



## IMPACT OF WIND SPEED AVERAGING TIME ON THE TREND DETECTION IN A CHANGING CLIMATE

Li, Sihan<sup>1,3</sup> and Irwin, Peter<sup>2</sup>

<sup>1</sup> RWDI Consulting Engineers and Scientists (RWDI), Canada

<sup>2</sup> RWDI Consulting Engineers and Scientists (RWDI), Canada

<sup>3</sup> [sihan.li@rwdi.com](mailto:sihan.li@rwdi.com)

**Abstract:** A previous study by the authors indicates that historical wind speeds may exhibit a statistically significant trend attributable to a changing climate. While data quality control was scrutinized, uncertainties from several causes were identified. These include changes of observation averaging time, data standardization for terrain effects and other large scale cyclical effects such as El Niño. Each of these factors deserves investigation. As a following up study, this paper focuses on the first factor. The wind speed archived by Environmental and Climate Change Canada is reported as a “nominal” hourly mean wind speed at each hour. This nominal hourly mean wind speed is not an average over an entire hour. The most widely used dataset, HLY01, began by recording the one minute mean wind speed before the top of each hour, but changed to recording the two-minute mean wind speed before the top of each hour in 1985. This would indicate that the wind speed data prior to 1985 could be biased compared to those after 1985. This paper studies several stations in Ontario to investigate the impact of the change of averaging time on the detection of a trend in historical extreme wind speeds. The ratio of one-minute wind speed to two-minute average wind speed is treated as a random variable and Monte Carlo simulation is employed to randomly correct the one-minute wind speed to two-minute mean wind speed before extracting the annual maximum wind speed. Trend analysis is then applied to detect trends in the corrected extreme wind speeds. The impact of the change of averaging time on detected trends is discussed.

### 1 Introduction

A previous study (Li et al. 2017) shows that there could be statistically significant trends of extreme wind speeds for some major cities in Canada. Climate change could be one possible reason. However, other factors could introduce uncertainties in the trend detection. These factors include change of observation averaging time, bias in the process of data standardization and natural variations caused by larger scale cyclical effects inherent in the climate. The change of observation averaging time occurs for the wind speed dataset in HY01 that is widely used to determine the design wind speed. The HY01 dataset provides nominal hourly mean wind speeds represented by the last 1-minute mean wind speed in each hour before year 1985, after which the dataset started recording the last 2-minute mean wind speed. This inconsistency in averaging time for the nominal hourly mean wind speed introduces uncertainty in the trend detection. The second factor is the possible bias introduced by the data standardization process. There are two parameters involved in the standardization process, anemometer height and exposure. The height standardization adjusts the wind speed measurement at specific height to 10-m standard height by assuming either a power law or a log law for the mean wind velocity profile. The exposure standardization applies factors to convert wind speed for specific wind directions from site terrain condition to the standard open terrain condition. For height standardization, if the information reported for anemometer height at the meteorological station is accurate, the uncertainty due to height correction should be very small. Uncertainty

from exposure standardization could be greater mainly because changes historically in the surrounding terrain (e.g. expanding suburbs) can be difficult to fully account for. The last possible uncertainty is due to natural long period cyclical climate variations, e.g., El Niño\La Niña. In this study, the impact of possible uncertainties from the change of observation averaging time in trend detection for the extreme wind speed will be focussed on. The other two factors will be the subject of future research.

Both the last 1-minute average and last 2-minute average wind speeds before the top of each hour are historically used in the HY01 dataset to represent the nominal hourly mean wind speed. Intuitively, the last 1-minute mean wind speed ( $V_{1m}$ ) in an hour could be either greater or less than the last 2-minute mean wind speed ( $V_{2m}$ ) in the same hour. This is simply because the  $V_{2m}$  is the average of the last two 1-minute mean wind speed. The relation between the  $V_{1m}$  and the  $V_{2m}$  is a random variable due to the stochastic nature of the wind speed. This indicates that even if the annual maximum of the  $V_{1m}$  before the top of each hour is extracted, it may not necessarily guarantee that such annual maximum nominal hourly mean wind speed is greater than that extracted from the  $V_{2m}$  at the top of each hour in the same year. This randomness could have an impact on the extracted annual maximum wind speed and consequently on the trend detection. To understand this impact, the statistical relation between the  $V_{1m}$  and  $V_{2m}$  needs to be established. As there is an expected correlation between the  $V_{1m}$  and  $V_{2m}$  in the same hour, the relation between these two values may not be well represented by conventional gust factor formulae, which provide relations to predict the ratio of gust speed of a given duration to the mean speed averaged over one hour. This is because even if the peak value of one minute mean wind speed in the hour occurs at the last 1-minute, it does not guarantee that the peak value of the two-minute mean has its peak value at the last 2-minute in the same hour.

