

Electronic Thesis and Dissertation Repository

8-18-2020 3:00 PM

Tornado Occurrence Modelling and Equivalent Tornado Design Wind Profile for Canada

Qian Huang, *The University of Western Ontario*

Supervisor: Hong, Hanping, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Civil and Environmental Engineering

© Qian Huang 2020

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Civil Engineering Commons](#), and the [Environmental Engineering Commons](#)

Recommended Citation

Huang, Qian, "Tornado Occurrence Modelling and Equivalent Tornado Design Wind Profile for Canada" (2020). *Electronic Thesis and Dissertation Repository*. 7288.

<https://ir.lib.uwo.ca/etd/7288>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

The tornadoes can be devastating to structures and infrastructure systems. They can cause fatalities and economic losses. In this thesis, I systematically assess the tornado hazard for Canada. I first investigate the use of three spatial point processes (i.e., Poisson, zero-inflated Poisson, and negative binomial models) in a hierarchical structure to predict the tornado occurrence rate by considering the annual cloud-to-ground lightning flash density and annual thunderstorm days as covariates. I consider an existing historical tornado catalogue with reported tornado events up to 2009 and include the population bias effect in estimating the tornado occurrence rate. Both the Bayesian technique and the maximum likelihood method are used to estimate the model parameters. It was shown that the NB model outperforms the other two judged based on statistical criteria such as the Akaike information criterion or Bayesian information criterion.

As the existing tornado catalogue is not up to date, a new tornado catalogue with reported events up to 2019 is developed. By using the existing and new tornado catalogue, the modelling of tornado occurrence rate focused on southern Ontario – a region that is prone to tornado in Canada – is carried out. Also, an assessment of the rate of the tornado striking a site of interest is estimated by incorporating the statistics of the tornado path characteristics. The striking rate and an adopted probabilistic tornado wind field model are then used to estimate the tornado wind velocity hazard maps. It is shown that for a point-like structure, a simple tornado design wind profile could be developed based on the specified return period value of the tornado wind velocity at different heights. This concept is extended for a line-like structure based on the return period values of the bending moment and shear force along the height of the structure. In all cases, the tornado design wind profile is expressed as a function of the tornado wind speed at 10 m height.

The development of the tornado wind hazard map and tornado design wind profile is extended to Canada. For the development, the Poisson model, zero-inflated Poisson model, and negative binomial model, as well as the adaptive Gaussian kernel smooth technique, are used to model tornado occurrence. The modelling again takes into account the tornado reporting bias due to population density and uses the cloud-to-ground lightning flash density and the thunderstorm

days as the explanatory variables. The tornado wind velocity hazard maps for Canada, in terms of wind velocity at 10 m height above the ground surface for different return period T , $V_T(10)$, are presented. Most importantly, it is shown that a simple equivalent tornado design wind profile can still be developed for regions with significant different tornado activities. The developed tornado design wind profile only depends on $V_T(10)$, which greatly facilitates its potential use for structural design subjected to tornado wind load and its possible implementation in structural design code and standards. It is shown that the tornado winds could dominate the wind hazard as the length of the footprint of an infrastructure system or the area of the footprint of a structure increases. This indicates that the consideration of tornado winds is necessary for a spatially extended building complex and infrastructure system that are critical for the safe operation of the society.

Keywords

Tornado, stochastic point process, negative binomial process, striking rates, equivalent wind profile, tornado wind hazard map.

Summary for Lay Audience

Tornadoes are a very destructive natural hazard. In this thesis, a systematic assessment for tornado wind hazard is developed. The steps for the assessment involve modelling the spatial distribution of tornado occurrence rate, estimating the tornado path striking a structure, and the evaluation of the tornado wind velocity when the structure is hit by a tornado. Since the tornado touchdown, tornado path, and tornado wind field are all random, the tornado hazard evaluation needs to take the randomness into account.

The modelling of tornado occurrence rate for Canada as well as for regions with significant tornado activities is carried out. For the modelling, the information on the population density, the thunderstorm activities, and the lightning activities are incorporated. An existing tornado catalogue and a newly developed tornado catalogue for this study are used. The best model among several considered models is identified based on statistical criteria. Also, tornado wind velocity maps are developed by considering the uncertainties in the tornado occurrence, the tornado damage path, and the tornado wind field. In addition, a simple-to-implement tornado design wind profile is developed. It is shown that although the design for tornado winds for a single small normal building may not be necessary based on the standard safety requirement, it could be very important for a large building complex and infrastructure with an extended footprint.

Co-Authorship Statement

The materials presented in this thesis are based on three joint-authored articles.

A version of Chapter 2, co-authored with H.P. Hong and W.J. Jiang, was submitted to a peer-reviewed journal for possible publication.

A version of Chapter 3, co-authored with H.P. Hong and W.J. Jiang, was submitted to a peer-reviewed journal for possible publication.

A version of Chapter 4, co-authored with H.P. Hong and others, will be submitted to a peer-reviewed journal for possible publication.

Acknowledgments

First and foremost, I would like to express my most sincere gratitude to my supervisor – Professor Han-ping Hong, for his undiminished encouragement, long-standing support and unfailing patience throughout my entire PhD life. His professionalism, broad knowledge and meticulous supervision not only guarantee the accomplishment of my research work but also enhance my problem-solving skills and help me to develop more scientific insights. The technologies and rigorous attitudes I gained in the PhD study are assets that will accompany me throughout the rest of my life.

I would also like to thank my thesis committee members: Dr. Wenxing Zhou, Dr. Horia Hangan, Dr. Jiandong Ren and Dr. Arnold Yuan for their careful reading and insightful comments which help to improve this thesis. Moreover, I am grateful to the knowledgeable faculty and diligent staff at the Department of Civil and Environmental Engineering for their support in various aspects. I acknowledge the financial support from the China Scholarship Council (CSC) and the University of Western Ontario.

I am always indebted to my parents and my husband. Their endless and unselfish love is the origin of motivation that makes me move forward. With their mental and financial support, I can pursue what I love without any worries.

Special thanks go to my friends and colleagues in London Ontario. I am so lucky to have them around me so that I can embrace an interesting and colorful non-academic life. The cheerful moments shared with them together will absolutely be part of happy memories and enduring friendships for the years to come.

Table of Contents

Abstract	ii
Summary for Lay Audience	iv
Table of Contents	vii
List of Tables	x
List of Figures	xii
List of Appendices	xvi
Nomenclature	xvii
Chapter 1	1
1 Introduction	1
1.1 Background and Literature review.....	1
1.2 Objectives of the current study	3
1.3 Chapter organization.....	4
References	5
Chapter 2.....	8
2 Statistical assessment of spatial tornado occurrence for Canada: modelling and estimation	8
2.1 Introduction.....	8
2.2 Spatial tornado occurrence modelling and model parameter estimation	11
2.2.1 Modelling the tornado occurrence as a spatial point process	11
2.2.2 Parameter estimation.....	14
2.3 Tornado catalogue and data used for estimating the tornado occurrence rate	16
2.4 Tornado occurrence rate estimation for Canada	19
2.4.1 Maximum likelihood estimates.....	19
2.4.2 Bayesian inference	25
2.5 Summary and conclusions	30

References	31
Chapter 3.....	36
3 Development of a simple equivalent tornado wind profile for structural design and evaluation	36
3.1 Introduction.....	36
3.2 Tornado occurrence rate and striking rate modelling	39
3.2.1 Tornado catalogue, explanatory variables, and population density for southern Ontario.....	39
3.2.2 Estimation of tornado occurrence rate for southern Ontario	43
3.2.3 Characteristics of tornado path and striking rate	48
3.3 Estimation of tornado wind velocity hazard	50
3.3.1 Assessment of probability distribution of tornado wind velocity	50
3.3.2 Assessment of probability distribution of bending moment and shear force	52
3.4 Tornado wind hazard evaluation, disaggregation, and mapping	55
3.4.1 Quantile of tornado wind velocity along the height.....	55
3.4.2 Disaggregation of tornado wind velocity hazard	56
3.4.3 Tornado hazard mapping	58
3.5 Adequacy of wind profile for the line-like structure	60
3.6 Conclusions.....	62
References	63
Chapter 4.....	68
4 Tornado wind hazard mapping and tornado design wind profile for Canada.....	68
4.1 Introduction.....	68
4.2 Developing tornado occurrence rate model	71
4.2.1 Tornado catalogue.....	71
4.2.2 Population density, cloud-to-ground lightning flash density and thunderstorm days.....	74

4.2.3	Stochastic modelling for tornado occurrence rate	77
4.2.4	Estimating model parameters to predict tornado occurrence rate.....	80
4.3	Tornado striking rate and database of probability distributions of wind velocity and its effects	85
4.3.1	Analysis procedure.....	85
4.3.2	Estimated tornado striking rate and assessed probability distributions of $V(z)$, $M(z,H)$ and $S(z,H)$	87
4.4	Tornado hazard mapping and tornado wind profile.....	90
4.4.1	Tornado hazard assessment for a single site	90
4.4.2	Tornado hazard mapping	95
4.4.3	Discussions	102
4.5	Conclusions.....	103
	References	105
Chapter 5	111
5	Conclusions and future research	111
5.1	Conclusions.....	111
5.2	Future research.....	112
Appendices	114
Appendix A:	Structure of the Bayesian hierarchical model.....	114
Appendix B:	Convergence issue with different probability detection model in MCMC	114
Appendix C:	Tornado statistics for southern Ontario	118
Curriculum Vitae	119

List of Tables

Table 2.1. Estimated model parameters considering ACGLF density as the covariate (the first entry represents the estimated value, and the second entry represents the estimated standard deviation of the model parameter).	21
Table 2.2. Estimated model parameters considering ATD as the covariate (the first entry represents the estimated value, and the second entry represents the estimated standard deviation of the model parameter).	23
Table 2.3. Estimated model parameters considering the ACGLF density and ATD as the covariates (the first entry represents the estimated value, and the second entry represents the estimated standard deviation of the model parameter).	24
Table 2.4. Estimated statistics of the model parameters based on the marginal posterior distribution.	29
Table 3.1. Reported tornadoes based on their intensity and the updated frequency distribution of tornado intensity.	40
Table 3.2. Cases considered for the statistical analysis of tornado occurrence (see Eq. (3.4). z_{1j} is used to represent ACGLF, and z_{2j} is used to represent ATD).	47
Table 3.3. Estimated model parameters for Cases A1 and B1 based on Bayesian hierarchical modelling technique (the first and second entries represent the mean and standard deviation of the model parameters).	47
Table 4.1. Models considered and the corresponding likelihood function	79
Table 4.2. Combinations of cases to be considered for the statistical analysis of tornado occurrence rate (z_{1j} is used to represent ACGLF density and z_{2j} is used to represent ATD)..	81
Table 4.3. Calculated AIC and BIC for the cases described in Table 2.	81
Table 4.4. Estimated model parameters for selected models (the first and second entries represent the estimated mean and standard deviation, respectively).	83

Table 4.5. Statistics of R_{ST} (The first and second entries represent the mean and the standard deviation). 99

List of Figures

Figure 2.1. Spatial distribution of reported historical tornado events from 1980 to 2009..... 16

Figure 2.2. Spatially varying population density (number of people per km²) based on 2001 Canadian census data: a) based on counties; b) based on regular grid cells..... 17

Figure 2.3. Plot of the ACGLF density (per year per km²) using data from 1999 to 2009. ... 18

