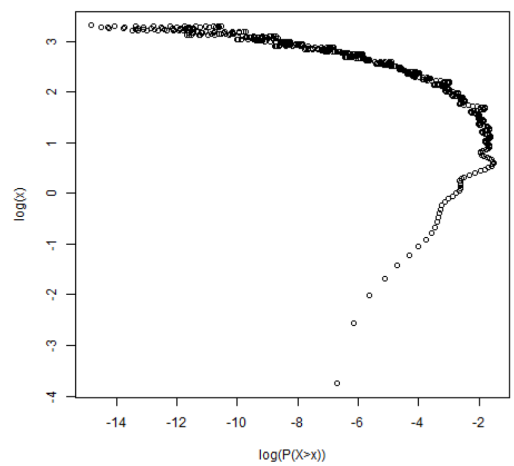

Figure 19: Log-log plot [36] for station ID = 6501 (HA3 is best-fit according to KS)

The log-log curve [36] in Figure 18 is not linear (which it is for Class C), hence we refer to mean excess plot (top right in Figure 20), the slope of which is positive. Hence, the tail behavior is classified as Class D, i.e. sub-exponential distributions (Table 7), according to the decision support system outlined in Figure 3 [40]. The top candidates to model tail behavior of such a sample are HA3, HB3, G2, P3, extreme value type 1 (EV1) or Gumbel distribution and lognormal (LN).

Table 7: Slope of mean excess plot, Shapiro-Wilk (SW) log-normality test, Modified Mann-Kendall (MMK) test, Jackson test statistics and p-values for selected stations (5251, 6501)

ID	**SW** [7] **Log-normality test statistic**	**SW** [7] **Log-normality test p-value**	**MMK** [6] **statistic**	**MK** [4] **p-value**	**MMK** [6] **p-value**	**Jackson`s test** [38] **statistic**	**Jackson`s test** [38] **p-value**	**Slope of mean excess plot**	**Tail behavior**
5251	0.93481231	2.28E-34	-0.2021858	0.02166128	0.12758335	1.56332114	1	2.68440656	Class D

6501	0.97841636	6.03E-21	-0.2513369	0.03794744	0.08424816	1.59586205	1	4.94858319	Class D

Keeping this in mind, we focus on how well HA3 and W2 model mean wind speed, which according to the sample is 4.51 m/s. The turquoise curve in Figure 14 represents the normal distribution with a mean of 4.51 and standard deviation of 2.62 (Table 9). Due to its symmetry, the mean of a normal distribution is also its median (that equally divides the area under the PDF is two) and its mode, i.e. most frequent or expected value. That is not the case for skewed distributions (all other candidate curves in the Figure 14). The red curve corresponds to W2 and the green one – to HA3. While Figure 14 does not answer the question which of the candidates is the best at modelling wind speed in all of its range, Figure 16 shows which model is closer to the histogram (better fit at the tail). The fit in the right tail must be responsible for overall improved fit (Figure 16), where it is clear that the red W2 curve is closer to the histogram bars, rather than the green HA3. The opposite can be noticed in Figure 17 (tail behavior for station 6501, where HA3 dominates).

Table 8: Sample statistics for selected stations (5251, 6501)

ID	Sample minimum, x_{min}	Sample maximum, x_{max}	δ_1	δ_2	Sample mean, μ	Sample median	Sample mode, x_M	Sample standard deviation, s	Sample skewness, $\sqrt{\beta_1}$	Sample kurtosis, β_2	Sample excess kurtosis	Coef. of variation, C_v	Coef. of skewness, C_s
5251	0.56	26.94	0.19	0.23	4.51	4.17	1.67	2.62	0.86	3.79	0.79	1.52	0.05
6501	0.28	26.94	0.18	0.22	4.57	4.17	3.06	2.75	1.3	5.41	2.41	1.66	0.06

In the case of station 6501, the log-log curve [36] in Figure 24 is again not linear, and the slope of mean excess plot (top right in Figure 21) is again positive. Therefore, the tail behavior is also classified as Class D (Table 10) according to decision support system outlined in Figure 3 [40]. HA3 is among the top candidates to model tail behavior, given that is the research question.

Table 9: Some sample statistics estimations based on estimated model parameters

Distribution	Mean, μ	Mode, x_m	Standard deviation, σ
N2($\mu, \boldsymbol{\sigma}$)	μ	μ	σ
HA3($\boldsymbol{a}, \boldsymbol{m}, \boldsymbol{\nu}$)	$\frac{mK_{\nu+1}}{K_\nu}$	$m\left[\frac{\nu-1}{2a}+\sqrt{1+\frac{\nu-1^2}{2a}}\right]$	$\mathrm{m}/K_\nu\sqrt{K_\nu K_{\nu+2}-K_{\nu+1}^2}$
HB3($\boldsymbol{a}, \boldsymbol{m}, \boldsymbol{\nu}$)	$\frac{mef_{\nu+1/2}(\alpha)}{ef_\nu(\alpha)}$	-	$\frac{m^2(ef_\nu ef_{\nu+1}-ef_{\nu+1/2}^2)}{ef_\nu^2}$
W2($\boldsymbol{k}, \boldsymbol{\lambda}$)	$\lambda\Gamma(1+1/k)$	$\lambda(k-1/k)^{1/k}$	$\lambda\sqrt{\Gamma\left(1-\frac{2}{k}\right)-\Gamma\left(1+\frac{1}{k}\right)^2}$

For a mindful comparison of PDFs in Figures 14-15, let's compute the mode, mean and standard deviation of some distributions of particular interest for the selected stations. A shortcut for the calculation is given in Table 9.

Table 10: Estimated parameters and moments for HA3, HB3, W2, GG3, and N2 distributions for selected stations (5251, 6501)

CS station ID	Station ID	Station name	Province	HA3(α, *m, v*)	HB3(α, *m, v*)	W2(k, λ)	N2(μ, σ)
1	5251	QUEBEC/JEAN LESAGE INTL A	QUEBEC	α = 0.5916372 m = 1.059223 v = 2.309711	α = -5.276964 m = 11.20424 v = 1.279034	λ = 5.091835 k = 1.81496 $\widehat{x_M}$ = 3.2755	μ = $\widehat{x_M}$ = 4.510853

				$\widehat{x_M}$ = 3.404 μ = 3.6375 σ = 2.6859	μ = 4.5109 σ = 5.2166	μ = 4.53 σ = 0	σ = 2.620128
19	6501	WESTERN HEAD	NOVA SCOTIA	α = 0.6325211 m = 1.126856 v = 2.331912 $\widehat{x_M}$ = 3.4997 μ = 8.0118 σ = 9.6874	-	λ = 5.157046 k = 1.767681 $\widehat{x_M}$ = 3.2173 μ = 4.5904 σ = 0	μ = $\widehat{x_M}$ = 4.568837 σ = 2.754406

The results of the calculation of the first and second non-central moments based on formulae given in Table 9 are presented among with the values of the estimated model parameters in Table 10.

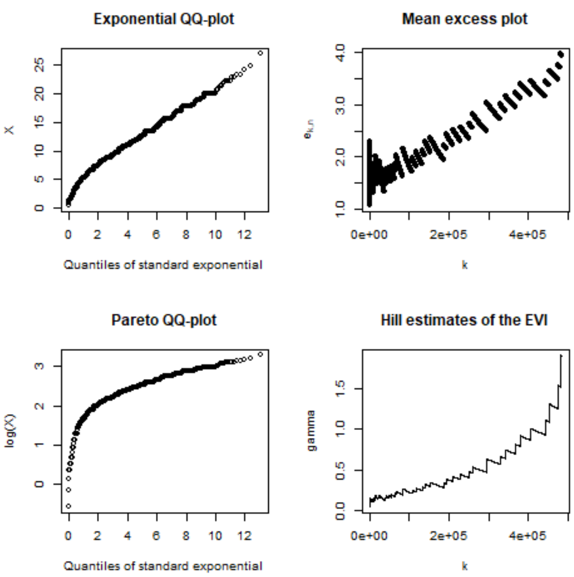

Figure 20: Exponential QQ plot, mean excess plot, Pareto QQ plot, Hill estimate of the extreme value indices (EVI) plot [39] for station ID = 5251 (W2 is best-fit according to KS)

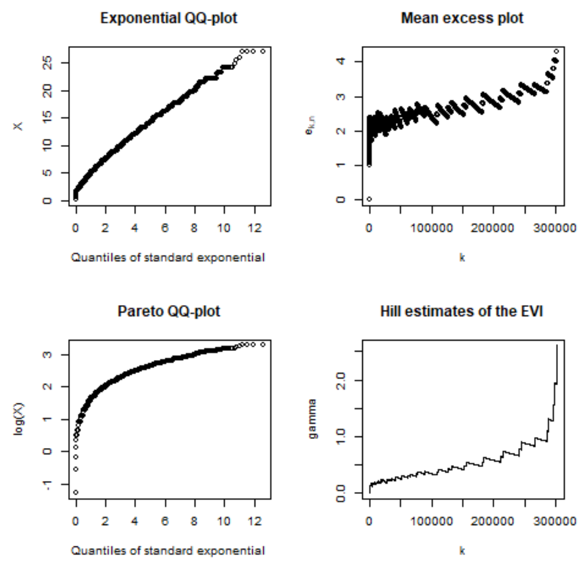

Figure 21 Exponential QQ plot, mean excess plot, Pareto QQ plot, Hill estimate of the EVI plot [39] for station ID = 6501 (HA is best-fit according to KS)

The exponential QQ-plots of Figures 20 and 21 (top left) are curves, but quite close to the line of slope 1, suggesting using the exponential distribution as a candidate model. The mean excess plots also support this argument, as they are not converging in both figures, while they are constant for exponential tail behavior [2]. On the other hand, Jackson's test [38] of exponentiality results with p-values of 1 for both stations suggests that there is not enough evidence to reject the null hypothesis of exponentiality. The Pareto-QQ plots of both figures confirm that the samples from selected stations do not exhibit Pareto type tail behavior (Class B). If they did, we would expect to see a line with a slope of 1 there instead of the curve. The Hill's ratio plots of both figures also do not converge to a constant, which suggests that Class C is not applicable for modelling tail behavior, because Hill's ratio is a consistent estimator of α in case when "*the tail is regularly varying (Class C) with tail index α*" [2]. The log-log plots support this as well, as they are linear for Class C, which is not the case for Figures 18 and 19. The Shapiro-Wilk [7] normality test of logarithm-transformed sample data rejects the null hypothesis of log normality (Class Lognormal) for any confidence level. This is the reason lognormal distributions were not originally considered among the candidates in the case study.