The objective of this study is to investigate the impact of the change of observation averaging time on trend detection in the extreme wind speeds. For this purpose, a probabilistic model is first derived from observations to model the relation between the  $V_{2m}$  and  $V_{1m}$ . Such a probabilistic model is applied to the dataset before year 1985 by randomly converting the observed  $V_{1m}$  before the top of each hour to the  $V_{2m}$  in same hour using the Monte Carlo technique. The statistically corrected annual maximum wind speeds prior to 1985 then are combined with the annual maximum 2-minute mean wind speeds after 1985 and used to conduct the trend analysis. The impact of the change of averaging time on the estimated return period wind speeds is also investigated.

## 2 Characteristics of the ratio of $V_{1m}$ to $V_{2m}$

Eight stations studied in Li et al. (2017) are used in this study. These stations are at the airports for Sudbury, Ottawa, Toronto, Waterloo, Hamilton, London, Windsor and Niagara Falls (U.S.). The geographical locations for these stations are shown in Figure 1. The mean hourly wind speeds for these stations prior to 1985 are 1-minute mean wind speeds prior to the top of the hour. As the 2-minute mean wind speed prior to the top of the hourly is used to represent the mean hourly wind speed after 1985, this study converts the  $V_{1m}$  to  $V_{2m}$  for records prior to 1985.

To convert  $V_{1m}$  to a  $V_{2m}$  in the same epoch, the relation between these two average speeds first needs to be built. A dataset with long term continuous 1-minute mean wind speed records is ideal for this purpose. However, to the authors' best knowledge, there is no such dataset available close to the studied stations. Therefore, this study employs the Automatic Surface Observation System (ASOS) one-minute dataset (<ftp://ftp.ncdc.noaa.gov/pub/data/asos-onemin>. Last visit on Oct. 30, 2017). The ASOS one-minute dataset provides a moving average 2-minute mean wind speed recorded at each one minute. Although this dataset does not directly provide the 1-minute mean wind speed, it provides continuous 2-minute moving average record that can be used to derive an approximate relation between the 2-minute mean wind speed and 1-minute mean wind speed as established below.

Assume the 1-minute mean wind speed was continuously recorded in an hour. There will then be 60 1-minute mean wind speeds for each hour. The  $i^{\text{th}}$  record in the hour can be denoted as  $V_{1m,i}$ . Similarly, the moving average 2-minute mean wind speed record in the hour in ASOS dataset can be denoted as  $V_{2m,i}$ . The last 2-minute mean wind speeds ( $V_{2m,60}$ ) can then be calculated from,

$$[1] V_{2m\_60} = (V_{1m\_59} + V_{1m\_60})/2$$

In this study, a relation is defined such that,

$$[2] V_{2m\_60} = V_{1m\_60} + \varepsilon_{60}$$

where  $\varepsilon_{60}$  is a random correction term.

Replacing  $V_{2m\_60}$  in Eq. (2) by  $V_{2m\_60}$  from Eq. (1), the correction term can be expressed by,

$$[3] \varepsilon_{60} = (V_{1m\_59} - V_{1m\_60})/2$$

Although there are no  $V_{1m}$  data available for stations in the studied areas to derive a probabilistic model for this correction term, it is noted that a relation exists in,

$$[4] \delta = V_{2m\_59} - V_{2m\_60} = (V_{1m\_58} - V_{1m\_59})/2 + (V_{1m\_59} - V_{1m\_60})/2$$

The mean values of  $\varepsilon$  and  $\delta$  over many hours have to tend to zero and from Equation [4] it can be deduced that the standard deviations are related by

$$[5] \sigma_{\delta} = \sigma_{\varepsilon} \sqrt{2(1+R)}$$

where  $R$  = correlation coefficient between  $\varepsilon_{59}$  and  $\varepsilon_{60} = \frac{\overline{\varepsilon_{59}\varepsilon_{60}}}{\sigma_{\varepsilon}^2}$ . If  $R = 0$  then Equation [5] implies

$\sigma_{\delta} = \sigma_{\varepsilon} \sqrt{2}$  but it seems probable that there will be some correlation. So in general  $R$  will not be zero. It seems reasonable to assume that the probability distributions of  $\delta$  and  $\varepsilon$  will be of the same form, Therefore, knowledge of the standard deviation of  $\delta$ , combined with Equation [5] can in principle be used to infer the probability distribution of  $\varepsilon$ . Knowledge of the correlation coefficient  $R$  is also needed and this will be discussed later.