Figure 2.4. a) Locations of the stations with thunderstorm winds recorded from 1980 to 2009; b) the estimated ATD per cell..... 19

Figure 2.5 Predicted tornado occurrence rate (per year per 10⁴ km²) by considering ACGLF density as the covariate and based on: a) Poisson model, b) ZIP model and c) NB model.... 22

Figure 2.6. Predicted tornado occurrence rate (per year per 10⁴ km²) by using the NB model a) using ATD as the covariate, and b) using both the ACGLF density and ATD as the covariates. 25

Figure 2.7. Results from the BGR diagnostics for r , α_0 , α_1 , α_2 , β and δ^2 considering the NB model with the ACGLF density and ATD as the covariates. 27

Figure 2.8. The marginal posterior distributions for r , α_0 , α_1 , α_2 , β and δ^2 by considering the NB model with the ACGLF density and ATD as the covariates..... 28

Figure 3.1. Tornadoes occurred in southern Ontario: a) For events from 1980 to 2009, b) for events from 1980 to 2019. 40

Figure 3.2. Estimated ACGLF, ADT: a) estimated ACGLF (per year per km²) using data from 1999 to 2009, b) Estimated ATD based on the identifier used in ECCC DLY04 database from 1980 to 2009, and c) Estimated ATD based on the identifier used in ECCC DLY04 database from 1980 to 2017..... 41

Figure 3.3. Estimated population density (number of people/km²): a) population density based on 2001 Canadian census data, e) population density based on 2011 Canadian census data..... 43

Figure 3.4. Estimated spatially varying tornado occurrence rate (number/(year×km²)) by using adaptive Gaussian kernel smoothing: a) based on TC80-09, and b) based on TC80-19. 43

Figure 3.5. Estimated spatially varying tornado occurrence rate (number/(year×km²)) based on the NB model with the model parameters shown in Table 3: a) based on Case A1, and b) based on Case B1. 47

Figure 3.6. Steps of calculating the striking rates for different F_{Aj} 49

Figure 3.7. Spatially varying tornado striking rate with intensity F_{Aj} by using the tornado occurrence rate shown in Figure 5a. 50

Figure 3.8. Steps to evaluate the probability distributions of wind velocity, and bending moment and shear force for a line-like structure. 51

Figure 3.9. Assessed $P(V(z) < v | F_{Aj})$ for selected z values. 52

Figure 3.10. Estimated $P(M(z, H) < m | F_{Aj})$ the normalized bending moment equals $M(z, H) / (C_D \rho_{air} B / 2)$ 54

Figure 3.11. Estimated $P(S(z, H) < s | F_{Aj})$, where the normalized shear force equals $S(z, H) / (C_D \rho_{air} B / 2)$ 54

Figure 3.12. Estimated quantiles of $V(z)$, $V_T(z)$, (3-s gust wind velocity) for London, Ontario, and fitted Eq. (9) for z up to 80 m. 56

Figure 3.13. Identified wind profile based on disaggregation analysis for $T = 5000$ and 10000 years. 58

Figure 3.14. Tornado wind velocity, $V_T(10)$, (3-s gust wind velocity) map for southern Ontario. 59

Figure 3.15. Variation of model parameters α_1 and α_2 as functions of $V_T(10)$ 60

Figure 3.16. Estimated $M_T(z,H)$ and $S_T(z,H)$ for London Ontario.	62
Figure 4.1. Spatial distribution of the reported tornado event: a) Reported events in EC-T80-09, and b) Reported events in T80-19.....	73
Figure 4.2. Spatially smoothed tornado occurrence rate (per (year $\times 10^5$ km 2)) based on the tornado catalogues for the reported tornado events with locations depicted in Figure 1: a) Based on reported events depicted in Figure 1a, and b) Based on reported events depicted in Figure 1b.	75
Figure 4.3. Spatially varying population density (number of people per km 2): a) Based on 2001 Canadian census data, and b) based on 2011 Canadian census data.	75
Figure 4.4. Estimated ACGLF density (per year per km 2) using data from 1999 to 2009.	76
Figure 4.5. Estimated ATD (number of days/year): a) using records from 1980 to 2009 and b) using records from 1980 to 2017.	77
Figure 4.6. Predicted tornado occurrence rate (per year per 10 4 km 2) based on the NB model: a) using EC-T80-09 and ACGLF as the explanatory variable; b) using EC-T80-09 and ACGLF and ATD as the explanatory variables; c) using T80-19 and ACGLF as the explanatory variable; b) using T80-19 and ACGLF and ATD as the explanatory variables.	84
Figure 4.7. Identified contour lines with the annual mean tornado occurrence rate (per km 2 per year) equal to 5×10^{-5} , (I: Region prone to significant tornadoes), 10^{-5} (II: Region prone to tornadoes), and 10^{-6} (III: Region with rare tornado occurrence).	85
Figure 4.8. Spatially varying annual striking rate γ_{Aj} (number per 10 5 years) using occurrence rate shown in Figure 6b.....	88
Figure 4.9. Spatially varying annual striking rate γ_{Aj} (number per 10 5 years) using occurrence rate shown in Figure 2b.....	89
Figure 4.10. Estimated $P(V(z) < v F_{Aj})$, $P(M(z, H) < m F_{Aj})$ and $P(S(z, H) < s F_{Aj})$..	90

Figure 4.11. Estimated $V_T(z)$, for Toronto and Winnipeg and comparison of the probability distributions of annual maximum synoptic wind velocity and of tornado wind velocity (the wind speed is presented in terms of the 3-second gust mean wind velocity (km/h)). 92

Figure 4.12. Values of α_1 and α_2 as functions of $V_T(10)$ for Toronto and Winnipeg 93

Figure 4.13. Estimated $M_T(z, H)$ and $S_T(z, H)$ and their approximations. 95

Figure 4.14. Estimated wind hazard maps: a) to c) Tornado wind hazard maps by considering the tornado occurrence rate shown in Figure 6b (i.e., striking rates shown in Figure 8) and for $T = 10,000, 50,000,$ and $100,000$ years, respectively; d) synoptic wind hazard map for $T = 500$ years. 96

Figure 4.15. Tornado wind velocity hazard maps based on the occurrence rate model shown in Figure 2b (i.e., the striking rate shown in Figure 9): a) for $T = 50,000$ years, and b) for $T = 100,000$ years. 97

Figure 4.16. Spatially varying annual striking rate γ_{Aj} (number per 10^5 years) using occurrence rate shown in Figure 6b for a structure with footprint represented by a circle with a diameter of 100 m. 99

Figure 4.17. Values of α_1 and α_2 as functions of $V_T(10)$ and fitted simple empirical relations. 101

Figure 4.18. Calculated reduction factor if the equivalent tornado wind profile defined by Eqs. (10), (13) and (14) is used to calculate the bending moment and shear force: a) for bending moment, b) for shear force. 102

List of Appendices

Appendix A: Structure of the Bayesian hierarchical model	114
Appendix B: Convergence issue with different probability detection model in MCMC	114
Appendix C: Tornado statistics for southern Ontario	118

Nomenclature

Chapter 2

Y_{ij}	Reported number of tornado occurrence before the time t at the j -th or cell
N_{ij}	Actual number of the tornado occurrence
K_{tji}	A Bernoulli random variable
p_{ij}	Probability that the i -th tornado occurrence at the j -th site is reported
λ_j	Time-independent occurrence rate
β	Model parameter
α_{kj}	Model parameter
γ_j	Population density
z_{kj}	Explanatory variables
ε_j	Error term of the prediction equation
n_{Aj}	Number of adjacent cells of the j -th cell
π_{ij}	The additional probability that the tornado counts is equal to zero
δ^2	Model parameter
σ^2	Model parameter
$\ln(\lambda_j)$	Actual tornado occurrence rate
p_{NB}	Probability of non-occurrence
η_j	Reported tornado occurrence rate
r	Model parameter

L	Likelihood function
M_c	Total number of cells considered
y_j	Reported tornado occurrences
z_{1j}	ACGLF density
p_{ij}	Probability of reporting of an actually occurred tornado
z_{2j}	ATD density
α_0	Model parameter
α_1	Model parameter
α_2	Model parameter

Chapter 3

F_i	Tornado intensity, Fujita scale
EF_i	Tornado intensity, Enhanced Fujita scale
N_{ij}	Number of tornado events
r	Model parameter
p_{NB}	Probability of non-occurrence
μ_j	Mean of tornado events at a counting cell
Y_{ij}	Reported tornado events
p_{ij}	Probability of an occurred tornado within the counting cell
γ_j	Population density
η_j	Number of tornadoes per year per km ²

β	Model parameter
z_{kj}	Explanatory variables
ε_j	Residual
n_{Aj}	Number of adjacent cells of the j -th cell
δ^2	Model parameter
α_0	Model coefficient
α_k	Model coefficient
M_c	Total number of cells for the considered region
γ_{Aj}	Tornado striking rate
L	Maximum likelihood function
θ	Set of parameters used for the parametric model
F_{Aj}	Actual tornado intensity
B	Structure width
H	Structure height
M	Bending moment
S	Shear force
ϕ	Wind direction
C_D	Drag coefficient
ρ_{air}	Air density
r_{Ai}	Striking rate for F_i tornado

$V_T(z)$	T -year return period value of tornado wind velocity
$V_{TA}(z)$	Equivalent tornado wind profile
T	Return period
α_1	Model parameter
α_2	Model parameter
ξ	Positive scaling factor

Chapter 4

Y_{tj}	Reported tornado occurrence in the j -th cell for the observation period of t years
N_{tj}	Actual number of the tornado
X_{tji}	Independent and identically distributed random variable according to the Bernoulli distribution
p_{tj}	Probability of reporting an occurred tornado event
λ_j	Tornado occurrence rate per year
α_k	Model parameter
α_0	Model parameter
n_K	Number of explanatory variables
π_{tj}	Zero tornado occurrence probability
p_{NB}	Probability of non-occurrence
r	Model parameters
γ_j	Population density

ξ	Model parameters
L	Maximum likelihood function
θ	Set of parameters used for the parametric model
p	Total number of model parameters to be estimated
α_1	Model parameter
α_2	Model parameter
β	Model parameter
F_{Aj}	Actual tornado intensity
F_i	Reported or classified tornado intensity
$V(z)$	Maximum wind speed experienced by the structure
H	Structure height
$M(z,H)$	Maximum bending moment
$S(z,H)$	Maximum shear
$v(x \theta)$	Wind velocity
θ	Motion direction
C_D	Drag coefficient
γ_{Aj}	Annual striking rate
T	Return period
$V_T(z)$	T -year return period value of tornado wind velocity
$M_{TA}(z, H)$	T -year return period value of bending moment

$S_{TA}(z, H)$	T -year return period value of shear force
D	Circle diameter of represented structure footprint
R_{ST}	Ratio of the striking rate

Chapter 1

1 Introduction

1.1 Background and Literature review

High-intensity wind events such as tornadoes and downbursts can be very destructive to structures and infrastructure systems. A tornado is a rapidly rotating column of air extending from the base of a thunderstorm down to the ground. Statistics indicate that about 100 tornadoes per year occur in Canada. An incomplete tornado catalogue is available in Canada. According to Etkin et al. (2001), the notes compiled by M.J. Newark indicate that the first report of a tornado in Canada, which occurred on 1 July 1792, leveled all houses in its path and uprooted trees along a narrow track. The first systematic Canadian tornado catalogue was established by Newark (1984), indicating that “much of Canada east of the large Manitoba lakes is an extension of the great tornado-prone region centered over the south-central United States.” The database was subsequently updated by Sills et al. (2004) for Ontario. Tornado catalogue for Prairie and Northern Region in Canada from 1826 to 2010 was compiled by Patrick McCarthy (2011, personal communication with H.P. Hong). Additional information on tornado occurrences in Canada is described in the Environment and Climate Change Canada (ECCC) website (<https://open.canada.ca/data/en/dataset/fd3355a7-ae34-4df7-b477-07306182db69>), and in Wikipedia (https://en.wikipedia.org/wiki/List_of_Canadian_tornadoes_and_tornado_outbreaks#cite_ref-ehfs_1-0; & https://en.wikipedia.org/wiki/List_of_21st-century_Canadian_tornadoes_and_tornado_outbreaks). The path orientation, length, and width for each listed tornado event are not always available.