4 Conclusions and future work

In this research paper, the application of the HDF [65] was introduced to model mean wind speed. Long-term measured wind speed data sets of 126 stations in the East of Canada were examined in the case study. The effectiveness of the HDF was assessed against the W2 distribution, and 9 other distributions commonly used for mean wind speed modelling. The measures of effectiveness or GOF used in the case study included the LL, AIC, BIC, KS, AD statistics, and the NRMSE. The GOF results show favorable results of the HDF performance for the task, which answers the main research question of the paper.

The most common best-fit model is the HDF, found best in 30-51% of the stations according to different GOF metrics (Table 6). The second to best-fit is the three-parameter Burr model (best-fit in 9-44% of the stations according to different GOF metrics) and GG3 (13% on average). What is significant is that no station data could be fitted to the HIB3 model.

HDF (HA3/HB3) fit was found superior to the most commonly used model for the task of wind resource assessment, i.e. W2, 54%-100% of the time according to different GOF metrics (see Table 6). HA3 and HB3 are recommended to be model candidates in WRE studies.

Modelling maximum wind speed for the extended case study with the HDF is planned to be carried out in subsequent study, as AD, the GOF favoring right tail fit, showed improved fit for HB3. The case study can also be extended by adding other distribution parameter estimation methods apart from the MLE used in this case study. Dvorak et al. [8] points out that HDF does not seem to be *“as rich in a variety of forms and types as the Pearson system and they do not fulfill some particular requirements as ease of computation”* [8]. Future efforts should also focus on the continuation of this case study with the goal of comparing HDF with the Pearson distribution system.

Data availability

The data used in the case study are provided in [45] for reproducibility of results. Possible extensions of this case study are encouraged.

Declaration of competing interest

There is no conflict of interest regarding this publication.

CRediT authorship contribution statement

Olga Tsvetkova: Methods, Writing - Original draft preparation.
Taha B.M.J. Ouarda: Idea & Conceptualization, Methods, Funding acquisition, Supervision, Writing - Reviewing and Editing.

References

[1] F. Chebana, S. El Adlouni, and B. Bobée, "Mixed estimation methods for Halphen distributions with applications in extreme hydrologic events."

[2] S. El Adlouni and B. Bobée, *Halphen distribution family with application in hydrological frequency analysis*. Water Resources Publications, LLC, 2017.

[3] National Bureau of Standards, *Handbook of mathematical functions with formulas, graphs and mathematical tables*, 10th Print. United States Department of Commerce, 1972.

[4] H. B. Mann, "Nonparametric Tests Against Trend," *Econometrica*, vol. 13, no. 3, p. 245, Jul. 1945, doi: 10.2307/1907187.

[5] "Package 'modifiedmk' Title Modified Versions of Mann Kendall and Spearman's Rho Trend Tests," 2021, doi: 10.1017/CBO9781107415324.004.

[6] K. H. Hamed and A. Ramachandra Rao, "A modified Mann-Kendall trend test for autocorrelated data," *J. Hydrol.*, vol. 204, no. 1–4, pp. 182–196, Jan. 1998, doi: 10.1016/S0022-1694(97)00125-X.

[7] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3–4, pp. 591–611, Dec. 1965, doi: 10.1093/BIOMET/52.3-4.591.

[8] F. Dvorak, V., Bobée, B., Boucher, S., Ashkar, "Halphen distributions and related systems of frequency functions," Quebec, 1988.

[9] L. Perreault, B. Bobée, and P. F. Rasmussen, "Halphen Distribution System. I: Mathematical and Statistical Properties," *J. Hydrol. Eng.*, vol. 4, no. 3, pp. 189–199, Jul. 1999, doi: 10.1061/(ASCE)1084-0699(1999)4:3(189).

[10] L. Perreault, B. Bobée, and P. F. Rasmussen, "Halphen Distribution System. II: Parameter and Quantile Estimation Estimation of extreme precipitations View project Statistical post-processing of ensemble weather forecasts View project," *Artic. J. Hydrol. Eng.*, 1999, doi: 10.1061/(ASCE)1084-0699(1999)4:3(200).

[11] F. Chebana, S. El Adlouni, and B. Bobée, "Method of moments of the Halphen distribution parameters," *Stoch. Environ. Res. Risk Assess.*, vol. 22, no. 6, pp. 749–757, 2008, doi: 10.1007/S00477-007-0184-4.

[12] S. El Adlouni and B. Bobée, "Sampling Techniques for Halphen Distributions," *J. Hydrol. Eng.*, vol. 12, no. 6, pp. 592–604, Nov. 2007, doi: 10.1061/(ASCE)1084-0699(2007)12:6(592).

[13] Y. Hundecha, A. St-Hilaire, T. B. M. J. Ouarda, S. El Adlouni, and P. Gachon, "A nonstationary extreme value analysis for the assessment of changes in extreme annual wind speed over the gulf of St. Lawrence Canada," *J. Appl. Meteorol. Climatol.*, vol. 47, no. 11, pp. 2745–2759, Nov. 2008, doi: 10.1175/2008JAMC1665.1.

[14] L. Zhang and V. P. Singh, "Revisiting the Application of Halphen Distributions in Flood Frequency Analysis," *J. Hydrol. Eng.*, vol. 26, no. 12, p. 04021042, Oct. 2021, doi: 10.1061/(ASCE)HE.1943-5584.0002133.

[15] S. El Adlouni, F. Chebana, and B. Bobée, "Generalized extreme value versus Halphen system: Exploratory study," *J. Hydrol. Eng.*, vol. 15, no. 2, pp. 79–89, 2010, doi: 10.1061/(ASCE)HE.1943-5584.0000152.

[16] R. Delhome, R. Billot, and N. E. El Faouzi, "Travel time statistical modeling with the Halphen distribution family," *J. Intell. Transp. Syst. Technol. Planning, Oper.*, vol. 21, no. 6, pp. 452–464, Nov. 2017, doi: 10.1080/15472450.2017.1326115.

[17] C. Jung and D. Schindler, "Wind speed distribution selection – A review of recent development and progress," *Renewable and Sustainable Energy Reviews*, vol. 114. Elsevier Ltd, p. 109290, Oct. 01, 2019, doi: 10.1016/j.rser.2019.109290.

[18] J. A. A. Carta, P. Ramírez, and S. Velázquez, "A review of wind speed probability distributions used in wind energy analysis. Case studies in the Canary Islands," *Renew. Sustain. Energy Rev.*, vol. 13, no. 5, pp. 933–955, Jun. 2009, doi: 10.1016/j.rser.2008.05.005.

[19] T. B. M. J. Ouarda and C. Charron, “Distributions of wind speed in a northern environment,” in *2018 9th International Renewable Energy Congress, IREC 2018*, May 2018, pp. 1–3, doi: 10.1109/IREC.2018.8362453.

[20] J. Zhou, E. Erdem, G. Li, and J. Shi, “Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites,” *Energy Convers. Manag.*, vol. 51, no. 7, pp. 1449–1458, Jul. 2010, doi: 10.1016/j.enconman.2010.01.020.

[21] L. Wang, J. Liu, and F. Qian, “Wind speed frequency distribution modeling and wind energy resource assessment based on polynomial regression model,” *Int. J. Electr. Power Energy Syst.*, vol. 130, p. 106964, Sep. 2021, doi: 10.1016/j.ijepes.2021.106964.

[22] H. Shi, Z. Dong, N. Xiao, and Q. Huang, “Wind Speed Distributions Used in Wind Energy Assessment: A Review,” *Front. Energy Res.*, vol. 9, p. 790, Nov. 2021, doi: 10.3389/FENRG.2021.769920/BIBTEX.

[23] N. Aries, S. M. Boudia, and H. Ounis, “Deep assessment of wind speed distribution models: A case study of four sites in Algeria,” *Energy Convers. Manag.*, vol. 155, pp. 78–90, Jan. 2018, doi: 10.1016/j.enconman.2017.10.082.

[24] I. Usta and Y. M. Kantar, “Analysis of some flexible families of distributions for estimation of wind speed distributions,” *Appl. Energy*, vol. 89, no. 1, pp. 355–367, Jan. 2012, doi: 10.1016/J.APENERGY.2011.07.045.

[25] J.-Y. Shin, T. B. M. J. Ouarda, and T. Lee, “Heterogeneous mixture distributions for modeling wind speed, application to the UAE,” *Renew. Energy*, vol. 91, pp. 40–52, Jun. 2016.

[26] S. El Adlouni, F. Chebana, and B. Bobée, “Generalized Extreme Value versus Halphen System: Exploratory Study,” *J. Hydrol. Eng.*, vol. 15, no. 2, pp. 79–89, Feb. 2010, doi: 10.1061/(ASCE)HE.1943-5584.0000152.

[27] MathWorks, “Matlab R2018b.” [Online]. Available: https://www.mathworks.com/products/new_products/release2018b.html.

[28] “Package ‘GPBayes’ Title Tools for Gaussian Process Modeling in Uncertainty Quantification,” 2022, Accessed: Sep. 07, 2022. [Online]. Available: https://github.com/pulongma/GPBayes/issues.

[29] “R Core Team, R: A Language and Environment for Statistical Computing.” R Foundation for Statistical Computing, Vienna, Austria, 2020, [Online]. Available: https://www.r-project.org/.

[30] M. K. Robin S Hankin, “Package ‘hypergeo’ Title The Gauss Hypergeometric Function,” 2016.

[31] G. Morlat, “Les lois de probabilités de Halphen,” 1956. Accessed: Jul. 15, 2022. [Online]. Available: http://www.sfds.asso.fr/publicat/rsa.htm.

[32] L. Perreault, B. Bobée, and P. F. Rasmussen, “Halphen Distribution System. I: Mathematical and Statistical Properties,” *J. Hydrol. Eng.*, vol. 4, no. 3, pp. 189–199, Jul. 1999, doi: 10.1061/(ASCE)1084-0699(1999)4:3(189).