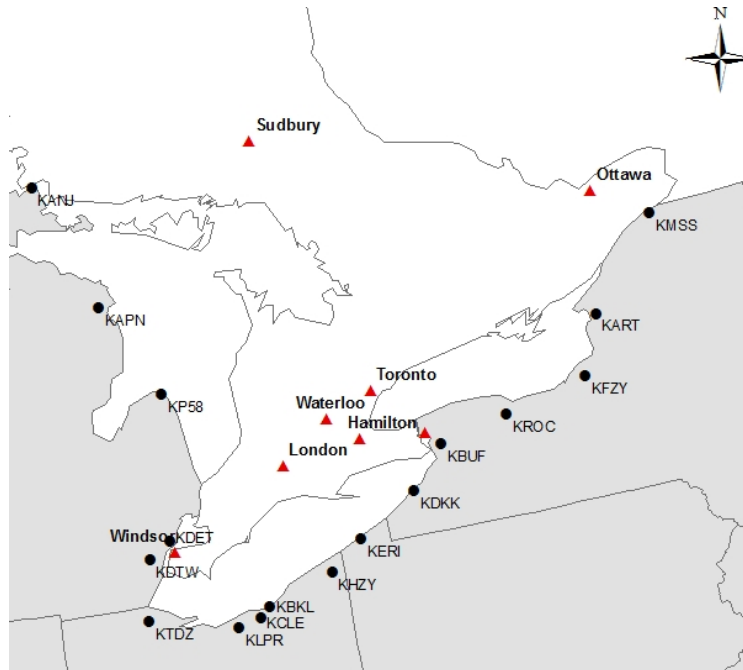


Figure 1: Geographical location of studied sites and ASOS 1-min dataset stations

To derive a probabilistic model for  $\delta$ , several stations from the ASOS 1-min dataset are selected. These ASOS stations are geographically close to the studied stations in HY01 but located in the US. There are 17 ASOS stations in this area close to the studied sites as shown in Figure 1. Two of these stations (KBUF and KROC) are excluded from this study due to data quality issue. Consequently, 15 ASOS 1-minute datasets are selected in this study to derive the probabilistic model for  $\delta$ . Before calculating  $\delta$ , a threshold wind speed is determined from the eight HY01 stations. The threshold value is equal to the minimum value of the annual maximum wind speeds for eight studied station less 1.5 times the corresponding standard deviation of the annual maximum wind speeds. The application of a threshold value is because the extreme wind speed is mostly of concern in this study. By filtering out the low wind speeds, the statistics derived from high wind speeds can better address the issue studied in this paper. The threshold wind speed is finally found to be 20 km/h.

There are about 10 to 15 years of data for each of these ASOS stations. All wind speed records in the ASOS stations have been converted into km/h before any analysis in this study. The skewness for all stations is close to zero. The kurtosis for these stations ranges from 3.4 to 5.1. These indicate that the Normal distribution, while not exact, could be used as an approximation to model  $\delta$ . It was found that the variability of  $\sigma_\delta$  amongst different stations was small. Therefore, the mean value of  $\sigma_\delta = 2.32$  km/hr was used in this study for all HY01 stations.

Although there is no continuous one-minute mean wind speed dataset available in Ontario, 14 years of continuous 1-minute mean wind speed record is obtained from a meteorological station located at Bratt's Lake, Saskatchewan. As there is no comparison that can be made to validate that the statistics calculated from this data could be representative for the wind climate of Ontario, this dataset is only used to examine the assumptions made in Eq. (3) through Eq. (5). Similarly, a threshold wind speed of 20km/h was also applied to this dataset. Values of  $\epsilon$  and  $\delta$  were calculated from this continuous one-minute mean wind speed record. The skewness values for  $\epsilon$  and  $\delta$  were found to be -0.21 and -0.15, respectively. The kurtosis values for  $\epsilon$  and  $\delta$  were 5.74 and 5.83, respectively. These observations indicate that Normal distribution is, while not an exact fit, is a reasonable approximation for modeling  $\epsilon$  and  $\delta$ . The standard deviations for  $\epsilon$  and  $\delta$  were 2.70 km/hr and 3.20 km/hr, respectively. The ratio of latter to former is 1.19, which is less than  $\sqrt{2}$  and implies a correlation coefficient for  $\epsilon$  of -0.29. The -0.29 correlation coefficient was therefore assumed in relating  $\sigma_\epsilon$  to  $\sigma_\delta$ . Considering this adjustment, the final standard deviation used to model  $\epsilon$  in Eq. (2) is equal to  $2.32/1.19 = 1.95$  km/hr.