To assess the tornado hazard, a simple-to-use tornado wind field model is needed to reduce the freedom present in the full-scale observations. The wind field model should allow for the evaluation of wind velocities along the three dimensions analytically and statistically. That means, given a point in the wind field, its experienced vertical, tangential and radial wind speeds could be calculated and the parameters of wind field could be sampled from the known distributions. There are numerous attempts in the literature toward establishing

such tornado wind field model. For example, as early as 1880s, a vortect model developed by Rankine (1882) mimics the tornado-like flow behavior. Kuo (1971) modelled the three-dimensional flow in the boundary layer of a tornado-like vortex. Dunn and Twisdale (1979) extended Kuo (1971) by containing variable flow features which are significant in missile trajectory prediction and are consistent with tornado path geometry statistics. Quite recently, Gillmeier et al. (2018) reviewed the existing tornado wind field models and the fundamental assumptions behind them comprehensively. In this thesis, we adopt the model of Dunn and Twisdale (1979), which is elaborated in Twisdale et al. (1981), to depict the tornado wind flow as well as the uncertainties associated with it.

A qualitative tornado hazard assessment for Ontario was presented by Newark (1984) using the established database, indicating that, on average, there is an F3 tornado in every five years in southwestern Ontario. Based on very simplistic and practical assumptions, Newark (1991) suggested design wind velocity pressure by considering the design basis tornadoes for different regions. More sophisticated modelling and assessment for the tornadic wind hazard for southern Ontario were carried out in Banik et al. (2007). Their estimate was focused on the estimation of the tornado occurrence rate, tornado intensity distribution, probabilistic models of the tornado path length and path width, and distribution of the annual maximum tornado wind velocity at a site within southern Ontario. It was considered that the probabilistic tornado wind field in Twisdale (1978) and Twisdale et al. (1981) can be used adequately to estimate tornadic wind hazard, and that the spatial inhomogeneity of tornado occurrence near a site of interest can be ignored. They concluded that, at 10 m height, the tornado winds do not dominate wind hazard for specifying the design wind load for the buildings belonging to the normal building category defined in the National Building Code of Canada (NBCC) (NRC 2015). However, the tornado winds dominate the wind hazard for the design of very high consequence structures (e.g., nuclear power plants) or transmission lines extended over a few tens of km (Banik et al. 2008). Similar conclusions were arrived in Tan and Hong (2010) by considering inhomogeneity of tornado occurrence near the site of interest. Besides, it was observed that the consideration of uncertainty in tornado occurrence rate through the application of the hierarchical Bayesian modelling technique (Wikle and Anderson 2003) only affects negligibly the estimated tornadic wind hazard for southern Ontario (Tan 2008). The use

of the Bayesian hierarchical model, including the consideration of the impact of local population density on observed tornado occurrences and the predictive relationship between the cloud-to-ground (CG) lightning flash, were considered in Cheng et al. (2013) in mapping the tornado occurrence for Canada. Cheng et al. (2016) further applied such an approach to model the tornado occurrence in North America. However, the probabilistic tornadic wind-velocity hazard map for Canada is unavailable.

One of the possible weaknesses of the tornadic wind hazard assessment presented in Banik et al. (2007) and Tan and Hong (2010) is that the tornadic wind hazard assessment was presented only in terms of wind velocity at 10 m height. The use of wind hazard expressed in terms of 10 m height wind velocity for synoptic winds in structural design codes is well-accepted practice since the along height wind profile of synoptic winds is relatively consistent and can be represented by log wind profile or power-law wind profile (NRC 2015). However, this is not the case for tornadoes (Twisdale et al. 1981), Wurman and Gill (2000), Hangan and Kim (2008), and Natarajan and Hangan (2012), and Refan and Hangan (2016). Therefore, the tornadic wind hazard maps expressed in terms of 10 m height wind velocity or of the tornado occurrence rate are not very practical for engineers to design new and evaluate existing structure and infrastructure systems. For such a tornadic wind-velocity hazard map to be useful for Canada, additional investigation and recommendations on how to codify the tornadic wind profile can be valuable for structural analysis and design.

1.2 Objectives of the current study

The main objectives of this study include:

- 1) To develop a predicting model for tornado occurrence incorporating richer and more reliable meteorological information and taking in account the population bias effect as well as the overdispersion feature of data.
- 2) To map the tornado hazard over the southern Ontario, as well as the whole Canada, using the newly developed spatial tornado occurrence model.

3) To develop a simple parametric equivalent along-height tornado profile for evaluating tornadic wind loading.

1.3 Chapter organization

In Chapter 2, a preferred predicting model for the spatially varying tornado occurrence rate is developed for Canada. The development takes into account the most commonly used spatial stochastic models, and the underreporting due to population density. It incorporates the annual average cloud-to-ground lightning flash (ACGLF) density and annual average thunderstorm days (ATD) as covariates in the predicting model. The model parameters estimation is carried out by using both the maximum likelihood method and the Bayesian inference.

In Chapter 3, a tornado occurrence rate model is first developed by considering the reporting bias due to population density and by using the cloud-to-ground lightning flash density and thunderstorm days per year as the explanatory variables. The estimation of the quantile of the tornado wind velocity hazard along the height for southern Ontario – a tornado-prone region in Canada – is then carried out. Also, a simple equivalent along height tornado design wind profile for a line-like structure is proposed by considering the bending moment and shear force along the height of the structure. Moreover, the concept of disaggregation analysis is employed to identify the scenario tornado wind profiles.

In Chapter 4, a systematic assessment of tornado wind velocity hazard is presented for Canada based on reported historical tornado events. For the assessment, the Poisson model, zero-inflated Poisson model, and negative binomial model, as well as the adaptive Gaussian kernel smooth technique, are used to model tornado occurrence. The modelling again takes into account the tornado reporting bias due to population density and uses the cloud-to-ground lightning flash density and the thunderstorm days as the explanatory variables. The tornado wind velocity hazard maps for Canada, in terms of wind velocity at 10 m height above the ground surface and return period T , $V_T(10)$, are developed based on an available probabilistic tornado wind field model. Also, a simple equivalent tornado

design wind profile is developed based on the return period value of tornado wind speed at different heights. The developed tornado design wind profile only depends on $V_T(10)$, which greatly facilitates its potential use for structural design subjected to tornado wind load and its possible implementation in structural design code and standards.

Finally, Chapter 5 summarizes the main conclusions drawn from the previous chapters. Contributions of this work are also highlighted in this chapter. Recommendations are given for the future work.

References

- Banik, S. S., Hong, H. P., & Kopp, G. A. (2007). Tornado hazard assessment for southern Ontario. *Canadian Journal of Civil Engineering*, 34(7), 830-842.
- Banik, S. S., Hong, H. P., & Kopp, G. A. (2008). Assessment of tornado hazard for spatially distributed systems in southern Ontario. *Journal of wind engineering and industrial aerodynamics*, 96(8-9), 1376-1389.
- Cheng, V. Y., Arhonditsis, G. B., Sills, D. M., Auld, H., Shephard, M. W., Gough, W. A., & Klaassen, J. (2013). Probability of tornado occurrence across Canada. *Journal of Climate*, 26(23), 9415-9428.
- Cheng, V. Y., Arhonditsis, G. B., Sills, D. M., Gough, W. A., & Auld, H. (2016). Predicting the climatology of tornado occurrences in North America with a Bayesian hierarchical modelling framework. *Journal of Climate*, 29(5), 1899-1917.
- Dunn, W. L., & Twisdale, L. A. (1979). A synthesized windfield model for tornado missile transport. *Nuclear Engineering and Design*, 52(1), 135-144.
- Etkin, D., Brun, S. E., Shabbar, A., & Joe, P. (2001). Tornado climatology of Canada revisited: tornado activity during different phases of ENSO. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 21(8), 915-938.

- Gillmeier, S., Sterling, M., Hemida, H., & Baker, C. J. (2018). A reflection on analytical tornado-like vortex flow field models. *Journal of Wind Engineering and Industrial Aerodynamics*, 174, 10-27.
- Hangan, H., & Kim, J. D. (2008). Swirl ratio effects on tornado vortices in relation to the Fujita scale. *Wind and Structures*, 11(4), 291-302.
- Natarajan, D., & Hangan, H. (2012). Large eddy simulations of translation and surface roughness effects on tornado-like vortices. *Journal of Wind Engineering and Industrial Aerodynamics*, 104, 577-584.
- Newark, M. J. (1991). A design basis tornado. *Canadian Journal of Civil Engineering*, 18(3), 521-524.
- Newark, M.J. 1984. Canadian tornadoes, 1950 -1979. *Atmosphere-Ocean*, 22 :343-353.
- NRC (2015) National Building Code of Canada. Institute for Research in Construction, National Research Council of Canada, Ottawa.
- Rankine, W. J. M. (1882). A manual of applied physics. Charles Griff and Co.: London, UK.
- Refan, M., & Hangan, H. (2016). Characterization of tornado-like flow fields in a new model scale wind testing chamber. *Journal of Wind Engineering and Industrial Aerodynamics*, 151, 107-121.
- Sills, D. M., Scriver, S. J., & King, P. W. S. (2004, October). The tornadoes in Ontario project (TOP). In *Preprints from the 22nd AMS Conference on Severe Local Storms, Hyannis, Mass, American Meteorological Society, CD-ROM B (Vol. 7)*.
- Tan, L. (2008). Assessment of tornado hazard maps for southern Ontario, M.E.Sc. Thesis, the University of Western Ontario, London, Ontario, Canada.
- Tan, L., & Hong, H. P. (2010). Influence of spatial inhomogeneity of tornado occurrence on estimated tornado hazard. *Canadian Journal of Civil Engineering*, 37(2), 279-289.

- Twisdale, L.A. (1978). Tornado data characterization and wind speed risk. *ASCE Journal of Structural Engineering*, 104(10), 1611- 1630.
- Twisdale, L.A., Dunn, W.L., Horie, Y., and Davis, T.L. (1981). Tornado Missile Simulation and design methodology. Vol.2, Model verification and database updates, Report prepared by Research Triangle Institute, Research Triangle Park, North Carolina, August.
- Wikle, C. K., & Anderson, C. J. (2003). Climatological analysis of tornado report counts using a hierarchical Bayesian spatiotemporal model. *Journal of Geophysical Research: Atmospheres*, 108(D24).
- Wurman, J., & Gill, S. (2000). Finescale radar observations of the Dimmitt, Texas (2 June 1995), tornado. *Monthly weather review*, 128(7), 2135-2164.