[33] T. Werner and C. Upper, “Time Variation in the Tail Behavior of Bund Future Returns,” *J. Futur. Mark.*, vol. 24, no. 4, pp. 387–398, 2004, doi: 10.1002/fut.10120.

[34] T. Werner and C. Upper, “Time Variation in the Tail Behaviour of Bunds Futures Returns,” *SSRN Electron. J.*, Dec. 2002, doi: 10.2139/SSRN.362320.

[35] S. El Adlouni, B. Bobée, and T. B. M. J. Ouarda, “On the tails of extreme event distributions in hydrology,” *J. Hydrol.*, vol. 355, no. 1–4, pp. 16–33, Jun. 2008, doi: 10.1016/J.JHYDROL.2008.02.011.

[36] J. Beirlant, T. De Wet, and Y. Goegebeur, “A goodness-of-fit statistic for Pareto-type behaviour,” *J. Comput. Appl. Math.*, vol. 186, no. 1, pp. 99–116, Feb. 2006, doi: 10.1016/J.CAM.2005.01.036.

[37] P. Embrechts, C. Klüppelberg, and T. Mikosch, “Modelling Extremal Events,” *Model. Extrem. Events*, 1997, doi: 10.1007/978-3-642-33483-2.

[38] O. A. Y. Jackson, “An Analysis of Departures from the Exponential Distribution,” *J. R. Stat. Soc. Ser. B*, vol. 29, no. 3, pp. 540–549, Sep. 1967, doi: 10.1111/J.2517-6161.1967.TB00716.X.

[39] B. M. Hill, “A Simple General Approach to Inference About the Tail of a Distribution,” *https://doi.org/10.1214/aos/1176343247*, vol. 3, no. 5, pp. 1163–1174, Sep. 1975, doi: 10.1214/AOS/1176343247.

[40] E. Adlouni, B. Bobée, and O. Samoud, “GUIDE FOR THE USE OF THE DECISION SUPPORT SYSTEM,” no. C, pp. 1–13, 2018.

[41] B. Bobée and F. Ashkar, “The gamma family and derived distributions applied in hydrology,” p. 203, 1991.

[42] T. B. M. J. Ouarda, C. Charron, and F. Chebana, “Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and L-moment ratio diagram methods, with a case study,” *Energy Convers. Manag.*, vol. 124, pp. 247–265, 2016.

[43] H. Akaike, “Statistical predictor identification,” *Ann. Inst. Stat. Math. 1970 221*, vol. 22, no. 1, pp. 203–217, Dec. 1970, doi: 10.1007/BF02506337.

[44] S. G, “Estimating the Dimension of a Model,” *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, 1978, doi: 10.2307/2958889.

[45] O. Tsvetkova, “Eastern Canada hourly wind speed dataset 126 stations,” 2022. .

[46] “Government of Canada,” *Historic data*. https://climat.meteo.gc.ca/historical_data/search_historic_data_e.html (accessed Oct. 02, 2019).

[47] O. Seidou, J. J. Asselin, and T. B. M. J. Ouarda, “Bayesian multivariate linear regression with application to change point models in hydrometeorological variables,” *Water Resour. Res.*, vol. 43, no. 8, p. 8401, Aug. 2007, doi: 10.1029/2005WR004835.

[48] T. B. M. J. Ouarda *et al.*, “Probability distributions of wind speed in the UAE,” *Energy Convers. Manag.*, vol. 93, pp. 414–434, Mar. 2015.

[49] N. Ayuketang Arreyndip and E. Joseph, “Generalized Extreme Value Distribution Models for the Assessment of Seasonal Wind Energy Potential of Debuncha, Cameroon,” *J. Renew. Energy*, vol. 2016, pp. 1–9, 2016, doi: 10.1155/2016/9357812.

[50] T. B. M. J. Ouarda and C. Charron, “On the mixture of wind speed distribution in a Nordic region,” *Energy Convers. Manag.*, vol. 174, no. August, pp. 33–44, 2018.

[51] H. Ounis and N. Aries, “On the wind resource in Algeria: Probability distributions evaluation:,” *https://doi.org/10.1177/0957650920975883*, vol. 235, no. 5, pp. 1187–1204, Nov. 2020, doi: 10.1177/0957650920975883.

[52] O. Tsvetkova and T. B. M. J. Ouarda, “A review of sensitivity analysis practices in wind resource assessment,” *Energy Convers. Manag.*, vol. 238, p. 114112, Jun. 2021, doi: 10.1016/j.enconman.2021.114112.

[53] K. S. R. Murthy and O. P. Rahi, “A comprehensive review of wind resource assessment,” *Renewable and Sustainable Energy Reviews*, vol. 72. Elsevier Ltd, pp. 1320–1342, 2017.

[54] H. E. AKYUZ and H. GAMGAM, “Statistical Analysis of Wind Speed Data with Weibull, Lognormal and Gamma Distributions,” *Cumhur. Sci. J.*, vol. 38, no. 4, pp. 68–76, Dec. 2017, doi: 10.17776/CSJ.358773.

[55] Z. H. Hulio, W. Jiang, and S. Rehman, “Techno-Economic assessment of wind power potential of Hawke’s Bay using Weibull parameter: A review,” *Energy Strateg. Rev.*, vol. 26, p. 100375, Nov. 2019.

[56] M. Sumair, T. Aized, S. A. R. Gardezi, S. U. ur Rehman, and S. M. S. Rehman, “A novel method developed to estimate Weibull parameters,” *Energy Reports*, vol. 6, pp. 1715–1733, Nov. 2020, doi: 10.1016/j.egyr.2020.06.017.

[57] K. Azad, M. Rasul, P. Halder, and J. Sutariya, “Assessment of wind energy prospect by Weibull distribution for prospective wind sites in Australia,” in *Energy Procedia*, 2019, vol. 160, pp. 348–355.

[58] J. C. Aririguzo and E. B. Ekwe, “Weibull distribution analysis of wind energy prospect for Umudike, Nigeria for power generation,” *Robot. Comput. Integr. Manuf.*, vol. 55, pp. 160–163, Feb. 2019, doi:

10.1016/J.RCIM.2018.01.001.

[59] F. Veronesi, S. Grassi, and M. Raubal, “Statistical learning approach for wind resource assessment,” *Renew. Sustain. Energy Rev.*, vol. 56, pp. 836–850, Apr. 2016, doi: 10.1016/J.RSER.2015.11.099.

[60] I. Tizgui, F. El Guezar, H. Bouzahir, and B. Benaid, “Wind speed distribution modeling for wind power estimation: Case of Agadir in Morocco:,” *https://doi.org/10.1177/0309524X18780391*, vol. 43, no. 2, pp. 190–200, Jun. 2018, doi: 10.1177/0309524X18780391.

[61] D. Villanueva and A. Feijóo, “Wind power distributions: A review of their applications,” *Renew. Sustain. Energy Rev.*, vol. 14, no. 5, pp. 1490–1495, 2010, doi: 10.1016/j.rser.2010.01.005.

[62] B. Safari, “Modeling wind speed and wind power distributions in Rwanda,” *Renew. Sustain. Energy Rev.*, vol. 15, no. 2, pp. 925–935, Feb. 2011, doi: 10.1016/J.RSER.2010.11.001.

[63] N. Masseran, A. M. Razali, and K. Ibrahim, “An analysis of wind power density derived from several wind speed density functions: The regional assessment on wind power in Malaysia,” *Renew. Sustain. Energy Rev.*, vol. 16, no. 8, pp. 6476–6487, Oct. 2012, doi: 10.1016/J.RSER.2012.03.073.

[64] I. Mert and C. Karakuş, “A statistical analysis of wind speed data using Burr, generalized gamma, and Weibull distributions in Antakya, Turkey,” *Turkish J. Electr. Eng. Comput. Sci.*, vol. 23, no. 6, pp. 1571–1586, Jan. 2015, doi: 10.3906/elk-1402-66.

[65] O. Tsvetkova and T. B. M. J. Ouarda, “Quasi-Monte Carlo technique in global sensitivity analysis of wind resource assessment with a study on UAE,” *J. Renew. Sustain. Energy*, vol. 11, no. 5, 2019, doi: 10.1063/1.5120035.

Appendix I: Case study stations overview

Table 11: Case study stations overview

CS ation ID	Station ID	Station name	Province	Latitude	Longitude	Elevati on, m.	First Year	Last Year	Years collec ted	Number of data points
1	5251	QUEBEC/JEAN LESAGE INTL A	QUEBEC	46.8	-71.38	74.4	1953	2013	60	485707
2	5293	VALCARTIER A	QUEBEC	46.9	-71.5	167.6	1978	2018	40	72439
3	5415	MONTREAL/PIERRE ELLIOTT TRUDEAU INTL A	QUEBEC	45.47	-73.75	35.97	1953	2013	60	502624
4	5530	SHERBROOKE A	QUEBEC	45.43	-71.68	241.4	1962	2005	43	320292
5	5668	BLANC SABLON	QUEBEC	51.42	-57.22	19.2	1967	1982	15	78840
6	5669	LOURDES DE BLANC SABLON A	QUEBEC	51.45	-57.18	37.18	1970	2014	44	220175
7	5707	MANICOUAGAN A	QUEBEC	50.65	-68.83	406.3	1961	1971	10	29826
8	5839	RIVIERE DU LOUP	QUEBEC	47.8	-69.55	148.4	1965	1980	15	121201
9	6029	CHIB3OUGAMAU CHAPAIS A	QUEBEC	49.77	-74.53	387.1	1982	2014	32	199814
10	6047	LA GRANDE RIVIERE A	QUEBEC	53.63	-77.7	195.1	1976	2012	36	304160
11	6053	MATAGAMI A	QUEBEC	49.77	-77.82	281.3	1973	1991	18	112387
12	6058	NITCHEQUON	QUEBEC	53.2	-70.9	536.1	1953	1985	32	235079
13	6081	VAL-D'OR A	QUEBEC	48.06	-77.79	337.4	1955	2012	57	463422
14	6089	CAPE HOPES ADVANCE	QUEBEC	61.08	-69.55	73.2	1953	1971	18	46645
15	6205	MISCOU ISLAND (AUT)	NEW BRUNSWICK	48.01	-64.49	4	1964	2018	54	291390
16	6256	ST LEONARD A	NEW BRUNSWICK	47.16	-67.83	242.2	1985	2014	29	153857
17	6330	COPPER LAKE	NOVA SCOTIA	45.38	-61.97	96.9	1953	1974	21	97014
18	6335	DEBERT A	NOVA SCOTIA	45.42	-63.45	44.2	1953	1960	7	65086
19	6501	WESTERN HEAD	NOVA SCOTIA	43.99	-64.66	10.06	1959	2018	59	301754
20	6557	ARGENTIA A	NEWFOUNDLAND	47.3	-54	15.5	1976	1986	10	49995
21	6581	BUCHANS A	NEWFOUNDLAND	48.85	-56.83	276.1	1953	1965	12	55225
22	6584	BURGEO 2	NEWFOUNDLAND	47.62	-57.62	10.6	1994	2006	12	94017
23	6708	ST ALBANS	NEWFOUNDLAND	47.87	-55.85	13.4	1968	1983	15	79823
24	6709	ST ANDREWS	NEWFOUNDLAND	47.77	-59.33	10.7	1953	1966	13	60136
25	6710	ST ANTHONY	NEWFOUNDLAND	51.37	-55.58	17.4	1953	1965	12	54653
26	6711	ST ANTHONY	NEWFOUNDLAND	51.37	-55.6	11.5	1970	1996	26	128993
27	6723	ST LAWRENCE	NEWFOUNDLAND	46.92	-55.38	48.5	1966	1996	30	180363
28	6740	STEPHENVILLE A	NEWFOUNDLAND	48.53	-58.55	24.7	1953	2014	61	439895
29	6756	TWILLINGATE	NEWFOUNDLAND	49.68	-54.82	4.9	1953	1967	14	68357
30	6773	CARTWRIGHT	NEWFOUNDLAND	53.71	-57.04	14.3	1953	2015	62	442305
31	6778	MAKKOVIK A	NEWFOUNDLAND	55.08	-59.19	70.4	1985	2015	30	119736
32	6783	MARY'S HARBOUR A	NEWFOUNDLAND	52.3	-55.83	10.6	1983	2013	30	116606
33	6797	SAGLEK	NEWFOUNDLAND	58.33	-62.59	516	1989	2018	29	145767