Table 1. Statistics of  $\delta$  for 15 ASOS dataset

| Station | Mean  | $\sigma_\delta$ | Skewness | Kurtosis | Station | Mean  | $\sigma_\delta$ | Skewness | Kurtosis |
|---------|-------|-----------------|----------|----------|---------|-------|-----------------|----------|----------|
| KANJ    | -0.34 | 2.33            | -0.03    | 3.47     | KERI    | -0.26 | 2.53            | -0.05    | 3.73     |
| KAPN    | -0.35 | 2.34            | -0.05    | 3.89     | KFZY    | -0.29 | 2.3             | 0.01     | 3.56     |
| KART    | -0.14 | 2.05            | -0.02    | 4.33     | KHZY    | -0.31 | 2.25            | -0.14    | 4.38     |
| KBKL    | -0.33 | 2.44            | 0.06     | 3.44     | KLPR    | -0.16 | 2.17            | -0.14    | 5.11     |
| KCLE    | -0.19 | 2.24            | -0.07    | 3.99     | KMSS    | -0.18 | 2.25            | -0.09    | 4.06     |
| KDET    | -0.20 | 2.26            | -0.14    | 4.28     | KP58    | -0.27 | 2.46            | 0.02     | 3.85     |
| KDKK    | -0.22 | 2.48            | 0.01     | 3.58     | KTDZ    | -0.22 | 2.38            | -0.04    | 3.84     |
| KDTW    | -0.21 | 2.37            | 0.06     | 4.57     |         |       |                 |          |          |

### 3 Trend detection considering bias caused by inconsistent averaging time

Similar as that in Li et al. (2017), the annual maximum wind speed for eight studied stations as shown in Figure 1 are extracted to conduct trend detection. It was found in Li et al. (2017) that the t-test and the Mann-Kendall Tau trend analysis both perform equivalently in detecting trends for annual maximum wind speeds. Therefore, only the t-test was used in this study due to its simplicity of application. The t-test trend detection can be simply done by fitting the analyzed random variable,  $Y$ , into a linear function of time ( $T$ ) such that  $Y = aT + b$ . The t-score for this test can then be calculated by,

$$[5] t_{score} = \hat{a} / \left( \sqrt{\sum_{i=1}^N (y_i - \hat{a}x_i - \hat{b})^2 / (n - 2)} / \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \right)$$

where  $n$  is the total number of years of record used in trend analysis.

The value of 0.05 for  $t_{score}$  is used as a critical value (Craparo, 2007). Any detected trend with a t-score value below 0.05 can be considered as statistically significant. The trend detection analysis is conducted for original annual maximum wind speeds for each of eight stations first. The slope of the fitted linear regression line is shown as a red line in Figure 2 for each station. The detected trend for the original annual maximum wind speeds (i.e. with no correction applied before 1985 for averaging time effect) for each station is same as that observed in Li et al. (2017).

Also, the conversion of 1-minute average wind speed to a 2-minute mean wind speed was conducted on the wind speed records prior to 1985 for each station. The conversion was randomly applied to the peak wind speeds of independent storms. Annual maximum wind speeds were then extracted from the corrected peak wind speeds of the independent storms for records prior to 1985. The annual maximum wind speeds for records after 1985 are extracted from the original wind speed records. By conducting this conversion, the averaging time of the mean wind speed could be considered as consistently 2-minute throughout the entire record. Such a 2-minute mean wind speed prior to the top of each hour is considered as a surrogate to represent the hourly mean wind speed of the hour. The effect of using the last 2-minute mean wind speed as a surrogate for the true hourly mean wind speed on the trend detection and calibration of design wind speed are out of the scope of this study. The same trend analysis was carried out for these corrected annual maximum wind speed samples for each station. For each station, the random conversion was carried out 10,000 times. For each conversion, one trend analysis was carried out. The ranges of slopes of the fitted linear regression lines are illustrated in Figure 2 as gray zones.

The linear regression line can be expressed by,

$$[6] Y = aX + b$$

where  $Y$  is the annual maximum wind speed and  $X$  is the number of years from the first year.

In trend detecting, the slope parameter,  $a$ , is of particular interest. Therefore, Equation [6] is re-arranged and normalized by the total number of years.

$$[7] Z = (Y - b) / X_n = a t$$

where  $t = X / X_n \in (0, 1]$ .

The slopes of regression lines in Figure 2 indicate the linear change of annual maximum wind speed (km/h) per year.

It can be observed that Hamilton, Niagara Falls, Toronto, Waterloo and Windsor have the largest variability of the detected linear trend, indicated by having the widest gray zones. However, the conversion of 1 minute to two minute averages does not change the direction of the trends observed at any of the stations.

The t-score was evaluated for each trend analysis. When the t-score value is less than 0.05, it was considered as statistically significant. The histogram of the evaluated trends for each trend analysis is

plotted in Figure 3. It can be observed that for stations in Niagara Falls, Toronto and Waterloo, the smallest t-scores for these stations were greater than 0.05.