Chapter 2

2 Statistical assessment of spatial tornado occurrence for Canada: modelling and estimation

2.1 Introduction

Tornadoes can cause a significant amount of damage and fatalities. Statistics indicate that around 50 to 100 tornadoes per year are reported in Canada. However, the bias and underreporting of tornado events make the database to be incomplete, and the development of a predictive model for the tornado occurrence challenging. It seems that the first reported tornado in Canada, which occurred on 1 July 1792, leveled all houses in its path and uprooted trees along a narrow track (Etkin et al. 2001). A systematic Canadian tornado catalogue was established by Newark (1984), indicating that part of Canada is an extension of the great tornado-prone region centered over the south-central United States. Sills et al. (2004) updated this database for Ontario. The most recent released tornado catalogue by Environment and Climate Change Canada (ECCC) is available at <https://open.canada.ca/data/en/dataset/fd3355a7-ae34-4df7-b477-07306182db69> (last accessed January 21, 2020). It includes the tornadoes that occurred from 1980 to 2009. The inclusion of the tornado data after 1980 could be justified since the quality control of the information on the tornado events archived after the development of Canada's regional severe weather offices around 1980 (Etkin et al. 2001) is likely to be better.

The first tornado occurrence rate modelling for Canada was presented by Newark (1984) based on a tornado catalogue from 1950 to 1979. The obtained rate was used to identify regions in Canada that are prone to the tornado hazard. Southern Ontario is one of the regions. Historical tornado events indicate that the tornado occurrence in the region is not spatially homogenous (King et al. 2003) because of the lake breeze effect resulting in a preferred southwest-to-northeast pattern of tornadic activity. A statistical assessment of the tornado occurrence for the region was presented in Banik et al. (2007) and Tan and Hong (2010). The inhomogeneous tornado occurrence rate was estimated by applying the Gaussian kernel estimation technique (Silverman 1986) in Tan and Hong (2010). They further assumed that the tornado occurrence follows a Poisson process to map tornado

induced wind velocity hazard. For the same region, Tan (2008) adopted the zero-inflated Poisson (ZIP) process with a hierarchical Bayesian modelling technique advocated by Wikle and Anderson (2003) to model the tornado occurrence and applied the Markov Chain Monte Carlo (MCMC) method (Gilks et al. 1996; Brooks et al. 2011) to estimate the model parameters to define the tornado occurrence rate. It was observed that the estimated tornado occurrence rate based on the hierarchical Bayesian model is sensitive to the considered catalogue and the discretization scheme of the spatial region (i.e., the size of the cells used to cover the region).

Studies on the tornado occurrence rate for Canada were presented by Cheng et al. (2013, 2016) based on a Bayesian hierarchical modelling framework. For their analysis, the tornado occurrence is represented by using the Poisson process or ZIP process. Most interestingly, they considered the reporting bias due to population density and used the annual average cloud-to-ground lightning flash (ACGLF) density to estimate the tornado occurrence rate. The consideration of ACGLF density as the covariate is justified since the analysis presented in Reap and MacGorman (1989) indicated that the flash density is related to severe local storms and that the flash density and the tornado occurrence rate are positively correlated. The relation between the cloud-to-ground (CG) lightning flash and the tornado occurrence and tornado characteristics were also investigated in several studies (Branick and Doswell 1992; Knapp 1994; MacGorman and Burgess 1994; Perez et al. 1997; Carey and Rutledge 2003; Carey and Buffalo 2007). The focuses of these studies include the identification of a potential correlation between CG lightning pattern and tornado genesis, the assessment of the regional variability of the lightning polarity in severe storms, and the investigation of the potential benefit of using lightning strike data for forecast and warning operations. Since the lightning polarity in severe storms varies from region to region, Cheng et al. (2013) adopted ACGLF density as the covariate for predicting the tornado occurrence rate. This is because ACGLF density is an ideal surrogate variable for thunderstorms (Huffines and Orville 1999), although there are not perfectly correlated.

Beside of using ACGLF density as the covariate, the use of annual average thunderstorm days (ATD) as covariate could aid the estimation of the tornado occurrence rate. For example, Yarbrough and Meentemeyer (1978) showed the correlation between

thunderstorm days with tornado occurrence in the eastern United States. Bissolli et al. (2007) correlated the tornado occurrence to ATD in the north-western part of Germany. However, it seems that the use of ATD or ATD in combination with ACGLF density as the covariates to develop a predicting equation for the tornado occurrence rate is not available in the literature.

In addition to the use of the Poisson process and ZIP process to model the tornado occurrence (Wen and Chu 1973; Twisdale et al. 1983; Wikle and Anderson 2003; Anderson et al. 2007; Banik et al. 2007; Tan and Hong, 2010; Tippett et al. 2012; Cheng et al. 2013, 2016), the negative binomial (NB) model was also employed to model tornado occurrence (Elsner and Widen 2014; Jagger et al. 2015). The use of the ZIP and NB models is preferable to the Poisson model for cases where the mean of the occurrence is less than the variability of the occurrence (i.e., the overdispersion of the occurrence is observed). However, a study on whether the ZIP model or NB model is the preferred model for the tornado occurrence in Canada has not been investigated.

The main objective of the present study is to develop a predicting model to map the spatially varying tornado occurrence for Canada. To achieve this objective, first, we review the formulation based on the Poisson, ZIP, and NB models used to model the tornado occurrence, by considering the effects of underreporting due to population bias and using possible covariates. We also provide likelihood functions based on these models that can be used to estimate the model parameters. By considering the reported historical tornado catalogue, the reporting bias associate with the spatially varying population density, ACGLF density, and ATD, the estimation of the model parameters is carried out by using the maximum likelihood method (MLM) and Bayesian inference. Also, a preferred model is identified based on AIC and BIC by considering the mentioned stochastic models and by using the ACGLF density alone, ATD alone or both the ACGLF density and ATD as covariates. The mapped tornado occurrence rate for Canada by using a preferred predicting model is suggested.

The remainder of this paper is structured as follows. Section 2 presents essential information on the Poisson, ZIP, and NB models that are used to model the tornado

occurrence. Formulations for estimating the model parameters based on MLM and Bayesian inference are also presented. Section 3 elaborates on the historical tornado catalogue, population density, ACGLF density, and ATD that are used to develop the prediction equation for the tornado occurrence model. The use of the models summarized in Section 2 and the data shown in Section 3 to estimate the spatially varying tornado occurrence rate is presented in Section 4. Based on the results of the statistical analysis and model parameter estimation, a preferred predictive model for the tornado occurrence rate is given, and a tornado occurrence rate map for Canada is recommended. Finally, conclusions are summarized in Section 5.

2.2 Spatial tornado occurrence modelling and model parameter estimation

2.2.1 Modelling the tornado occurrence as a spatial point process

A distinction is made between the number of reported and actually occurred tornadoes because of the possible underreporting. We first divide a geographical region of interest into mutually exclusive grid cells. Let Y_{tj} denote the reported number of tornado occurrence before the time t at the j -th site (or cell), which could be modelled as

$$Y_{tj} = \sum_{i=1}^{N_{tj}} K_{tji}, \quad (2.1)$$

where N_{tj} is the actual number of the tornado occurrence, and K_{tji} is a Bernoulli random variable defined by the probability mass function $P(K_{tji} = 1) = p_{tj}$ and $P(K_{tji} = 0) = 1 - p_{tj}$, in which p_{tj} is the probability that the i -th tornado occurrence at the j -th site is reported (i.e., probability of reporting or detection).

The tornado occurrence in time and space is uncertain. As mentioned in the introduction, the commonly used probabilistic models include the Poisson process, ZIP process, and NB model.

If the (temporarily) stationary Poisson model is adopted for the tornado occurrence, N_{tj} for a time interval t is distributed according to,

$$N_{ij} \sim \text{Poisson}(\lambda_j t), \quad (2.2)$$

where λ_j is the time-independent occurrence rate. By considering the possible underreporting, Y_{ij} defined in Eq. (2.1) is modeled using,

$$Y_{ij} \sim \text{Binomial}(N_{ij}, p_{ij}), \quad (2.3)$$

It is considered that the prediction equation for the occurrence rate for Y_{ij} , η_j , could be expressed as,

$$\ln \eta_j = \ln p_{ij} + \ln \lambda_j = -\beta / \gamma_j + \alpha_{0j} + \sum_{k=1}^{n_k} \alpha_{kj} z_{kj} + \varepsilon_j, \quad (2.4)$$

where $-\beta / \gamma_j$ represents $\ln p_{ij}$ (Wikle and Anderson 2003; Anderson et al. 2007); the remaining terms in the last equality are equal to $\ln \lambda_j$ (Cheng et al. 2013); γ_j denotes the population density for the j -th cell, β and α_{kj} for $k = 0, \dots, n_k$, are model parameters, and n_k represents the number of explanatory variables z_{kj} . The error term ε_j in Eq. (2.4) is modelled using the conditional autoregressive model (CAR) (Besag 1974; Besag and Kooperber 1995),

$$\varepsilon_j | C \square N(\zeta_j, 1 / (\delta^2 n_{A_j})), \quad (2.5)$$

in which $\varepsilon_j | C$ denotes ε_j conditioned on its adjacent cells, $N(\cdot, \cdot)$ denotes the normal distribution with the first and second arguments representing the mean (vector) and variance (covariance matrix), n_{A_j} is the number of adjacent cells of the j -th cell, and $\zeta_j = (1 / n_{A_j}) \sum_k \varepsilon_k$ with k denotes all the adjacent cells of the j -th cell, δ^2 is a model

parameter. The adopted parameterization by using δ^2 instead of $\sigma^2 = 1 / \delta^2$ and by using $p_{ij} = \exp(-\beta / \gamma_j)$ instead of $p_{ij} = \exp(-\beta / \exp(\gamma_j))$ (Cheng et al. 2013) is due to the numerical convergence problem encountered in Bayesian analysis that will be discussed in the following section (see appendix A). The construction of the joint normal distribution function of $[\varepsilon_j]$, denoted as $N(0, \Omega)$ with Ω representing the covariance matrix of $[\varepsilon_j]$, can be done based on the formulation given in Besag (1974) and Besag and Kooperber (1995). Note that Cheng et al. (2013, 2016) used the ACGLF density alone in Eq. (2.4) as a covariate. The use of both the ACGLF density and ATD as the covariates to estimate the tornado occurrence rate has not been reported in the literature.

The full structure of the Bayesian hierarchical model is put in the appendix. We remark that by ingoring the error term our predictive model could be embedded in a generalized linear model (GLM) framework. The GLM should not be confused with the general linear model or generalized least squares. In GLM, the typical approach to estimate the parameters (here are the coefficients in Eq. (2.4)) is the maximum likelihood method, which is to be discussed in detail in the next section.