34	6916	BATHURST A	NEW BRUNSWICK	47.63	-65.75	58.8	1994	2013	19	110852
35	6918	BUCTOUCHE CDA CS	NEW BRUNSWICK	46.43	-64.77	35.9	2005	2018	13	110505
36	6923	KEJIMKUJIK 1	NOVA SCOTIA	44.4	-65.2	125	1994	2018	24	157471
37	7026	ST STEPHEN (AUT)	NEW BRUNSWICK	45.22	-67.25	26.1	1977	2006	29	121782
38	7162	TRURO (AUT)	NOVA SCOTIA	45.37	-63.27	39.9	1994	2005	11	75524
39	7169	MCNABS ISLAND (AUT)	NOVA SCOTIA	44.6	-63.53	15.4	1999	2018	19	151731
40	7173	HART ISLAND (AUT)	NOVA SCOTIA	45.35	-60.98	8.2	1994	2018	24	129896
41	8456	CHEVERY	QUEBEC	50.47	-59.64	7.7	1994	2018	24	142423
42	8675	HEATH POINT	QUEBEC	49.09	-61.7	7	1994	2018	24	95265
43	8989	ILE D'ORLEANS	QUEBEC	47	-70.81	3.6	1994	2018	24	17078
44	8991	SAGONA ISLAND	NEWFOUNDLAND	47.37	-55.79	59.7	1994	2017	23	94300
45	8999	CAP-MADELEINE	QUEBEC	49.25	-65.32	29	1994	2018	24	132126
46	9002	EASTMAIN A	QUEBEC	52.23	-78.52	7.32	1992	2014	22	39448
47	10226	SALLUIT A	QUEBEC	62.18	-75.67	227.08	1969	2015	46	61606
48	10252	KANGIQSUJUAQ A	QUEBEC	61.59	-71.93	157.58	1992	2015	23	55420
49	10661	NORTHEAST MARGAREE (AUT)	NOVA SCOTIA	46.37	-60.98	54	1994	2018	24	40984
50	10684	KANGIQSUALUJJUAQ A	QUEBEC	58.71	-65.99	66.14	1992	2015	23	63006
51	10763	ILES DE LA MADELEINE	QUEBEC	47.43	-61.77	7.6	1993	2018	25	140456
52	10789	AKULIVIK A	QUEBEC	60.82	-78.15	23.16	1992	2015	23	64075
53	10792	GRAND ETANG	NOVA SCOTIA	46.55	-61.05	12.8	1994	2018	24	206943
54	10797	MISTOOK	QUEBEC	48.6	-71.72	112.5	1994	2018	24	68093
55	10800	SUMMERSIDE	PRINCE EDWARD ISLAND	46.44	-63.84	12.2	1994	2018	24	193635
56	10813	NAIN	NEWFOUNDLAND	56.55	-61.68	7.6	1994	2018	24	104903
57	10817	POOLS ISLAND	NEWFOUNDLAND	49.11	-53.58	19.3	1994	2018	24	188737
58	10818	GRATES COVE	NEWFOUNDLAND	48.17	-52.94	46.2	1994	2017	23	174062
59	10836	CAPE KAKKIVIAK	NEWFOUNDLAND	59.99	-64.17	551.1	1994	2018	24	94188
60	10841	CAPE KIGLAPAIT	NEWFOUNDLAND	57.14	-61.48	834.1	1994	2017	23	119333
61	10855	AUPALUK A	QUEBEC	59.3	-69.6	36.9	1992	2015	23	40788
62	10856	TASIUJAQ A	QUEBEC	58.67	-69.95	36.9	1993	2015	22	49777
63	10859	BRIER ISLAND	NOVA SCOTIA	44.29	-66.35	15.8	1969	2018	49	206693
64	10869	CAP-TOURMENTE	QUEBEC	47.08	-70.78	6	1994	2018	24	146543
65	10873	STE-ANNE-DE-BELLEVUE 1	QUEBEC	45.43	-73.93	39	1994	2018	24	57376
66	10909	MARTICOT ISLAND	NEWFOUNDLAND	47.33	-54.59	21.8	1994	2018	24	57922
67	10945	SHEARWATER JETTY	NOVA SCOTIA	44.63	-63.52	5.5	1994	2018	24	166933
68	26778	LAC BENOIT	QUEBEC	51.53	-71.11	549	1994	2018	24	90186
69	26797	CHAMOUCHOUANE	QUEBEC	49.28	-73.36	303.8	1994	2018	24	39032
70	26893	LA POCATIERE	QUEBEC	47.36	-70.03	31	1996	2018	22	48800
71	26968	KOUCHIB3OUGUAC CS	NEW BRUNSWICK	46.78	-65.02	21	2005	2018	13	56252

72	27115	ST JOHN'S WEST CDA CS	NEWFOUNDLAND	47.52	-52.78	114	1999	2013	14	115310
73	27195	CAP-CHAT	QUEBEC	49.11	-66.65	5	1996	2018	22	135316
74	27621	ST LAWRENCE	NEWFOUNDLAND	46.92	-55.38	48.5	1998	2006	8	66884
75	27712	WINTERLAND	NEWFOUNDLAND	47.14	-55.33	29.26	1999	2018	19	135268
76	27846	CONFEDERATION BRIDGE	PRINCE EDWARD ISLAND	46.23	-63.73	54.1	1999	2006	7	37758
77	29471	WRECKHOUSE	NEWFOUNDLAND	47.71	-59.31	31.7	2000	2006	6	55810
78	29493	GAGETOWN AWOS A	NEW BRUNSWICK	45.84	-66.45	50.6	2000	2018	18	94953
79	29714	LEMIEUX	QUEBEC	46.3	-72.06	97.2	1999	2018	19	136144
80	30164	SCHEFFERVILLE	QUEBEC	54.8	-66.8	517.2	2005	2018	13	64342
81	30165	MONTREAL/PIERRE ELLIOTT TRUDEAU INTL	QUEBEC	45.47	-73.74	32.1	2008	2018	10	66528
82	30168	NATASHQUAN A	QUEBEC	50.19	-61.81	11.7	2008	2018	10	68794
83	30308	HARRINGTON CDA CS	PRINCE EDWARD ISLAND	46.34	-63.17	53	2004	2018	14	119609
84	31829	LUNENBURG	NOVA SCOTIA	44.36	-64.3	3.8	2002	2018	16	131817
85	41575	TRACADIE	NOVA SCOTIA	45.61	-61.68	66.67	2003	2018	15	132171
86	41903	ST. PETERS	PRINCE EDWARD ISLAND	46.45	-62.58	29.7	2003	2018	15	128039
87	42013	GRANBY	QUEBEC	45.37	-72.77	86	2003	2018	15	50637
88	42083	NAPPAN AUTO	NOVA SCOTIA	45.76	-64.24	19.8	2003	2018	15	124848
89	42243	DEBERT	NOVA SCOTIA	45.42	-63.47	37.5	2003	2018	15	122873
90	43124	HALIFAX KOOTENAY	NOVA SCOTIA	44.59	-63.55	52	2004	2018	14	101129
91	43323	EDMUNDSTON	NEW BRUNSWICK	47.42	-68.32	154.2	2004	2018	14	103677
92	43404	OSBORNE HEAD DND	NOVA SCOTIA	44.61	-63.42	30	2004	2018	14	85804
93	43406	BEDFORD BASIN	NOVA SCOTIA	44.71	-63.63	3.5	2004	2018	14	97665
94	44363	UPPER STEWIACKE RCS	NOVA SCOTIA	45.23	-63.06	23.5	2005	2018	13	103267
95	44503	SYDNEY CS	NOVA SCOTIA	46.16	-60.04	62.5	2006	2018	12	106826
96	45047	BONAVISTA	NEWFOUNDLAND	48.67	-53.11	25.6	2006	2018	12	98733
97	45307	WRECKHOUSE	NEWFOUNDLAND	47.71	-59.31	31.7	2006	2018	12	101878
98	46007	BACCARO PT	NOVA SCOTIA	43.45	-65.47	4.6	2007	2018	11	95696
99	47587	SAINT-GERMAIN-DE-GRANTHAM	QUEBEC	45.83	-72.54	85	2008	2018	10	80411
100	47607	STEPHENVILLE RCS	NEWFOUNDLAND	48.56	-58.57	58	2008	2018	10	85459
101	48348	ST. ANTHONY	NEWFOUNDLAND	51.39	-56.07	32.9	2009	2018	9	74927
102	48568	FREDERICTON	NEW BRUNSWICK	45.87	-66.54	20.7	2010	2018	8	67566
103	48871	ST JOHNS WEST CLIMATE	NEWFOUNDLAND	47.51	-52.78	110	2010	2018	8	67305
104	49390	VAL D'OR A	QUEBEC	48.05	-77.78	337.4	2011	2018	7	53798
105	49491	MATAGAMI A	QUEBEC	49.76	-77.8	279.8	2011	2018	7	52561
106	49608	MONTREAL MIRABEL INTL A	QUEBEC	45.68	-74.04	82.3	2012	2018	6	49372
107	49629	CHURCHILL FALLS A	NEWFOUNDLAND	53.56	-64.11	439.5	2011	2018	7	52052