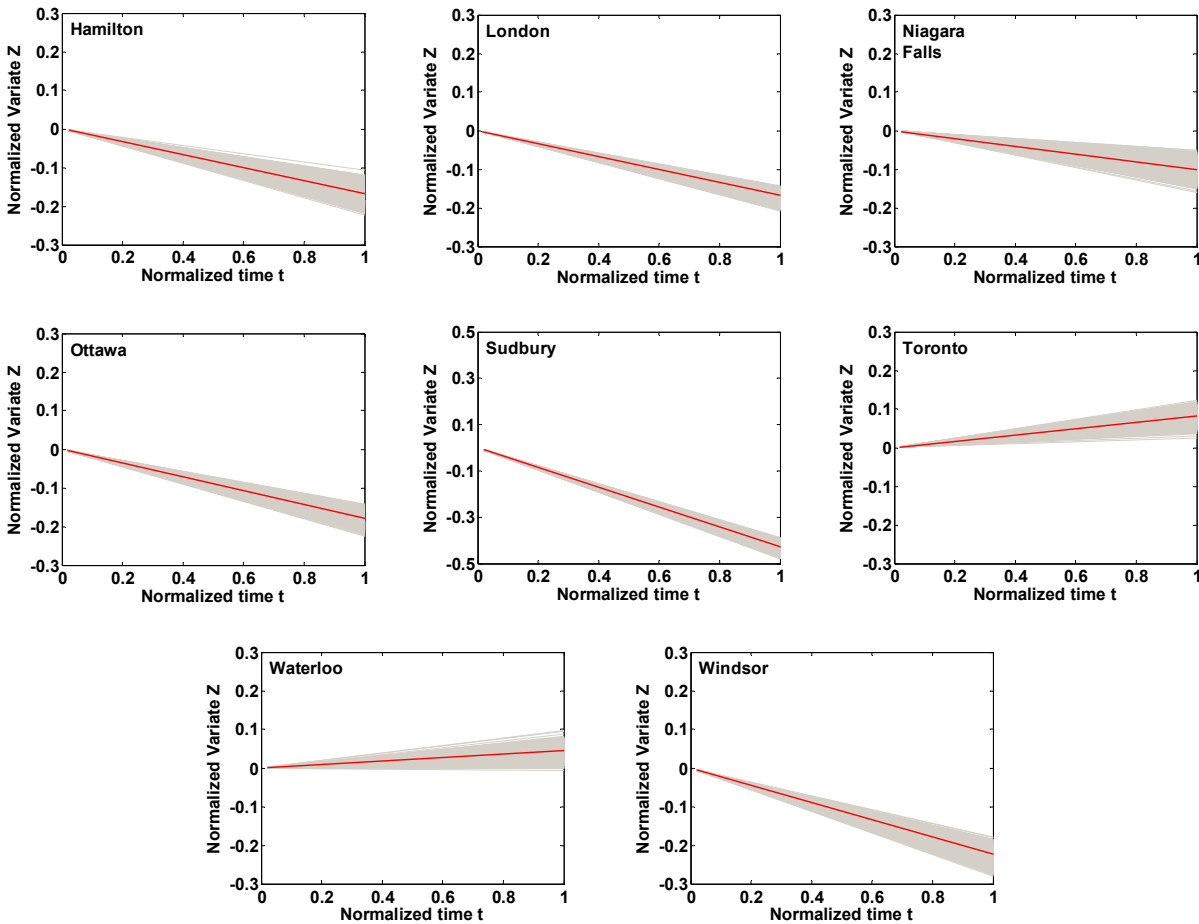


Figure 2: Normalized linear trend: line from original data; gray zone including correction for averaging time prior to 1985.

This indicates that there is no statistically significant trend being detected for these stations, even if the impact of the changes of averaging time is considered. For the stations in London, Ottawa, Sudbury and Windsor, the bulk of the simulations gave a t-score less than 0.05, indicating a significant trend. For Hamilton, there are chances that the trend could be considered as statistically significant. It can be also observed that if the original detected trend has a t-score close to the critical value, there will be more impact of the change of averaging time on the detected trend. For those cases where the t-score is either very large (close to 1.0) or very small, the variability of the detected trend caused by the change of averaging time is unlikely to change the trend detected in the original observations without averaging time correction.

The above trend detection was conducted for wind climate mixed with synoptic wind and thunderstorm wind. The same detection analysis was conducted for synoptic wind only, i.e., filtering out the thunderstorm winds. It was found that majority of the previous observations remain the same, except that the t-score for Hamilton, London and Windsor could be impacted. Figure 4 shows that after filtering thunderstorm wind, the smallest t-score for Hamilton becomes greater than 0.05, which indicates no statistically significant trend can be confirmed. For London and Windsor, although the majority of the t-scores are still smaller than 0.05, there are chances that the t-score could be greater than 0.05. This indicates that separating wind mechanism could have some impact on the statistical significance of detected trends for some stations.