If the ZIP model is considered, the distribution of N_{ij} can be written as (Lambert 1992; Wilke and Anderson 2003),

$$N_{ij} \square \begin{cases} \text{Poisson}(\lambda_j t) & \text{with probability } 1-\pi_{ij} \\ 0 & \text{with probability } \pi_{ij} \end{cases}, \quad (2.6)$$

where π_{ij} denotes the additional probability that the tornado counts is equal to zero. This model effectively inflates the probability of non-occurrence. The variance of the number of occurrences obtained from the ZIP model is higher than that obtained from the Poisson model; it is adequate to model the occurrence with overdispersion. By considering that the reported tornado occurrence follows the model discussed earlier, and that (Wikle and Anderson 2003)

$$\pi_{ij} = \exp(\xi) / (1 + \exp(\xi)), \quad (2.7)$$

the actual tornado occurrence rate $\ln(\lambda_j)$ shown in Eq. (2.4) is replaced by,

$$\ln \lambda_j = -\ln(1 + \exp(\xi)) + \alpha_{0j} + \sum_{k=1}^{n_K} \alpha_{kj} z_{jk} + \varepsilon_j. \quad (2.8)$$

An alternative to the Poisson and ZIP models is to use the NB model with the number of events N_{ij} distributed according to,

$$P(N_{ij} = n) = \frac{\Gamma(n + rt)}{n! \Gamma(rt)} p_{NBj}^r (1 - p_{NBj})^n, \quad (2.9)$$

where r is a real positive number, and p_{NB} is the probability of non-occurrence. For this distribution, the mean, μ_j , equals $rt(1 - p_{NBj}) / p_{NBj}$, and the variance equals $rt(1 - p_{NBj}) / p_{NBj}^2$. A simple algebraic manipulation shows that the variance can also be expressed as $\mu + \mu^2 / r$, indicating that the use of the NB model can cope with overdispersion. The overdispersion increase as r decreases. By adopting the model shown

in Eq. (2.1) which takes into account the reporting bias due to population density, it can be shown that Y_{ij} is still NB distributed but with p_{NBj} replaced by p_{NBBj} where

$$p_{NBBj} = \frac{p_{NBj}}{p_{ij} + p_{NBj}(1 - p_{ij})}. \quad (2.10)$$

Since the occurrence rate of Y_{ij} based on this model equals to $r(1 - p_{NBBj})/p_{NBBj}$, by equating this rate to η_j shown in Eq. (2.4) we obtained,

$$p_{NBBj} = \frac{r}{r + \eta_j}, \quad (2.11)$$

Although the considered reported tornado occurrence rate η_j by using the Poisson model and the NB model has the same functional form, the estimation of the model parameters by using the MLM depends on which model is considered. The likelihood functions used to estimate model parameters are given in the following section for the three considered models.

2.2.2 Parameter estimation

We first consider the Poisson model. In this case, the likelihood function L based on the model specified in Eqs. (2.1) to (2.4) for Y_{ij} can be written as

$$L = \prod_{j=1}^{M_c} \frac{(\eta_j(\Theta)t)^{y_j}}{y_j!} e^{-\eta_j(\Theta)t}, \quad (2.12)$$

where the notation $\eta_j(\Theta)$ is used to replace η_j to emphasize that η_j is a function of a set of model parameters $\Theta = (\bar{\alpha}, \beta)$, the vector $\bar{\alpha}$ denotes all α_{kj} , M_c is the total number of cells considered, and y_j is the reported tornado occurrences for the j -th cell. In formulating the likelihood function, the error term ε_j in Eq. (2.4) is neglected. The model parameters α_{kj} and β can then be estimated by maximizing L .

Similarly, if the ZIP model is considered the likelihood function L is given by

$$L = \left[\frac{1}{1 + \exp(\xi)} \right]^{M_c} \prod_{j=1}^{M_c} \left[\exp(\xi) I_{\{y_j=0\}} + \frac{(\eta_j(\Theta)t)^{y_j}}{y_j!} e^{-\eta_j(\Theta)t} \right], \quad (2.13)$$

where $I_{\{y_j=0\}}(y_j)$ equals 1 and 0 for y_j equal to zero and not equal to zero, respectively.

By considering the NB model and the model parameters shown in Eqs. (2.4) and (2.11), it can be shown that L is given by

$$L = \prod_{j=1}^{M_c} \left[\frac{\Gamma(y_j + rt)}{y_j! \Gamma(rt)} \left(\frac{r}{r + \eta_j(\Theta)} \right)^{rt} \left(\frac{\eta_j(\Theta)}{r + \eta_j(\Theta)} \right)^{y_j} \right]. \quad (2.14)$$

Again, the model parameters can be estimated by maximizing L shown in Eq. (2.13) for the ZIP model and in Eq. (2.14) for the NB model.

Instead of using the simple MLM to estimate the model parameters, a more sophisticated approach to estimate model parameters is based on Bayesian inference. This approach was employed in Wikle and Anderson (2003) to estimate the tornado occurrence rate for sites in the United States. It was also employed in Tan (2008) for southern Ontario and in Cheng et al. (2013, 2016) for Canada and the United States. It assesses the posterior probability distribution of the model parameters of tornado occurrence rate and updates the stochastic tornado occurrence process based on historical tornado catalogue through a series of conditional probabilistic models. The posterior distribution of the model parameters, [parameters of process, parameters | data], could be expressed as,

$$\begin{aligned} & [\text{parameters of the process, parameters} \mid \text{data}] \propto \\ & [\text{data} \mid \text{parameters of process, parameters}] [\text{parameters of process} \mid \text{parameters}] [\text{parameters}] \end{aligned} \quad (2.15)$$

where the symbol \propto denotes the proportionality, the brackets refer to the probability distribution, the vertical line separates the conditioned uncertain variable(s) and conditioning variables, [data | parameters of process, parameters] represents the data model, [parameter of process | parameters] represents the process model, and [parameters] denotes the parameter model. Since the posterior probability distribution is difficult to obtain analytically, the Markov Chain Monte Carlo (MCMC) sampling approach (Gilks et al. 1996; Brooks et al. 2011) is employed. The point estimate of the model parameters can be obtained based on the posterior distribution.

2.3 Tornado catalogue and data used for estimating the tornado occurrence rate

Newark (1984) established the systematic Canadian tornado catalogue for events that occurred from 1950 to 1979. The data was gathered from a variety of sources, including newspapers, local community histories, field surveys of tornado damage, Atmospheric Environment Service (AES) (which becomes Environment and Climate Change Canada (ECCC)) archive material, AES weather observations, data collected by provincial organizations, public archive material, and human memory. Although the catalogue is incomplete because of underreporting and only covers the areas with a population density greater than a threshold value, its use indicated that southwestern Ontario and southeastern Manitoba have the maximum tornado activities within Canada. A tornado catalogue from 1980 to 2009 was released by ECCC (see <https://open.canada.ca/data/en/dataset/fd3355a7-ae34-4df7-b477-07306182db69> (last accessed January 21, 2020)). No additional tornado catalogue was released by a government agency in Canada. In this catalogue, there are 1839 tornado events that occurred from 1980 to 2009 in Canada. Cheng et al. (2013) used this catalogue for their tornado occurrence assessment. The spatial distribution of the tornadoes included in the catalogue is shown in Figure 1, confirming the substantial tornado activities that occurred in southwestern Ontario and southeastern Manitoba.

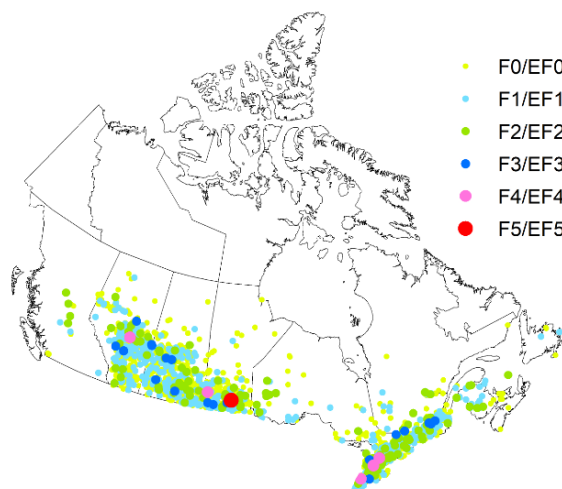


Figure 2.1. Spatial distribution of reported historical tornado events from 1980 to 2009.

Since the tornado reporting bias is considered to be a function of the population density (see Eq. (2.4)), a population density is estimated by using 2001 Canadian census data (<https://www12.statcan.gc.ca/english/census01/products/standard/popdwell/tables.cfm>).

The spatially varying population density according to census data for each **census subdivision** is depicted in Figure 2a. The cells defined by counties are very irregular and not sufficiently refined. If the squared grid cell system, with each cell covering approximately $30 \times 30 \text{ km}^2$ presented in Figure 2b, is considered, the obtained population density for each cell is also plotted in the same figure. This squared grid system, with the size of the cell that is within that considered in Cheng et al. (2013), is used throughout this study for predicting the tornado occurrence rate. The plot shown in Figure 2b indicates that the Canadian urban area is mostly concentrated near its southern border with the United State

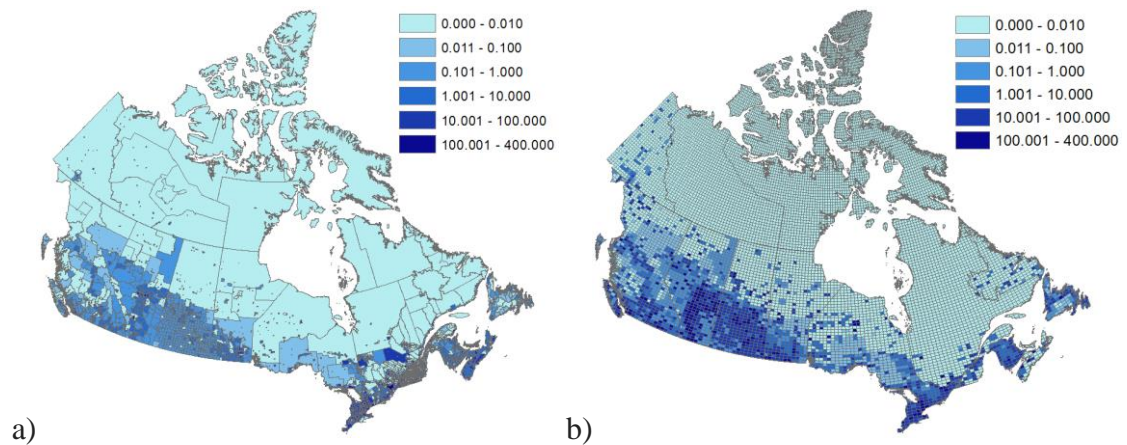


Figure 2.2. Spatially varying population density (number of people per km^2) based on 2001 Canadian census data: a) based on **census subdivisions**; b) based on regular grid cells.

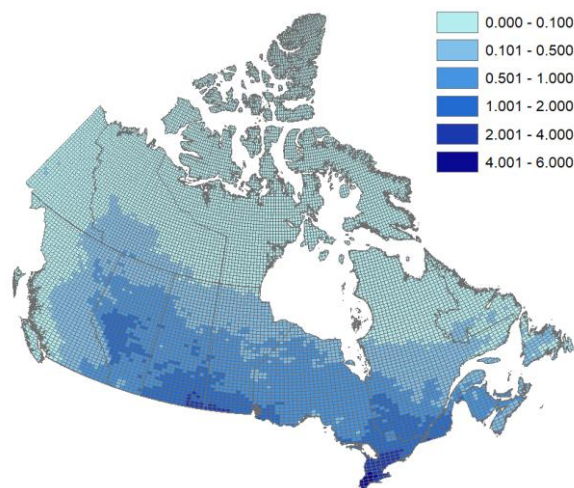


Figure 2.3. Plot of the ACGLF density (per year per km²) using data from 1999 to 2009.