108	49648	LA GRANDE 4 A	QUEBEC	53.75	-73.68	306.3	2011	2018	7	48417
109	49649	SCHEFFERVILLE A	QUEBEC	54.81	-66.81	520.9	2012	2018	6	50055
110	49748	ESKASONI FIRST NATION AUTOMATIC WEATHER STATION	NOVA SCOTIA	45.92	-60.65	28	2011	2018	7	55871
111	50089	ST. JOHN'S INTL A	NEWFOUNDLAND	47.62	-52.75	140.5	2012	2018	6	54126
112	50309	MONCTON INTL A	NEW BRUNSWICK	46.11	-64.68	70.7	2012	2018	6	52220
113	50310	SAINT JOHN A	NEW BRUNSWICK	45.32	-65.89	108.8	2012	2018	6	51561
114	50620	HALIFAX INTL A	NOVA SCOTIA	44.88	-63.51	145.4	2012	2018	6	49913
115	50621	CHARLOTTETOWN A	PRINCE EDWARD ISLAND	46.29	-63.12	48.5	2012	2018	6	49891
116	50677	CORMACK RCS	NEWFOUNDLAND	49.32	-57.39	165.8	2012	2018	6	44830
117	50678	MILLERTOWN RCS	NEWFOUNDLAND	48.82	-56.54	203.6	2013	2018	5	41723
118	50719	OTTAWA GATINEAU A	QUEBEC	45.52	-75.56	64.3	2012	2018	6	30377
119	51418	BATHURST A	NEW BRUNSWICK	47.63	-65.74	58.8	2013	2018	5	40616
120	51457	QUEBEC INTL A	QUEBEC	46.79	-71.39	74.4	2013	2018	5	40628
121	51537	DOAKTOWN AUTO RCS	NEW BRUNSWICK	46.59	-66.01	43	2013	2018	5	40682
122	51638	LAC MEGANTIC	QUEBEC	45.6	-70.88	460.3	2013	2018	5	40458
123	52038	PUVIRNITUQ A	QUEBEC	60.05	-77.29	25.3	2014	2018	4	36621
124	52081	WASKAGANISH A	QUEBEC	51.47	-78.76	24.08	2014	2018	4	10602
125	52179	KUUJJUAQ A	QUEBEC	58.09	-68.42	39.9	2014	2018	4	37136
126	52759	STEPHENVILLE A	NEWFOUNDLAND	48.54	-58.55	24.7	2014	2018	4	32042

Appendix II: Tests statistics for all stations

Table 12: Slope of mean excess plot, Shapiro-Wilk (SW) log-normality test, Modified Mann-Kendall (MMK) test, Jackson test statistics and p-values for all stations

CS station ID	Station ID	SW [7] Log-normality test statistic	SW [7] Log-normality test p-value	MMK [6] statistic	MK [4] p-value	MMK [6] p-value	Jackson\`s test [38] statistic	Jackson\`s test [38] p-value	Slope of mean excess plot	Tail behhavior
1	5251	0.93481231	2.28E-34	-0.2021858	0.02166128	0.12758335	1.56332114	1	2.68440656	Class D
2	5293	0.95892703	2.48E-28	-0.0829268	0.45172647	0.2764187	1.52610075	1	-6.4991479	Class E
3	5415	0.95711197	7.16E-29	-0.1289617	0.14363575	0.20440144	1.56498987	1	-0.8616302	Class E
4	5530	0.9639304	9.56E-27	-0.1818182	0.0837127	0.06806273	1.61709338	1	1.35168316	Class D
5	5668	0.97638483	6.59E-22	-0.15	0.44404364	0.22268739	1.60151372	1	6.61710348	Class D
6	5669	0.96658447	7.79E-26	-0.0909091	0.45855875	0.28144999	1.63432095	1	8.33372025	Class D
7	5707	0.97584232	3.74E-22	-0.1272727	0.64042879	0.30672146	1.48347736	1	0.29553217	Class D
8	5839	0.96953715	9.38E-25	0.18333333	0.34441752	0.26813169	1.48320724	1	7.23863026	Class D
9	6029	0.93444108	1.91E-34	-0.0080645	0.96119873	0.95898441	1.50195529	1	10.7414845	Class D
10	6047	0.91188302	1.23E-38	-0.2192192	0.05790218	0.07938472	1.49275953	1	-6.7049647	Class E
11	6053	0.96111328	1.17E-27	-0.1461988	0.40110257	0.19297135	1.52632486	1	15.2726748	Class D
12	6058	0.96599302	4.83E-26	0.10606061	0.39410715	0.05112528	1.54777502	1	-7.8198381	Class E
13	6081	0.94150267	6.74E-33	0.07320024	0.42084374	0.14686764	1.52128921	1	-15.157217	Class E
14	6089	0.94865349	3.58E-31	0.04093567	0.83373524	0.64818442	1.5723862	1	5.85079177	Class D
15	6205	0.96736848	1.48E-25	0.02651515	0.84036442	0.74330691	1.55056855	1	0.80017988	Class D
16	6256	0.97336741	3.14E-23	-0.1895425	0.2888828	0.12257836	1.54894854	1	2.58657075	Class D
17	6330	0.95072477	1.23E-30	-0.0181818	1	1	1.54234048	1	-2.1288076	Class E
18	6335	0.94607421	8.15E-32	0.23809524	0.54800557	0.19931025	1.62006894	1	-1.0474461	Class E
19	6501	0.97841636	6.03E-21	-0.2513369	0.03794744	0.08424816	1.59586205	1	4.94858319	Class D
20	6557	0.98029839	5.39E-20	-0.3333333	0.45237036	0.21211346	1.45321575	1	6.63357691	Class D
21	6581	0.95384202	8.45E-30	-0.2	0.70711423	0.40509468	1.56308889	1	-9.9139578	Class E
22	6584	0.9736668	4.20E-23	-0.2	0.43627493	0.06472304	1.66649257	1	-8.5922124	Class E
23	6708	0.92647291	4.85E-36	0.22222222	0.46551217	0.16110606	1.57339242	1	-4.8109461	Class E
24	6709	0.97112435	3.85E-24	-0.1428571	0.76389059	0.41234628	1.6898481	1	1.74757555	Class D
25	6710	0.96755682	1.73E-25	0.06666667	1	1	1.43417475	1	3.41664655	Class D
26	6711	0.96167589	1.76E-27	-0.047619	0.84308523	0.81161818	1.43286562	1	8.93068579	Class D

27	6723	0.97842407	6.08E-21	-0.1809524	0.26387201	0.17409393	1.52634236	1	0.41597246	Class D
28	6740	0.95685804	6.04E-29	0.18530612	0.05869675	0.19090927	1.51811884	1	3.78472722	Class D
29	6756	0.9709129	3.18E-24	0	1	1	1.51564016	1	8.27069864	Class D
30	6773	0.9602558	6.31E-28	-0.0579592	0.55818465	0.44638122	1.55907095	1	-1.8456878	Class E
31	6778	0.97805488	4.02E-21	0.20879121	0.32442369	0.10657076	1.5150236	1	-3.5942856	Class E
32	6783	0.95959034	3.94E-28	-0.1538462	0.50215825	0.27939239	1.52503252	1	4.94084834	Class D
33	6797	0.97202713	8.83E-24	-0.0735294	0.71083593	0.6069561	1.64122849	1	-0.5934565	Class E
34	6916	0.96473836	1.79E-26	-0.1025641	0.66933401	0.38640725	1.52089551	1	10.2497783	Class D
35	6918	0.96337438	6.26E-27	-0.1025641	0.66933401	0.38907499	1.53675444	1	20.9562541	Class D
36	6923	0.95603741	3.50E-29	0.04575163	0.82021675	0.8066977	1.62969963	1	14.8891822	Class D
37	7026	0.94834736	2.99E-31	-0.1428571	0.51122035	0.06206496	1.60121212	1	-8.5639959	Class E
38	7162	0.94109217	5.43E-33	0.16666667	0.60216753	0.17375311	1.66850983	1	11.1839127	Class D
39	7169	0.94480077	4.00E-32	-0.2794118	0.12747693	0.18862426	1.58583059	1	-3.1506054	Class E
40	7173	0.93831946	1.30E-33	0.12380952	0.55261513	0.13194898	1.50947487	1	1.54263405	Class D
41	8456	0.95502865	1.81E-29	-0.0631579	0.72117631	0.72336297	1.56864683	1	12.2161765	Class D
42	8675	0.9377185	9.59E-34	0.20588235	0.26605176	0.05655811	1.52756212	1	11.5777406	Class D
43	8989	0.96600834	4.89E-26	0.4	0.46243273	0.2660457	1.58244215	1	11.5506313	Class D
44	8991	0.95604772	3.52E-29	0.16363636	0.53341651	0.2554553	1.5240639	1	4.90611012	Class D
45	8999	0.96159223	1.66E-27	0.14619883	0.40110257	0.39801429	1.58267489	1	-10.461789	Class E
46	9002	0.89741883	6.91E-41	0.09090909	0.57278231	0.46190887	1.34676725	1	0.40353873	Class D
47	10226	0.93861073	1.51E-33	-0.2066667	0.15425751	0.13930185	1.56984582	1	10.4143809	Class D
48	10252	0.97070993	2.65E-24	-0.0289855	0.86215561	0.74540156	1.5745926	1	2.43050786	Class D
49	10661	0.9366207	5.55E-34	-0.2	0.80649594	0.71409977	1.72997772	1	-2.5015868	Class E
50	10684	0.91253666	1.58E-38	0.16666667	0.26433733	0.13824615	1.43996805	1	-6.2704313	Class E
51	10763	0.94440289	3.22E-32	0.05882353	0.7618734	0.52398724	1.489577	1	11.0177912	Class D
52	10789	0.95888706	2.41E-28	0.02898551	0.86215561	0.72994085	1.44821935	1	11.5631234	Class D
53	10792	0.97838935	5.85E-21	-0.057971	0.70984368	0.51589066	1.62919164	1	12.4044824	Class D
54	10797	0.96144078	1.48E-27	0.12727273	0.64042879	0.22179148	1.60736136	1	0.43519723	Class D
55	10800	0.95481011	1.57E-29	0.01298701	0.95502646	0.89164261	1.52453336	1	-0.3561772	Class E
56	10813	0.94242542	1.10E-32	-0.030303	0.94532987	0.88324926	1.73424363	1	6.90223172	Class D
57	10817	0.94627557	9.12E-32	0.03030303	0.8656491	0.69245629	1.50555291	1	9.41346406	Class D