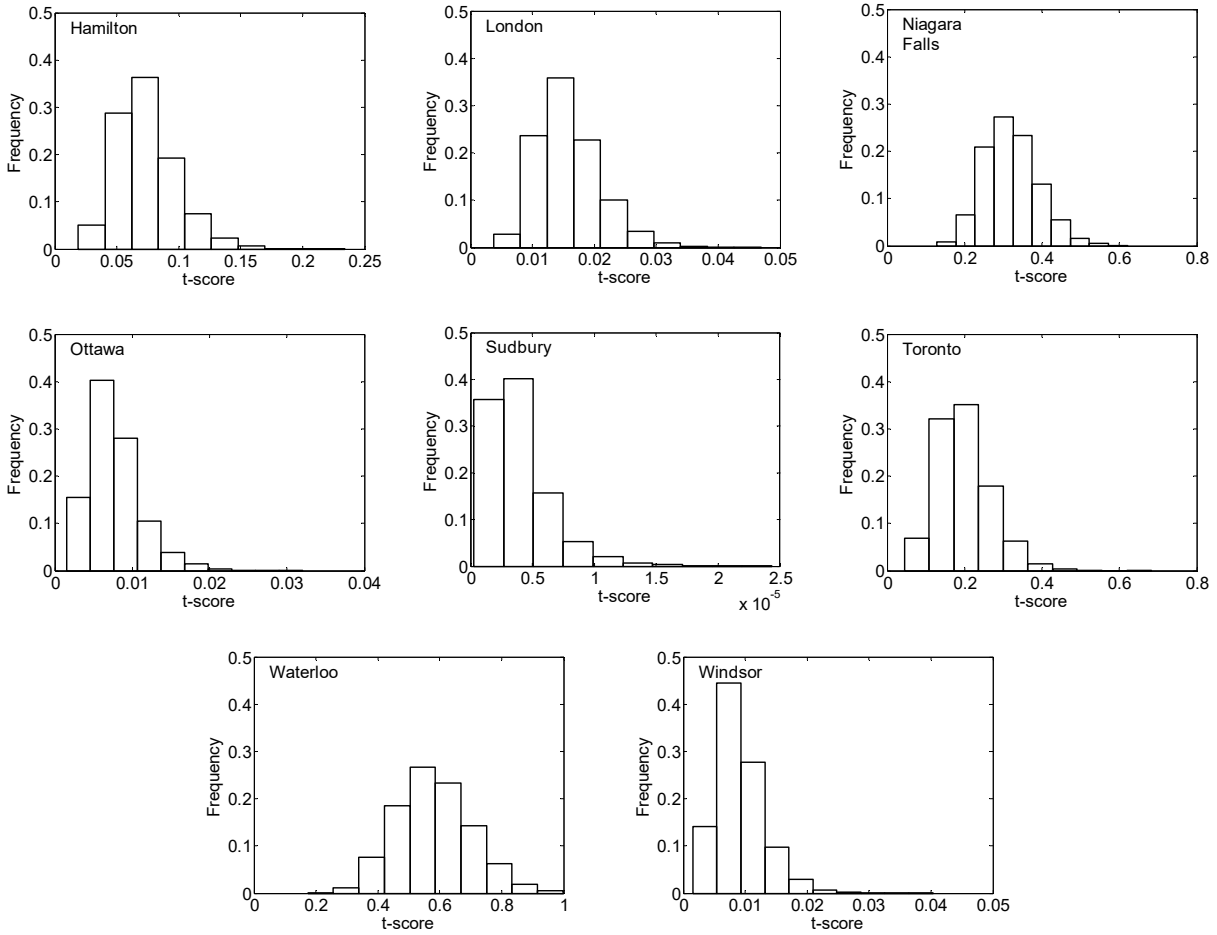


Figure 3: Histogram of t-score.