To explore possible covariates that can be used in Eq. (2.4) for predicting the tornado occurrence rate, the CG lightning data from 1999 to 2009 are obtained from the Canadian Lightning Detection Network. Details on the CG lightning data are described in Burrows and Kochtubajda (2010) and Shephard et al. (2013). The data are processed, and the obtained ACGLF density is shown in Figure 3. A visual inspection of the reported historical tornadoes shown in Figure 1 and the ACGLF density indicates that there are similarities in the spatial trends in the occurrence of the tornadoes and the ACGLF density, especially for southwestern Ontario and the southern border of Saskatchewan and Manitoba.

To identify the thunderstorm days, we use the wind records from ECCC DLY04 digital archive (see http://www.climate.weatheroffice.gc.ca/prods_servs/documentation_index_e.html#hly01).

In this database, the daily maximum wind speed was reported for each meteorological station, and an identifier indicating whether the wind is thunderstorm wind or non-thunderstorm wind is provided. The locations of the stations are identified in Figure 4a, where each of the stations has at least ten years of useable data from 1989 to 2009. The considered recording period is consistent with the period of the tornado catalogue used and shown in Figure 1, and the consideration of at least ten years of useable data is aimed at minimizing the statistical uncertainty in estimating ATD. The number of thunderstorm days is counted by using the thunderstorm identifier, and the ATD at each station is then

estimated. Spatial smoothing of ATD obtained for each of the stations shown in Figure 4a is carried out using the ordinary kriging (Johnston et al. 2003). The smoothed spatially varying ATD is presented in Figure 4b. An inspection of the reported historical tornadoes plotted in Figure 1, and the ATD shown in Figure 4b indicates that there are similarities in their spatial patterns, suggesting again that the ATD could be used as a covariate to predict the tornado occurrence rate.

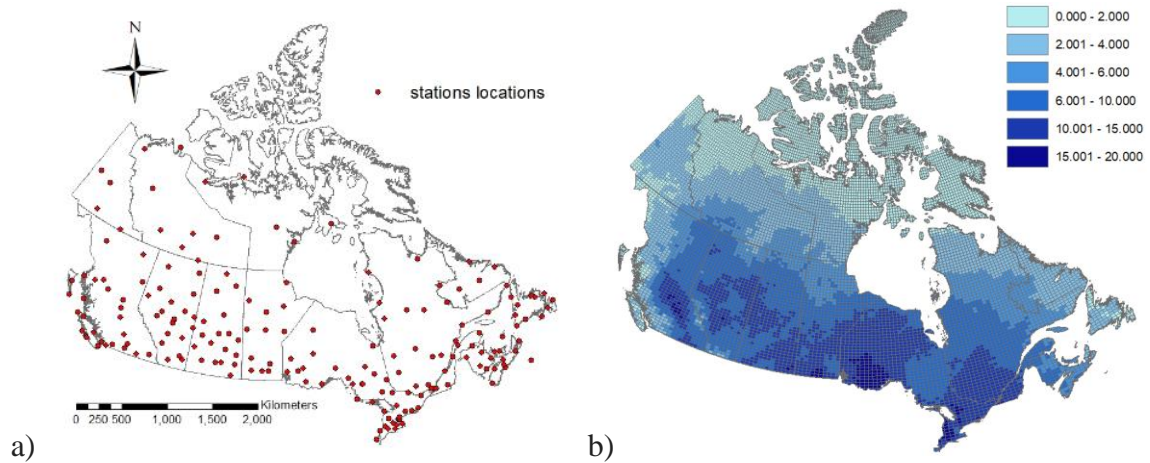


Figure 2.4. a) Locations of the stations with thunderstorm winds recorded from 1980 to 2009; b) the estimated ATD per cell.

2.4 Tornado occurrence rate estimation for Canada

2.4.1 Maximum likelihood estimates

If the historical catalogue contains a large amount of data, the model coefficients α_{kj} for the model shown in Eqs. (2.4) and (2.8) could be estimated by using the MLM and Bayesian method. However, since the historical tornado catalogue is short, it does not allow the estimation of the cell-dependent (i.e., spatially varying) model parameter α_{kj} with confidence. Therefore, for the numerical analysis presented in the following, it is assumed that $\alpha_{kj} = \alpha_k$. This assumption is consistent with that used in Cheng et al. (2013).

By considering the Poisson model and using the spatially varying ACGLF density alone as the covariate in Eq. (2.4) (i.e., $n_k = 1$ and z_{1j} represents the ACGLF density), the estimated

model parameters by maximizing L shown in Eq. (2.12) are obtained and presented in Table 1, where the population density γ_j in terms of person/km² and the ACFLF density in terms of per (year×km²) are used for the numerical analysis. In the same table, the estimated standard deviations of the estimated model parameters are also presented, indicating that they are relatively small. For the numerical analysis, the maximization of L is carried out by using the function “glm” available in **R** programming language for statistical computing (R Core Team 2016). Also, the Akaike information criterion (AIC) and Bayesian information criterion (BIC) of the considered model (i.e., AIC(model) and BIC(model)) (Burnham and Anderson 2002),

$$\text{AIC}(\text{model}) = 2k - 2\ln L \quad (2.16)$$

and

$$\text{BIC}(\text{model}) = k \ln(n) - 2\ln L \quad (2.17)$$

are calculated, where k is the total number of model parameters to be estimated, n is the sample size, and L is the maximized value of L . The calculated AIC(model) and BIC(model) are also included in Table 1. AIC and BIC could be treated as penalized-likelihood-based model selection tools if several models are considered. The model with the lowest AIC(model) or BIC(model) is the preferred model.

The estimation of the model parameters for the Poisson model in **R** programming language (R Core Team 2016) is repeated by considering L shown in Eq. (2.13) for the ZIP model and in Eq. (2.14) for the NB model. The estimated model parameters, AIC(model), and BIC(model) are also presented in Table 1. A comparison of the values of AIC(model) or the values of BIC(model) indicates that the NB model is preferred among the three considered models. This preference is followed by the ZIP model. The Poisson model is the least preferred model. In the table, the estimated r for the NB model is small, suggesting that there is a large overdispersion of tornado occurrence (see Eq. (2.9)). A comparison of the predicted tornado occurrence rates based on the models listed in Table 2.1 is shown in Figure 5. The plot indicates that the use of the NB model leads to a less spatially smeared tornado occurrence rate than that obtained by using the Poisson model or the ZIP model.

Table 2.1. Estimated model parameters considering ACGLF density as the covariate (the first entry represents the estimated value, and the second entry represents the estimated standard deviation of the model parameter).

Model	Model parameters							
	r	α_0	α_1	ξ	β	AIC	BIC	Annual rate for Canada
Poisson	---	-1.097; 0.037	0.778; 0.016	---	0.036; 0.001	6912	6934	157
ZIP	---	0.032; 0.053	0.500; 0.022	0.566; 0.065	0.036; 0.001	6230	6260	155
NB	0.347; 0.021	-2.699; 0.061	2.399; 0.044	---	0.026; 0.001	5359	5382	152

In all cases, the maps presented in Figure 5 indicate that the tornado activities in southwestern Ontario and southeastern Manitoba are most severe in Canada. The spatial trends of the tornado activities follow those exhibited by using the historical catalogue. The annual occurrence rate for Canada is calculated by integrating the occurrence rate over Canada shown in the plots in Figure 5. The obtained annual rate for Canada for each considered model is presented in Table 1. The ratio of the predicted rate to that calculated from the reported tornadoes (i.e., 1839/30) is about 2.5. This ratio is consistent with the value of 2.3 reported by Cheng et al. (2013) based on their model and using 25×25 km² grid cells. A discussion of the large ratio was justified in Cheng et al. (2013), citing the sparse population density and lack of radar coverage in some regions. Note that the adjustment of tornado occurrence is small since the probability of reporting of an actually occurred tornado p_{ij} (i.e., $p_{ij} = \exp(-\beta/\gamma_j)$) for $\beta = 0.036$ is greater than 90%, 99%, 99.4%, and 99.5% for the average population density γ_j greater than 0.35, 3.6, 6, and 7.2 (person/km²), respectively. p_{ij} is a decreasing function of β . These values agree with the observation (King 1997) that if the threshold of rural population density is about 6 (person/km²) then most tornadoes should be observed. The calculated p_{ij} also indicates that the underreporting is substantial only for regions with very low population densities.

In other words, the large ratio is partly ascribed to the underreporting in regions with very low population densities in Canada.

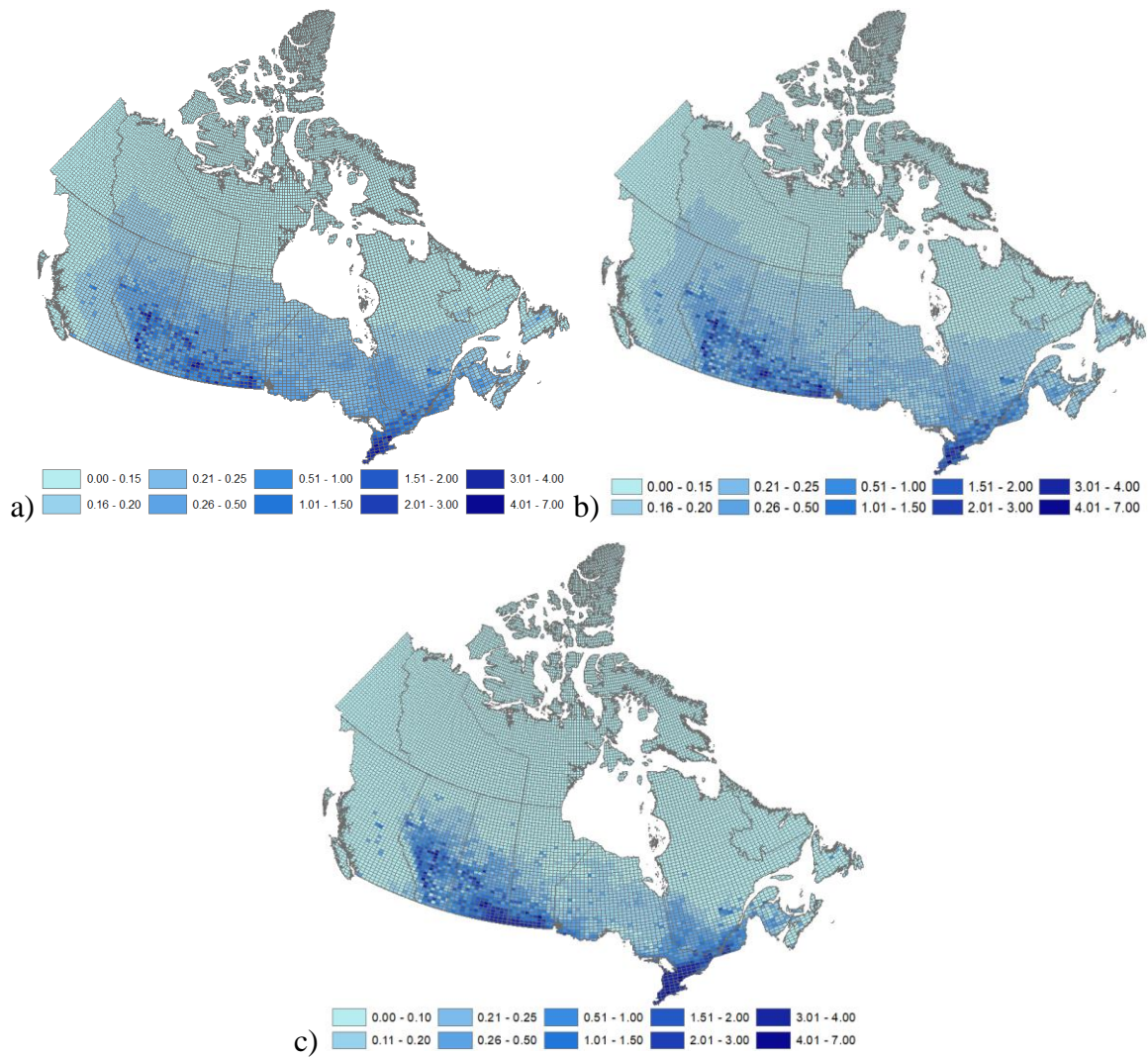


Figure 2.5 Predicted tornado occurrence rate (per year per 10^4 km^2) by considering ACGLF density as the covariate and based on: a) Poisson model, b) ZIP model and c) NB model.