58	10818	0.93784318	1.02E-33	-0.1052632	0.53760324	0.20811161	1.50378922	1	2.40317922	Class D	
59	10836	0.97289302	2.00E-23	0.01818182	1	1	1.6635247	1	10.688216	Class D	
60	10841	0.97383519	4.95E-23	-0.3186813	0.12531111	0.05495594	1.66216111	1	1.16627959	Class D	
61	10855	0.9284077	1.15E-35	-0.1304348	0.38530998	0.27006424	1.47728004	1	-6.3970974	Class E	
62	10856	0.954859	1.62E-29	0.08225108	0.61176003	0.13112966	1.49265256	1	2.07979197	Class D	
63	10859	0.95917041	2.93E-28	-0.0217391	0.90129784	0.8386627	1.57591341	1	11.141533	Class D	
64	10869	0.96832087	3.29E-25	-0.2526316	0.12728869	0.10749857	1.69799329	1	-3.9318051	Class E	
65	10873	0.95704567	6.85E-29	0.21428571	0.53618677	0.2126473	1.54223966	1	-12.28814	Class E	
66	10909	0.97110481	3.78E-24	0.04761905	1	1	1.58636234	1	0.84644596	Class D	
67	10945	0.95331019	6.03E-30	0.05263158	0.77956607	0.65972822	1.59589914	1	-1.1814594	Class E	
68	26778	0.96347796	6.77E-27	-0.2761905	0.16585666	0.06412279	1.59602339	1	-2.9646804	Class E	
69	26797	0.89400236	2.21E-41	-0.2857143	0.38647623	0.10706378	1.54789949	1	9.32743581	Class D	
70	26893	0.96193969	2.14E-27	-0.047619	1	1	1.66582712	1	-10.380662	Class E	
71	26968	0.8974126	6.89E-41	0.06666667	1	1	1.65950249	1	6.13878741	Class D	
72	27115	0.96106509	1.13E-27	-0.1794872	0.42771051	0.09447317	1.55867945	1	17.0724157	Class D	
73	27195	0.94275486	1.31E-32	0.01052632	0.97411775	0.97099606	1.65428844	1	1.02836412	Class D	
74	27621	0.95464	1.41E-29	-0.3571429	0.26551039	0.07362676	1.61023359	1	7.10547151	Class D	
75	27712	0.95806841	1.37E-28	-0.0666667	0.7665253	0.53407036	1.55972186	1	10.6881494	Class D	
76	27846	0.95209676	2.83E-30	0	1	1	1.4986696	1	6.06753571	Class D	
77	29471	0.9794812	2.04E-20	-0.0666667	1	1	1.71007256	1	8.86518404	Class D	
78	29493	0.97570922	3.25E-22	0.09090909	0.75549688	0.52960709	1.52563149	1	14.4016511	Class D	
79	29714	0.95068113	1.19E-30	0.05263158	0.77956607	0.50888243	1.57181801	1	11.1652229	Class D	
80	30164	0.97266038	1.60E-23	0.02564103	0.95135213	0.93011753	1.52752719	1	0.75446435	Class D	
81	30165	0.95574029	2.88E-29	0.15555556	0.59150504	0.11270953	1.50735904	1	19.0980643	Class D	
82	30168	0.96913896	6.64E-25	-0.1111111	0.72051479	0.49646544	1.55164422	1	14.8035902	Class D	
83	30308	0.95142679	1.88E-30	-0.010989	1	1	1.51485875	1	18.0578582	Class D	
84	31829	0.98062706	8.02E-20	0.12380952	0.55261513	0.25693653	1.59392287	1	-4.8610456	Class E	
85	41575	0.94137046	6.29E-33	-0.0666667	0.7665253	0.301116	1.50743351	1	17.2246441	Class D	
86	41903	0.97759499	2.42E-21	-0.0095238	1	1	1.51289751	1	-8.9899278	Class E	
87	42013	0.94118096	5.69E-33	-0.3333333	0.36752074	0.05295609	1.67075836	1	19.7744645	Class D	
88	42083	0.94991966	7.56E-31	-0.1208791	0.58407037	0.07955597	1.59434405	1	17.7210905	Class D	

89	42243	0.94230537	1.03E-32	-0.1208791	0.58407037	0.24056589	1.63991705	1	13.181518	Class D
90	43124	0.97768012	2.66E-21	0.03030303	0.94532987	0.90212139	1.54922473	1	-1.2272778	Class E
91	43323	0.93673974	5.88E-34	0	1	1	1.64618544	1	-15.868101	Class E
92	43404	0.9658352	4.26E-26	-0.1111111	0.72051479	0.57063966	1.55846282	1	10.2306174	Class D
93	43406	0.96171733	1.82E-27	-0.0545455	0.87626966	0.67052487	1.65899371	1	18.7592327	Class D
94	44363	0.94327203	1.74E-32	-0.2121212	0.37269149	0.06233324	1.69247158	1	-3.5261599	Class E
95	44503	0.95386022	8.55E-30	-0.0606061	0.83701148	0.68830643	1.52327199	1	12.9337324	Class D
96	45047	0.9636125	7.50E-27	0.05454546	0.87626966	0.74296076	1.49550741	1	-2.5480518	Class E
97	45307	0.95529246	2.15E-29	0.15151515	0.53713386	0.14139297	1.68700806	1	11.8812594	Class D
98	46007	0.97885514	9.92E-21	0.16363636	0.53341651	0.19059469	1.57092374	1	11.9156812	Class D
99	47587	0.95270444	4.13E-30	-0.0181818	1	1	1.59203912	1	-2.1523862	Class E
100	47607	0.95827514	1.58E-28	-0.0222222	1	1	1.60771755	1	5.40230921	Class D
101	48348	0.9588749	2.39E-28	0.11111111	0.75445418	0.43293719	1.51750012	1	10.0336228	Class D
102	48568	0.96175385	1.87E-27	0.21428571	0.53618677	0.05421063	1.54519674	1	-4.3571003	Class E
103	48871	0.96641287	6.78E-26	-0.3571429	0.26551039	0.08046745	1.51063435	1	10.9554423	Class D
104	49390	0.96717869	1.27E-25	0.07142857	0.90153863	0.85286992	1.46614106	1	-3.576914	Class E
105	49491	0.97142313	5.06E-24	0.14285714	0.71052302	0.37979686	1.47820763	1	17.8292863	Class D
106	49608	0.97964629	2.48E-20	-0.3333333	0.36752074	0.10222829	1.56477759	1	14.1662494	Class D
107	49629	0.96710891	1.20E-25	0.2	0.70711423	0.24216476	1.51836126	1	11.1672923	Class D
108	49648	0.96127137	1.31E-27	-0.0714286	0.90153863	0.72134696	1.49585029	1	-4.5533548	Class E
109	49649	0.972995	2.20E-23	0.33333333	0.36752074	0.0801209	1.51245995	1	-5.559103	Class E
110	49748	0.97478645	1.27E-22	-0.2	0.70711423	0.40641853	1.6603877	1	7.70248528	Class D
111	50089	0.9649493	2.11E-26	-0.2	0.70711423	0.38202473	1.48559667	1	5.29683996	Class D
112	50309	0.98712011	7.08E-16	-0.0666667	1	1	1.50755097	1	6.98969019	Class D
113	50310	0.97000601	1.41E-24	-0.0666667	1	1	1.58608973	1	-6.9961901	Class E
114	50620	0.98561749	6.77E-17	-0.0666667	1	1	1.48851245	1	9.56242402	Class D
115	50621	0.98000869	3.81E-20	0.06666667	1	1	1.51236611	1	1.38414564	Class D
116	50677	0.96281358	4.10E-27	-0.4	0.46243273	0.07480954	1.63645409	1	21.4041495	Class D
117	50678	0.95280991	4.41E-30	-0.2	0.80649594	0.62319154	1.63456231	1	-24.350595	Class E
118	50719	0.98259763	9.71E-19	0.04761905	1	1	1.55997813	1	-20.143569	Class E
119	51418	0.96771933	1.99E-25	0.2	0.80649594	0.57419188	1.6488314	1	6.21708212	Class D

120	51457	0.94780528	2.19E-31	-0.0666667	1	1	1.58867705	1	-1.1715843	Class E
121	51537	0.94138976	6.35E-33	0	1	1	1.69568491	1	26.9418209	Class D
122	51638	0.96686503	9.80E-26	-0.3333333	0.45237036	0.27651246	1.59400317	1	27.8832854	Class D
123	52038	0.98072059	8.99E-20	-0.2	0.80649594	0.58267765	1.4854465	1	15.0241735	Class D
124	52081	0.94354083	2.01E-32	0.4	0.46243273	0.05889361	1.41541359	1	12.4844673	Class D
125	52179	0.93580971	3.72E-34	0.2	0.80649594	0.59460175	1.53313478	1	13.9734908	Class D
126	52759	0.96215025	2.50E-27	0	1	1	1.6463994	1	9.84037738	Class D