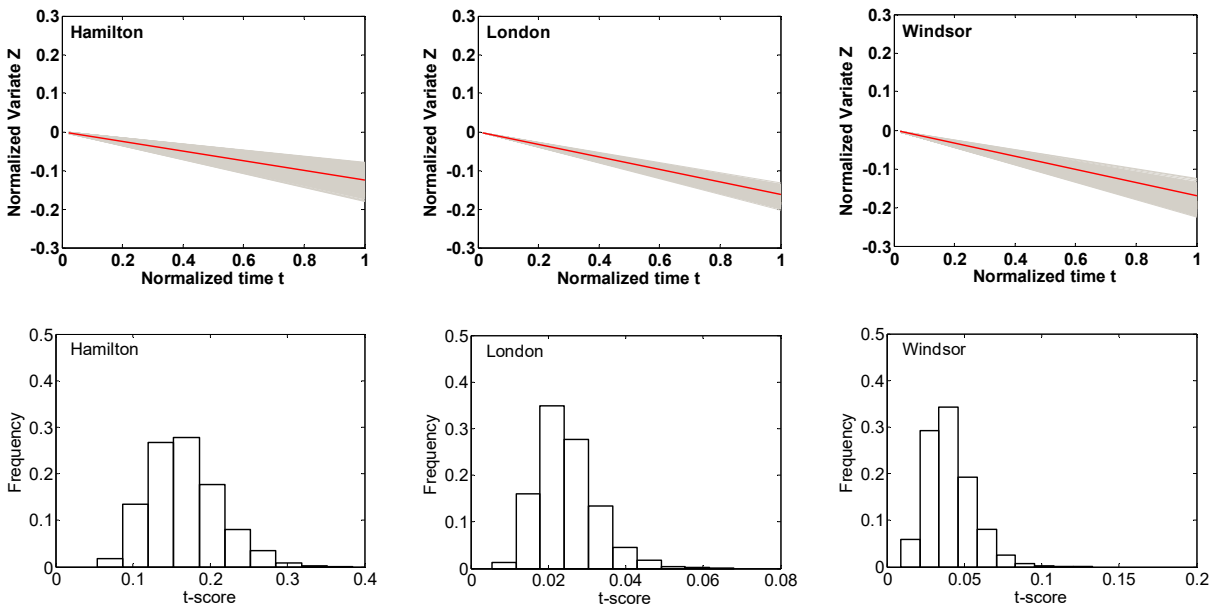


Figure 4. Linear trend and histogram of t-score for selected stations filtering thunderstorm wind

#### 4 Effect on design wind speeds

The final purpose of all trend analysis is to evaluate the effect of any detected trend on predicted extreme wind speeds as a function of return period. Therefore, a further analysis was conducted to investigate the effect of the change of averaging time on the extreme value analysis used to determine design wind speed.

In the extreme value analysis, the annual maximum wind speeds either from the original dataset or corrected to consistent 2-minute mean wind speeds were used to fit a Gumbel distribution. It should be noted that no separation of different wind mechanism was done before carrying out the extreme value analysis. As no significant increasing trend was detected for these stations, it is not urgent to apply non-stationary extreme value analysis. The return period wind speeds were calibrated through conventional extreme value analysis. The generalized least square method (Hong et al. 2014) was used to estimate the distribution parameters. The Gumbel distribution can be expressed by,

$$[6] F(x) = \exp(-\exp(-(x-u)/a))$$

where  $u$  and  $a$  are distribution parameters.