Rather than considering the ACGLF density as the covariate in Eq. (2.4), if the spatially varying ATD is considered as the covariate (i.e., in Eq. (2.4) $n_k = 2$ but $\alpha_1 = 0$ and z_{2j} represents the ATD density), by repeating the analysis that is carried out for the results presented in Table 1, the estimated model parameters, in this case, are summarized in Table 2.2. An inspection of the results presented in the table indicates that the preferred model

judged based on AIC(model) or BIC(model) is again the NB model, which is consistent with that observed from the results shown in Table 2.1. A comparison of AIC and BIC values presented Tables 2.1 and 2.2 indicates that the model with ACGLF density alone as the covariate is preferable to the model with ATD alone as the covariate. To further compare the NB model with ATD as the covariate to that with the ACGLF density as the covariate, the predicted tornado occurrence rate by using the former is presented in Figure 6a. An inspection of the mapped occurrence rates shown in Figures 5c and 6a indicates that the use of the ATD as the covariate results in a more smeared occurrence rate and increased tornado activities in regions with higher latitudes as compared to that obtained by using the ACGLF density as the covariate.

Table 2.2. Estimated model parameters considering ATD as the covariate (the first entry represents the estimated value, and the second entry represents the estimated standard deviation of the model parameter).

Model	Model parameters							
	r	α_0	α_2	ξ	β	AIC	BIC	Annual rate for Canada
Poisson	---	-2.283; 0.080	0.201; 0.066	---	0.033; 0.001	7324	7346	165
ZIP	---	-1.319; 0.103	0.187; 0.008	0.479; 0.062	0.030; 0.001	6174	6203	160
NB	0.275; 0.017	-3.407; 0.120	0.313; 0.011	---	0.028; 0.001	5659	5681	156

Now, consider that the models with both the ACGLF density and ATD as the covariates. The use of the MLM results in the estimated model parameters presented in Table 2.3. The preferred model is the NB model, as well. The ratio of the predicted annual occurrence rate for Canada by using this preferred model to that calculated from the reported tornadoes (i.e., 1839/30) is about 2.38.

The results presented in Tables 2.1 to 2.3 indicate that judged based on AIC and BIC, the preferred model is the NB model with the ACGLF density and ATD as the covariates. This

observation suggests that the use of two covariates is preferable than using the ACGLF density alone or ATD alone as the covariate. This observation is equally applicable if the Poisson model or the ZIP model is considered as well. An additional observation is that in all cases, the standard deviations of the estimated model parameters are relatively small, providing confidence on the point estimate of model parameters.

Using the NB model shown in Table 2.3, the predicted tornado occurrence rate is presented in Figure 6b. A comparison of the predicted occurrence rates shown in Figures 5c, 6a, and 6b indicate that they all follow the similar spatial trends. However, the extent of the area with the annual occurrence rate higher than 0.15 (per 10^4 km²) for Figure 6b is between those observed from Figures 5c and 6a.

Table 2.3. Estimated model parameters considering the ACGLF density and ATD as the covariates (the first entry represents the estimated value, and the second entry represents the estimated standard deviation of the model parameter).

Model	Model parameters								
	r	α_0	α_1	α_2	ξ	β	AIC	BIC	Annual rate for Canada
Poisson	---	-2.066; 0.080	0.606; 0.021	0.112; 0.008	---	0.031; 0.001	6689	6718	156
ZIP	---	-1.204; 0.101	0.328; 0.027	0.130; 0.009	0.355; 0.069	0.030; 0.001	6043	6080	153
NB	0.371; 0.023	-3.201; 0.102	1.834; 0.051	0.103; 0.011	---	0.025; 0.001	5326	5355	146

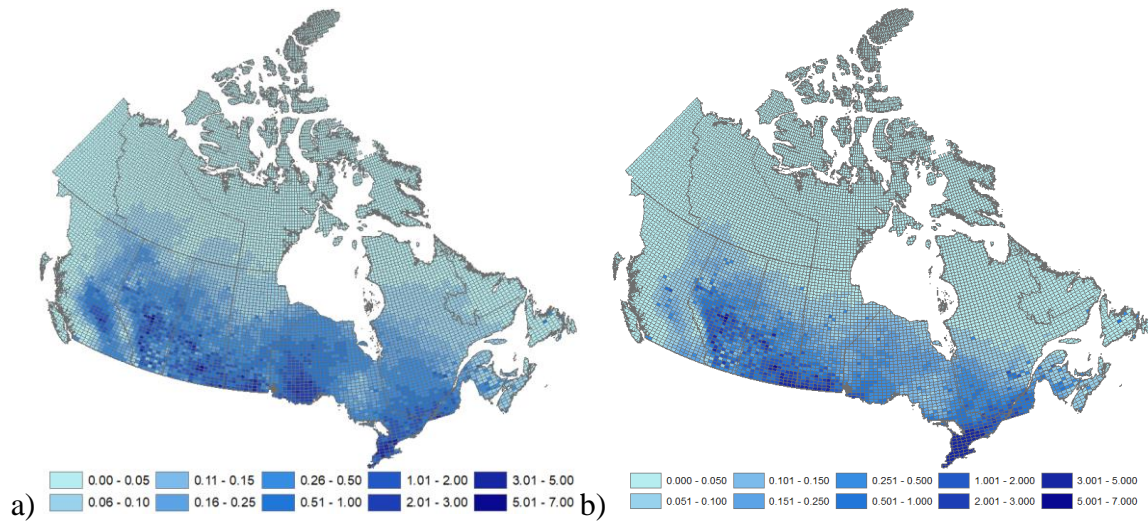


Figure 2.6. Predicted tornado occurrence rate (per year per 10^4 km^2) by using the NB model a) using ATD as the covariate, and b) using both the ACGLF density and ATD as the covariates.

2.4.2 Bayesian inference

The estimation based on Eq. (2.15) requires the evaluation of the posterior distribution given the data and prior distributions of the model parameters. For the numerical analysis to be carried out, the MCMC sampling approach is applied (Gilks et al. 1996; Brooks et al. 2011). In the present study, WinBUGS software (<http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml>) is used to implement the Metropolis-Hasting (MH) algorithm to evaluate Eq. (2.15). WinBUGS provides a graphical user interface and on-line monitoring and convergence diagnostics. The numerical analysis requires the assignment of the prior probability distributions as well as the initial values of the model parameters. The approach of monitoring convergence of MCMC output proposed by Gelman and Rubin (1992) is used for the analysis presented below. Moreover, a thinning by taking every 5th sample is adopted to reduce the possible correlation of the samples. Although analysis by considering the NB model with the ACGLF density alone, ATD alone, and both as the covariates is carried out, only the results by using both the ACGLF density and ATD as the covariates are presented. This is because the conclusions drawn from results obtained for other models are similar. The assigned “non-informative” prior for the analysis are: $r \sim \text{Gamma}(1, 0.01)$, $\alpha_0 \sim N(0, 10^4)$, $\alpha_1 \sim N(0, 10^4)$, $\alpha_2 \sim N(0, 10^4)$,

$\beta \sim N(0,10^4)$, and $\delta^2 \sim \text{Gamma}(1,0.01)$, where $\text{Gamma}(\cdot)$ denote the gamma distribution with the first and second arguments denote the shape and rate parameters. For simplicity, this is referred to as Case 1.

To illustrate the convergence of the MCMC, a plot of results of the Brooks-Gelman-Rubin (BGR) diagnostics is presented in Figure 7. As can be observed from the figure, the overall and within-chain variabilities become stable as the iteration increases. Moreover, the BGR diagnostic tends to become stable around one, suggesting that the convergence is achieved, and the generated samples are valid and can be used to represent the marginal posterior distributions.

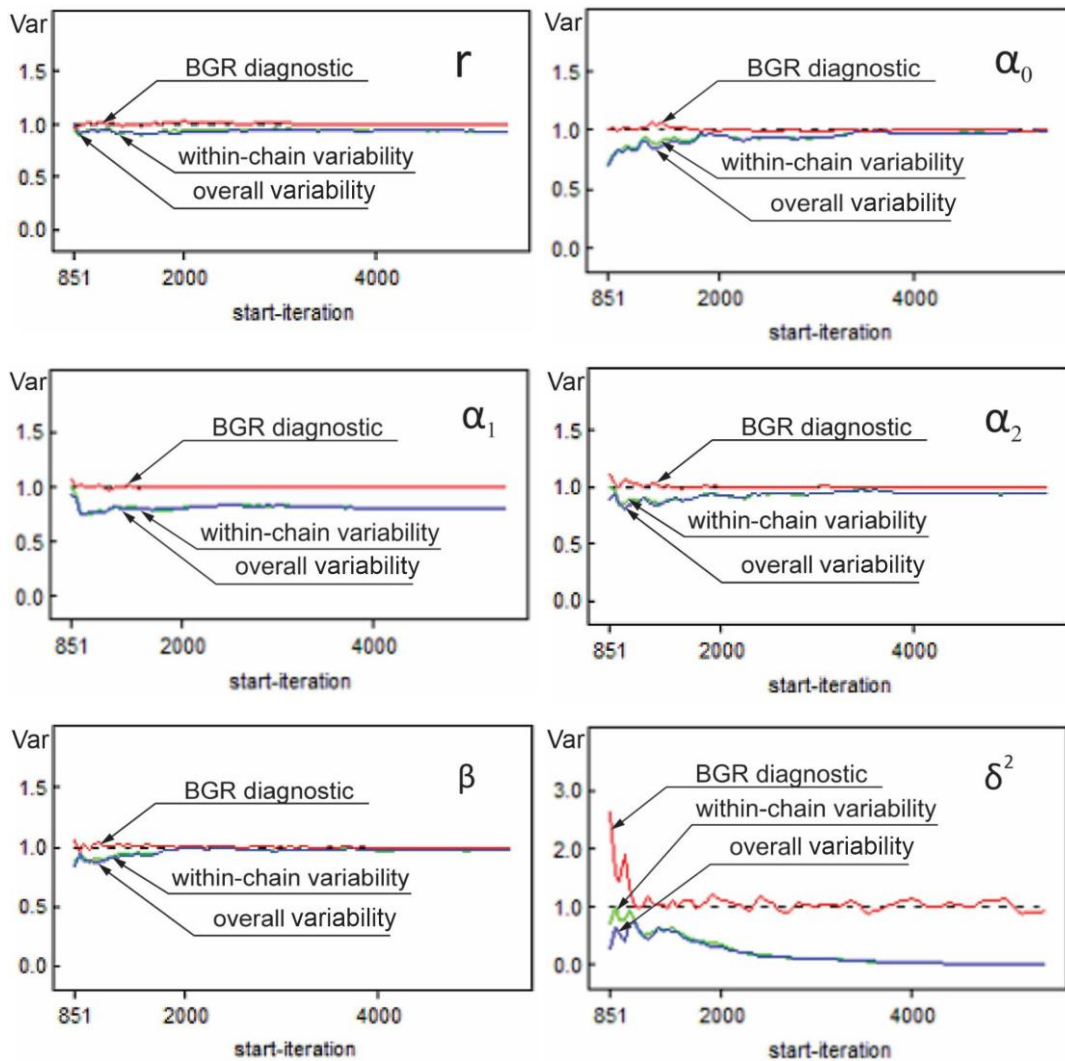
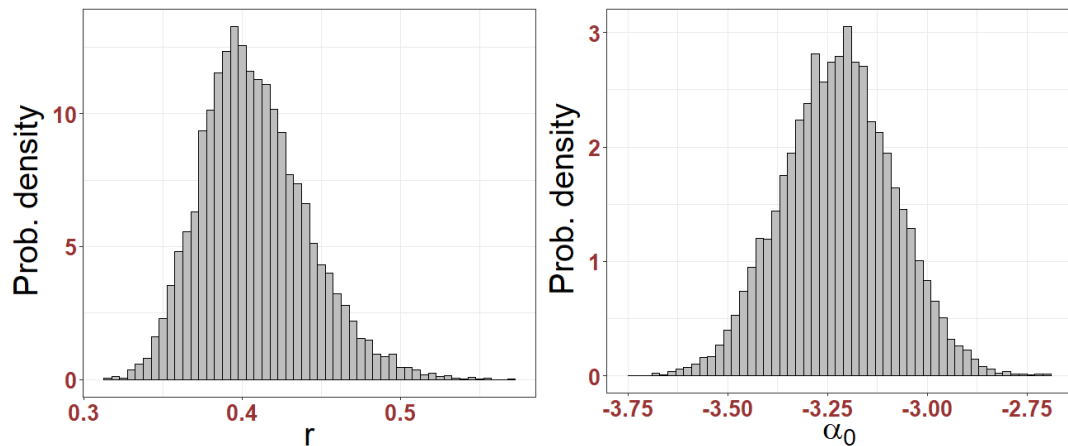