Appendix III: Estimated parameters of Halphen distributions for all stations

Table 13: Estimated parameters for HDF for all stations

CS station ID	Station ID	Station name	Province	HA, α, *m, v*	HB3, α, *m, v*
1	5251	QUEBEC/JEAN LESAGE INTL A	QUEBEC	4.0960920e-02 6.3681010e-02 2.3506790e-01	4.7960350e-02 7.2552360e-02 2.9643370e-01
2	5293	VALCARTIER A	QUEBEC	1.4083120e+03 2.5896970e+05 1.6308770e+10	-
3	5415	MONTREAL/PIERRE ELLIOTT TRUDEAU INTL A	QUEBEC	1.9568490e-02 1.8002560e-02 7.5457070e-02	4.4917490e-02 6.5283050e-02 2.3025260e-01
4	5530	SHERBROOKE A	QUEBEC	2.0408160e-01 5.7322030e-01 5.7978410e+00	-
5	5668	BLANC SABLON	QUEBEC	1.9942760e-01 3.8956660e-01 1.9516860e+00	6.0390000e-02 7.9205730e-02 1.0590340e-01
6	5669	LOURDES DE BLANC SABLON A	QUEBEC	1.5193530e-01 2.6545630e-01 1.0074200e+00	-
7	5707	MANICOUAGAN A	QUEBEC	9.0528080e-02 1.0671120e-01 2.0347530e-01	-
8	5839	RIVIERE DU LOUP	QUEBEC	-	2.6114340e-02 3.4029430e-02 3.6034810e-02
9	6029	CHIB3OUGAMAU CHAPAIS A	QUEBEC	-	3.6233620e-02 4.8820970e-02 1.4597630e-01
10	6047	LA GRANDE RIVIERE A	QUEBEC	4.1366030e-02 6.0564330e-02 3.1796050e-01	4.2471850e-02 5.6963900e-02 1.5588030e-01
11	6053	MATAGAMI A	QUEBEC	-	6.5000840e-02 9.8906660e-02 4.5366790e-01
12	6058	NITCHEQUON	QUEBEC	4.7310970e-02 6.5793580e-02 2.1147270e-01	6.1495790e-02 8.7638250e-02 2.9964400e-01
13	6081	VAL-D'OR A	QUEBEC	1.6289890e-02 1.6425830e-02 1.1574300e-01	2.9143840e-02 3.9855350e-02 1.2232030e-01
14	6089	CAPE HOPES ADVANCE	QUEBEC	-	2.8961960e-01 4.6447660e-01 2.3146830e+00
15	6205	MISCOU ISLAND (AUT)	NEW BRUNSWICK	-	1.0506120e-02 1.4077980e-02 1.1850240e-02
16	6256	ST LEONARD A	NEW BRUNSWICK	5.0397450e-02 7.2231860e-02 2.4157440e-01	6.0780910e-02 8.7571420e-02 3.1082730e-01
17	6330	COPPER LAKE	NOVA SCOTIA	9.4928350e-02 1.6058780e-01 7.7389680e-01	9.0600560e-02 1.2593830e-01 3.8563690e-01
18	6335	DEBERT A	NOVA SCOTIA	6.6416980e-02 7.3866250e-02 1.6801170e-01	4.3468270e-02 6.0002750e-02 1.0505590e-01

19	6501	WESTERN HEAD	NOVA SCOTIA	6.0899960e-02 1.0333150e-01 4.6217380e-01	-
20	6557	ARGENTIA A	NEWFOUNDLAND	-	5.1652060e-02 4.9957610e-02 1.6348540e-02
21	6581	BUCHANS A	NEWFOUNDLAND	9.8128220e-02 1.6974770e-01 1.0225960e+00	2.0359160e-01 3.4979720e-01 2.1773120e+00
22	6584	BURGEO 2	NEWFOUNDLAND	1.9485020e-01 3.4525260e-01 1.3358240e+00	2.1237790e-02 2.2771710e-02 2.6282380e-03
23	6708	ST ALBANS	NEWFOUNDLAND	7.3097140e-02 1.0174130e-01 2.9703730e-01	1.0764730e-01 1.6788190e-01 7.4284420e-01
24	6709	ST ANDREWS	NEWFOUNDLAND	2.0733220e-01 3.1900530e-01 9.2826400e-01	-
25	6710	ST ANTHONY	NEWFOUNDLAND	-	1.0617630e-01 1.3109820e-01 2.6263750e-01
26	6711	ST ANTHONY	NEWFOUNDLAND	-	6.0964460e-02 7.4327660e-02 1.5275310e-01
27	6723	ST LAWRENCE	NEWFOUNDLAND	5.2101160e-02 6.1000000e-02 4.0934150e-01	9.1132390e-02 1.2524660e-01 4.0448120e-01
28	6740	STEPHENVILLE A	NEWFOUNDLAND	5.0756620e-02 8.5084640e-02 4.5651790e-01	3.7315100e-02 4.7019120e-02 1.1464010e-01
29	6756	TWILLINGATE	NEWFOUNDLAND	8.9163510e-02 8.6294390e-02 3.3388190e-01	1.6516440e-01 2.1440220e-01 5.1498540e-01
30	6773	CARTWRIGHT	NEWFOUNDLAND	4.2631060e-02 6.0121290e-02 1.9755930e-01	6.1624990e-02 8.6846570e-02 2.7613620e-01
31	6778	MAKKOVIK A	NEWFOUNDLAND	7.0028460e-02 8.3931190e-02 1.8291390e-01	9.8111770e-02 1.2920470e-01 3.3670190e-01
32	6783	MARY'S HARBOUR A	NEWFOUNDLAND	-	8.0414140e-02 1.1510530e-01 4.3412210e-01
33	6797	SAGLEK	NEWFOUNDLAND	7.5023690e-02 7.8497970e-02 9.3288110e-02	5.6007030e-02 7.6859330e-02 1.0948200e-01
34	6916	BATHURST A	NEW BRUNSWICK	5.3963780e+01 2.3839240e+03 1.2748220e+07	-
35	6918	BUCTOUCHE CDA CS	NEW BRUNSWICK	-	6.3235090e-02 8.7303700e-02 2.6990200e-01
36	6923	KEJIMKUJIK 1	NOVA SCOTIA	7.4863250e-01 3.7263480e+00 1.1017300e+02	-
37	7026	ST STEPHEN (AUT)	NEW BRUNSWICK	5.8236170e-02 8.2554880e-02 2.2745240e-01	1.3563500e-01 3.0551030e-01 3.3881700e+00
38	7162	TRURO (AUT)	NOVA SCOTIA	7.2455650e-02 8.6015680e-02 1.4225030e-01	-

39	7169	MCNABS ISLAND (AUT)	NOVA SCOTIA	-	3.1817540e-02 4.1257120e-02 5.6448160e-02
40	7173	HART ISLAND (AUT)	NOVA SCOTIA	-	1.1394260e-01 2.4255710e-01 3.2220440e+00
41	8456	CHEVERY	QUEBEC	4.7561900e-02 9.1084110e-02 6.3204570e-01	9.2562520e-02 1.3295450e-01 4.4309490e-01
42	8675	HEATH POINT	QUEBEC	-	1.0280760e-01 1.2829140e-01 3.0476460e-01
43	8989	ILE D'ORLEANS	QUEBEC	2.4087090e-01 4.8699960e-01 2.8558840e+00	4.2226710e-01 7.7992660e-01 5.5383090e+00
44	8991	SAGONA ISLAND	NEWFOUNDLAND	-	1.6937920e-01 3.2045400e-01 3.1878980e+00
45	8999	CAP-MADELEINE	QUEBEC	3.4217110e-01 9.3983060e-01 9.2857910e+00	3.2439990e-02 4.1981350e-02 5.6105120e-02
46	9002	EASTMAIN A	QUEBEC	-	4.0105760e-02 4.7795100e-02 9.9738660e-02
47	10226	SALLUIT A	QUEBEC	-	2.5343350e-01 5.6593630e-01 6.8099010e+00
48	10252	KANGIQSUJUAQ A	QUEBEC	1.3464490e-01 1.8665440e-01 5.4410810e-01	2.3512000e-01 3.8990390e-01 2.1041580e+00
49	10661	NORTHEAST MARGAREE (AUT)	NOVA SCOTIA	4.5098350e-02 4.7617820e-02 7.9448980e-02	-
50	10684	KANGIQSUALUJJUAQ A	QUEBEC	-	7.1638320e-02 9.0782320e-02 2.1745460e-01
51	10763	ILES DE LA MADELEINE	QUEBEC	-	.4195460e-02 8.5880040e-02 2.6946900e-01
52	10789	AKULIVIK A	QUEBEC	-	8.1641210e-02 1.0324310e-01 2.5721770e-01
53	10792	GRAND ETANG	NOVA SCOTIA	4.5953530e-01 1.3570260e+00 1.5559980e+01	.9787620e-02 5.3995740e-02 7.4100400e-02
54	10797	MISTOOK	QUEBEC	1.4625750e-01 3.0110140e-01 1.6000780e+00	3.6381880e-02 4.8389110e-02 6.6623540e-02
55	10800	SUMMERSIDE	PRINCE EDWARD ISLAND	-	6.2252310e-02 9.3792990e-02 4.2790690e-01
56	10813	NAIN	NEWFOUNDLAND	4.2808770e-02 6.5045280e-02 1.3302980e-01	2.6383230e-02 3.8135840e-02 9.6604520e-02
57	10817	POOLS ISLAND	NEWFOUNDLAND	-	.7350330e-02 8.9581830e-02 2.6298180e-01
58	10818	GRATES COVE	NEWFOUNDLAND	-	2.7670880e-02 3.3744840e-02 4.2178880e-02