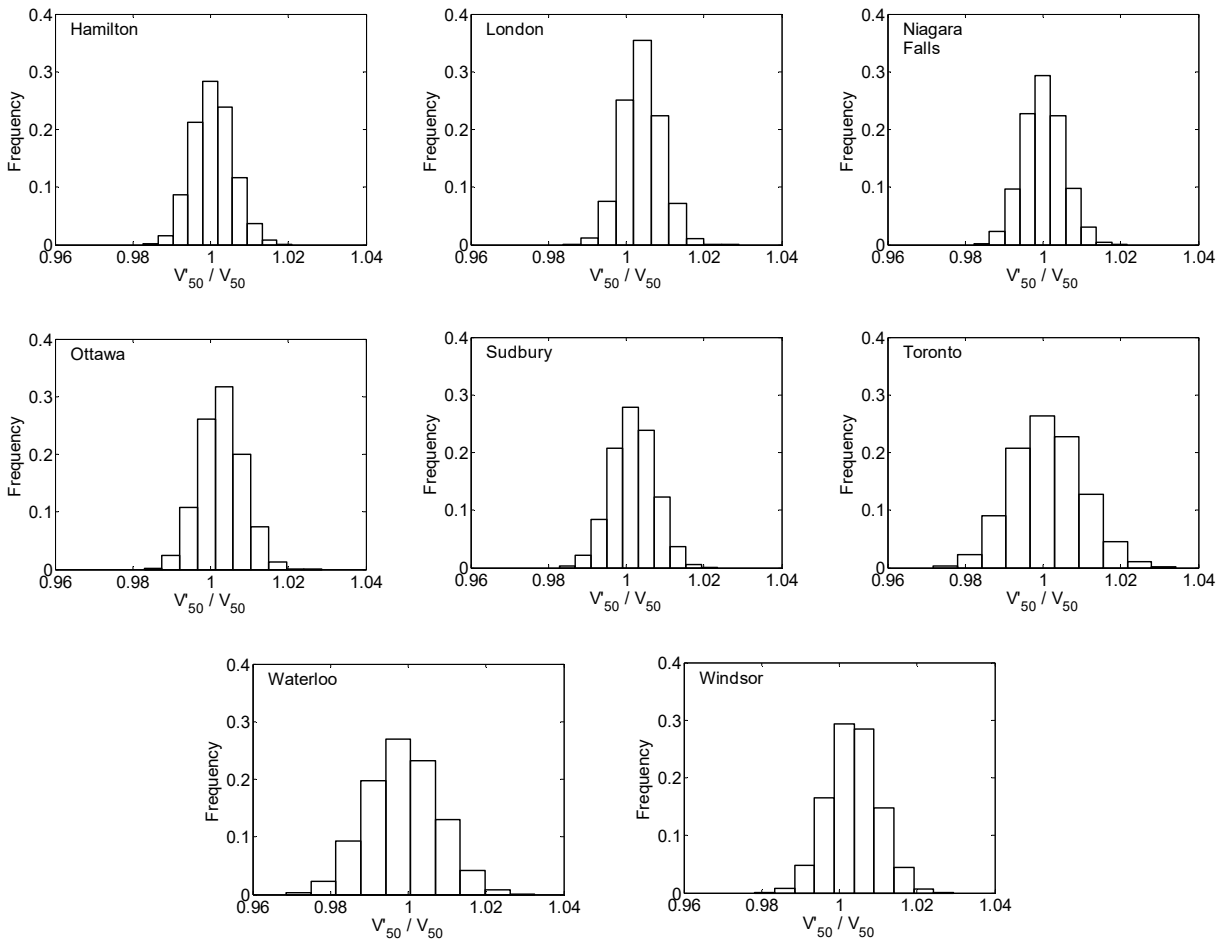


Figure 5: Ratio of 50-year design wind speed with correcting averaging time ( $V'_{50}$ ) to that calibrated from original dataset ( $V_{50}$ ).



The return period wind speeds can then be evaluated from the fitted Gumbel distribution using,

$$[7] V_T = u + a (-\ln(-\ln(1-1/T)))$$

where  $T$  = return period. In the present study,  $T = 50$  years, the current reference return period for structural loads in the National Building Code of Canada.

The histograms of ratio of the 50-year return period wind speed derived for each station considering conversion of averaging time ( $V'_{50}$ ) to that derived from original data ( $V_{50}$ ) are shown in Figure 5 based on the Monte Carlo simulation approach described previously. It can be observed in Figure 5, the variation of the design wind speed for each station is small. The range of the variation of the calculated design wind speed for all stations is within 3%. This uncertainty seems to be small for engineering practice compared to uncertainties caused by other factors. A similar observation was made by Morris (2009) who used one-minute continuous wind speed records.

## 5 Discussion and Conclusion

This study explores the effect of the change of averaging time of mean wind speed in historical records on the detection of trends in annual maximum speeds. The historical 1-minute mean wind speed was randomly converted to 2-minute mean wind speed, which is used to represent the hourly mean wind speed for the hour. For converting 1-minute mean wind speed to 2-minute mean wind speed, 15 ASOS one-minute wind speed datasets that contained 2-minute moving average wind speeds recorded at each 1-minute interval was used to derive a simple prediction model. The assumption for this prediction model was reasonably supported by employing a long term continuous 1-minute mean wind speed dataset obtained at Bratt's Lake Saskatchewan. 10,000 simulations were conducted for each station using Monte Carlo techniques to randomly convert 1-minute mean wind speed to 2-minute mean wind speeds for the record prior to 1985. Trend detection was carried out for each of these samples and its statistical significance determined using the t-test method. While it was found that the slope of the linear trend varied due to the change of averaging time, the variations were not enough to change the direction of trends. There are few chances that an insignificant trend can be converted into a significant one, or vice versa. Separation of wind mechanism (synoptic and thunderstorm) combined with averaging time effect has some impact on the significance of the detected trend for some stations. A further investigation was carried out to investigate the impact of the change of averaging time on the determination of design wind speeds. It was found that there could be up to 3% variation of the design wind speed due to the change of averaging time. This uncertainty seems to be small compared to other uncertainties that exist in the evaluation of design wind loads.

Other sources of uncertainty in trend detection identified in Li et al. (2017), data standardization for terrain effects and long period large scale cyclical climate influences, should be further studied in the future.

## Acknowledgements

Financial support from the National Research Council of Canada is gratefully acknowledged. The authors appreciate the assistance of Jarrett Philip, of Environmental and Climate Change Canada (ECCC), and Bob Morris, consultant to ECCC, in acquisition of some of the wind speed data used in present study.

## References

- Craparo, Robert M.* 2007. Significance level. In Salkind, Neil J. Encyclopedia of Measurement and Statistics 3. Thousand Oaks, CA: SAGE Publications. pp. 889–891. ISBN 1-412-91611-9
- Hong, H. P., Li, S. H., & Mara, T. G.* 2013. Performance of the generalized least-squares method for the Gumbel distribution and its application to annual maximum wind speeds. *Journal of Wind Engineering and Industrial Aerodynamics*, 119, 121-132.

*Li, S.H., Irwin, P., Kilpatrick, J., Gibbons, M., Valerie, S. 2017. A review of historical extreme wind speeds in a changing climate at some major Canadian Cities. CSCE, Vancouver, B.C., Canada.*

*Morris, C.A. 2009. Wind interim report on the updating of the design wind speeds in the National Building Code of Canada for the task Group on climatic loads. National Research Council of Canada, Ottawa, Canada.*