59	10836	CAPE KAKKIVIAK	NEWFOUNDLAND	1.1199000e-01 1.2052040e-01 1.5219240e-01	2.9628150e-02 3.1755050e-02 5.6517800e-04
60	10841	CAPE KIGLAPAIT	NEWFOUNDLAND	1.0767390e-01 1.2662320e-01 2.2653820e-01	-
61	10855	AUPALUK A	QUEBEC	-	1.1519660e-01 1.7196820e-01 8.6338450e-01
62	10856	TASIUJAQ A	QUEBEC	-	1.1821040e-01 1.5896810e-01 5.0575100e-01
63	10859	BRIER ISLAND	NOVA SCOTIA	-	1.1492390e-01 1.7364740e-01 6.9299050e-01
64	10869	CAP-TOURMENTE	QUEBEC	1.3062060e-01 2.2009230e-01 7.5090910e-01	-
65	10873	STE-ANNE-DE-BELLEVUE 1	QUEBEC	-	2.5135820e-02 3.1585770e-02 3.9765420e-02
66	10909	MARTICOT ISLAND	NEWFOUNDLAND	6.8048080e-02 9.1046800e-02 5.4671480e-01	1.8937340e-01 2.9404290e-01 1.2399300e+00
67	10945	SHEARWATER JETTY	NOVA SCOTIA	2.1620050e-01 5.7837510e-01 5.4367130e+00	2.5488450e-02 3.3506560e-02 4.4302850e-02
68	26778	LAC BENOIT	QUEBEC	3.2953370e-02 4.3649650e-02 7.5927250e-02	2.0281760e-02 2.9433860e-02 4.1651020e-02
69	26797	CHAMOUCHOUANE	QUEBEC	1.8443680e+01 5.8005400e+02 2.8594490e+06	-
70	26893	LA POCATIERE	QUEBEC	2.6710360e-01 5.2497890e-01 2.4396690e+00	2.0652060e-02 2.2147830e-02 3.4377600e-03
71	26968	KOUCHIB3OUGUAC CS	NEW BRUNSWICK	5.0148400e-01 1.5510880e+00 1.7979680e+01	-
72	27115	ST JOHN'S WEST CDA CS	NEWFOUNDLAND	-	2.5922010e-02 3.2959530e-02 4.2014980e-02
73	27195	CAP-CHAT	QUEBEC	9.0591850e-02 1.4349490e-01 4.8872480e-01	1.5937370e-02 1.7009780e-02 2.8356910e-03
74	27621	ST LAWRENCE	NEWFOUNDLAND	-	6.2602310e-02 8.3104840e-02 1.2260700e-01
75	27712	WINTERLAND	NEWFOUNDLAND	1.5389370e-01 2.9904850e-01 1.8002790e+00	1.1960630e-01 1.6809050e-01 5.2405130e-01
76	27846	CONFEDERATION BRIDGE	PRINCE EDWARD ISLAND	-	2.0165460e-01 3.5875890e-01 3.2526660e+00
77	29471	WRECKHOUSE	NEWFOUNDLAND	2.0815790e-01 3.0329920e-01 8.2013740e-01	-
78	29493	GAGETOWN AWOS A	NEW BRUNSWICK	6.0863220e-02 1.1976980e-01 1.4269900e+00	-

79	29714	LEMIEUX	QUEBEC	-	1.7830050e-02 2.3033880e-02 3.1168190e-02
80	30164	SCHEFFERVILLE	QUEBEC	-	8.9883800e-02 1.2495890e-01 4.0907370e-01
81	30165	MONTREAL/PIERRE ELLIOTT TRUDEAU INTL	QUEBEC	-	6.9968600e-02 9.7299770e-02 3.3081130e-01
82	30168	NATASHQUAN A	QUEBEC	-	1.1930340e-01 1.7649670e-01 6.9723780e-01
83	30308	HARRINGTON CDA CS	PRINCE EDWARD ISLAND	-	5.9855970e-02 8.1505220e-02 2.5346880e-01
84	31829	LUNENBURG	NOVA SCOTIA	2.6765990e-01 7.1111800e-01 6.5767320e+00	3.2191100e-02 4.2152670e-02 5.5504440e-02
85	41575	TRACADIE	NOVA SCOTIA	-	7.3717410e-02 1.3817520e-01 1.2569610e+00
86	41903	ST. PETERS	PRINCE EDWARD ISLAND	-	6.0274290e-02 8.1297270e-02 2.3972800e-01
87	42013	GRANBY	QUEBEC	3.7682400e-02 4.8197470e-02 1.5346340e-01	1.0280830e-02 1.1044290e-02 2.1161120e-03
88	42083	NAPPAN AUTO	NOVA SCOTIA	2.3258880e-01 5.9324750e-01 4.9474520e+00	2.7582740e-02 3.6151370e-02 4.8714100e-02
89	42243	DEBERT	NOVA SCOTIA	2.7141240e-02 3.3959900e-02 5.3349220e-02	1.7931920e-02 2.7213110e-02 4.0480980e-02
90	43124	HALIFAX KOOTENAY	NOVA SCOTIA	-	-
91	43323	EDMUNDSTON	NEW BRUNSWICK	3.0331570e-02 2.9419280e-02 4.1579050e-02	7.4831210e-03 7.9743930e-03 1.6850940e-03
92	43404	OSBORNE HEAD DND	NOVA SCOTIA	-	1.3610130e-01 1.9229310e-01 6.2518580e-01
93	43406	BEDFORD BASIN	NOVA SCOTIA	3.7319470e-02 5.1363000e-02 1.0364510e-01	2.4961400e-02 3.8602090e-02 5.9095880e-02
94	44363	UPPER STEWIACKE RCS	NOVA SCOTIA	1.2466300e-01 2.1552030e-01 7.5816780e-01	9.3288150e-03 1.1555020e-02 1.2565860e-02
95	44503	SYDNEY CS	NOVA SCOTIA	-	7.7940080e-02 1.0883370e-01 3.7136230e-01
96	45047	BONAVISTA	NEWFOUNDLAND	-	1.0772400e-01 1.5754540e-01 7.3780870e-01
97	45307	WRECKHOUSE	NEWFOUNDLAND	3.1334760e-01 5.8419080e-01 2.4541570e+00	2.0496150e-02 2.5279170e-02 2.7322820e-02
98	46007	BACCARO PT	NOVA SCOTIA	6.0310040e-02 5.9165220e-02 2.7357570e-01	1.5001800e-01 2.1623730e-01 7.3076240e-01
99	47587	SAINT-GERMAIN-DE-GRANTHAM	QUEBEC	3.5627790e-01 1.0371180e+00 1.1801460e+01	2.9970360e-02 3.9169160e-02 5.2411690e-02

100	47607	STEPHENVILLE RCS	NEWFOUNDLAND	3.9949440e-02 4.1251370e-02 1.6661300e-01	4.2540630e-01 1.5659820e+00 4.7005250e+01
101	48348	ST. ANTHONY	NEWFOUNDLAND	9.6758680e-02 1.2495770e-01 3.6566980e-01	1.2402560e-01 1.6421780e-01 4.4036730e-01
102	48568	FREDERICTON	NEW BRUNSWICK	6.3667670e+02 7.7284290e+04 1.7483850e+09	-
103	48871	ST JOHNS WEST CLIMATE	NEWFOUNDLAND	-	8.8774960e-02 1.4874570e-01 9.6980650e-01
104	49390	VAL D'OR A	QUEBEC	5.9515210e-02 6.9012130e-02 4.2156090e-01	8.4149140e-02 1.1193710e-01 3.0121460e-01
105	49491	MATAGAMI A	QUEBEC	7.2468250e-02 8.3266230e-02 3.7040140e-01	9.3736960e-02 1.2503180e-01 3.3799660e-01
106	49608	MONTREAL MIRABEL INTL A	QUEBEC	-	9.8981210e-02 1.6023650e-01 8.4302300e-01
107	49629	CHURCHILL FALLS A	NEWFOUNDLAND	8.7181020e+02 1.2191880e+05 7.6333290e+09	-
108	49648	LA GRANDE 4 A	QUEBEC	1.0289170e-01 1.1800690e-01 1.3496150e-01	1.0755790e-01 1.4433860e-01 3.9548500e-01
109	49649	SCHEFFERVILLE A	QUEBEC	1.1729280e-01 2.2447840e-01 3.0861460e+00	-
110	49748	ESKASONI FIRST NATION AUTOMATIC WEATHER STATION	NOVA SCOTIA	5.8718810e-02 6.0554950e-02 1.7704330e-01	-
111	50089	ST. JOHN'S INTL A	NEWFOUNDLAND		1.1107110e-01 1.4442320e-01 3.9981690e-01
112	50309	MONCTON INTL A	NEW BRUNSWICK	-	1.2010750e-01 1.6040150e-01 4.3046650e-01
113	50310	SAINT JOHN A	NEW BRUNSWICK	-	2.8202850e-02 4.1131630e-02 6.6148290e-02
114	50620	HALIFAX INTL A	NOVA SCOTIA	-	1.0418130e-01 1.3600670e-01 3.5507620e-01
115	50621	CHARLOTTETOWN A	PRINCE EDWARD ISLAND	-	1.0185580e-01 1.3648550e-01 3.9451140e-01
116	50677	CORMACK RCS	NEWFOUNDLAND	5.6303600e-02 8.0829840e-02 3.3088270e-01	3.9351990e-02 5.4126660e-02 7.7038110e-02
117	50678	MILLERTOWN RCS	NEWFOUNDLAND	6.3281930e-02 1.1580540e-01 5.0185220e-01	2.5830250e-02 3.5466360e-02 5.1339110e-02
118	50719	OTTAWA GATINEAU A	QUEBEC	-	3.9753410e-02 5.1038120e-02 7.0524010e-02
119	51418	BATHURST A	NEW BRUNSWICK	3.0862500e-01 7.0188710e-01 4.4385730e+00	1.6565970e-02 1.7611680e-02 3.5378290e-03

120	51457	QUEBEC INTL A	QUEBEC	-	4.3770700e-02 5.7719070e-02 8.5419130e-02
121	51537	DOAKTOWN AUTO RCS	NEW BRUNSWICK	2.8056040e-01 5.7873330e-01 2.8825240e+00	1.3739000e-02 1.4979260e-02 0.0000000e+00
122	51638	LAC MEGANTIC	QUEBEC	1.4304330e-01 3.1780230e-01 1.9973630e+00	3.4685200e-02 4.5715020e-02 6.0319740e-02
123	52038	PUVIRNITUQ A	QUEBEC	1.2918360e-01 1.6595800e-01 4.5664300e-01	1.6029300e-01 2.0614570e-01 4.9477990e-01
124	52081	WASKAGANISH A	QUEBEC	-	1.0165530e-01 1.2832630e-01 3.4127470e-01
125	52179	KUUJJUAQ A	QUEBEC	-	1.1703200e-01 1.7912550e-01 8.3719680e-01
126	52759	STEPHENVILLE A	NEWFOUNDLAND	1.0359370e+00 2.9547970e+00 3.1000970e+01	8.4928170e-02 1.1746310e-01 1.7004910e-01