Figure 2.7. Results from the BGR diagnostics for r , α_0 , α_1 , α_2 , β and δ^2 considering the NB model with the ACGLF density and ATD as the covariates.

The marginal posterior distributions of r , α_0 , α_1 , α_2 , β and σ^2 based on more than 10000 samples are shown in Figure 8 for Case 1. The mean, median, and standard deviation of the parameters calculated from the marginal posterior distributions are presented in Table 2.4. As can be observed that since δ^2 is not very large, the standard deviation of the error term in the predicting equation for the tornado occurrence rate is not small. Table 4 shows that the mean and median are practically the same, indicating that the posterior marginal probability distributions of the model parameters are almost symmetric, which can be observed in Figure 8. A comparison of the mean and standard deviation shown in Tables 3 and 4 for the NB model indicates that they agree well except the standard deviation for α_2 . This good agreement could be explained based on the relation between the maximum likelihood estimate and the so-called maximum a posteriori probability estimate (Robert 2007; Lee 2012). The latter tends to the former if the prior distribution tends to uniform. In other words, the good agreement between the maximum likelihood estimates and the point estimate from the Bayesian inference is likely due to the “non-informative” prior.



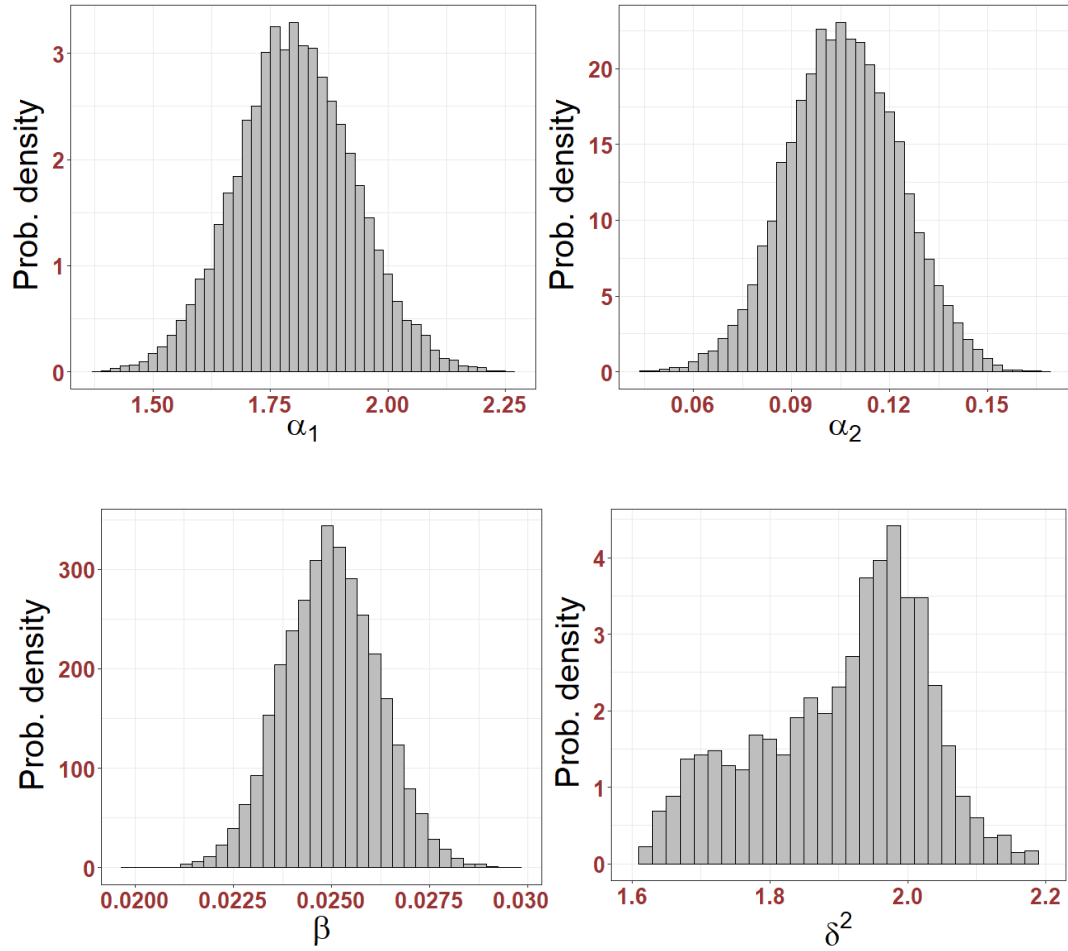


Figure 2.8. The marginal posterior distributions for r , α_0 , α_1 , α_2 , β and δ^2 by considering the NB model with the ACGLF density and ATD as the covariates.

It is seen from Figure 8 that the marginal posterior distributions for r , α_0 , α_1 , α_2 , and β are almost symmetric. The symmetry feature indicates that the posterior mode equals to the posterior mean or median, which echoes the result that the MLE estimators are close to the Bayesian estimators. The posterior distribution of r is positively skewed, and the posterior distribution of δ^2 is negatively skewed.

Table 2.4. Estimated statistics of the model parameters based on the marginal posterior distribution.

Parameter	Case	Model parameters					
		r	α_0	α_1	α_2	β	η^2
Mean	1	0.389	-3.198	1.824	0.103	0.025	1.938
	2	0.387	-3.170	1.841	0.101	0.025	2.008
	3	0.385	-3.189	1.838	0.103	0.025	1.972
Median	1	0.387	-3.198	1.824	0.103	0.025	1.953
	2	0.385	-3.171	1.839	0.100	0.025	2.037
	3	0.382	-3.187	1.836	0.101	0.025	1.976
Standard deviation	1	0.023	0.135	0.124	0.017	0.001	0.171
	2	0.022	0.132	0.128	0.016	0.001	0.165
	3	0.023	0.138	0.130	0.016	0.001	0.176

A few comments on using the model specified in Eq. (2.4) and the use of WinBUGS are now in order. First, instead of using $p_{ij} = \exp(-\beta/\gamma_j)$ in Eq. (2.4), $p_{ij} = \exp(-\beta/\exp(\gamma_j))$ used in Cheng et al. (2013) was considered. Unfortunately, the results from the BGR diagnostics indicate that the convergence cannot be achieved for such a model. Although Cheng et al. (2013) did not mention the diagnostics for convergence, and their model parameters differ from those given in the present study, the ratio of the predicted annual occurrence rate for Canada by using their model is consistent with that obtained in the present study as shown in the previous section. Second, the readily available probability distribution types in WinBUGS are limited. For example, one could not specify inverse gamma distribution using readily available distributions for $\sigma^2 = 1/\delta^2$. For this reason, we adopted the parameterization, as shown in the residual part in Eq. (2.5).

Since the posterior distribution of the model parameters in Bayesian inference is influenced by the selected prior distribution in many applications, a sensitivity analysis is carried out by considering two additional cases: Cases 2 and 3. For the sensitivity analysis, the prior distributions for α_0 , α_1 , and β are varied since they directly affect the estimated tornado occurrence rate by considering the model shown in Eq. (2.4). The assigned prior

distributions in Case 2 are $r \sim \text{Gamma}(1, 0.01)$, $\alpha_0 \sim N(0, 10^4)$, $\alpha_1 \sim \text{Gamma}(1, 0.01)$, $\alpha_2 \sim \text{Gamma}(1, 0.01)$, $\beta \sim \text{Gamma}(1, 0.01)$, and $\eta^2 \sim \text{Gamma}(1, 0.01)$. In Case 3, they are $r \sim \text{Gamma}(1, 0.01)$, $\alpha_0 \sim N(0, 10^4)$, $\alpha_1 \sim \text{TN}(0, 10^4)$, $\alpha_2 \sim \text{TN}(0, 10^4)$, $\beta \sim \text{TN}(0, 10^4)$, and $1/\sigma^2 \sim \text{Gamma}(1, 0.01)$, where $\text{TN}(\mu, \sigma^2)$ denotes the normal distribution $N(\mu, \sigma^2)$ but truncated at zero, which leads it to be a non-negative distribution.

Since the obtained posterior distributions for Cases 2 and 3 based on Bayesian inference are very similar to those presented in Figure 8, they are not plotted. The calculated statistics of the model parameters based on the samples of the posterior distributions are also presented in Table 4. The comparison of the results presented in the table indicates that the results for the considered application are not very sensitive to the assigned prior distributions of α_0 , α_1 , and β .

The results presented in this section indicate that the estimates of the model parameters needed to predict the tornado occurrence rate from the MLM are consistent with those from the Bayesian inference. It suggests that the use of MLM is adequate for the considered application. This is advantageous since the use of MLM is convenient, and its implementation is straightforward (see Eqs. (2.12) to (2.14)).

2.5 Summary and conclusions

The present study is focused on the development of a preferred predicting model for the tornado occurrence rate for Canada. Three stochastic models, namely the Poisson model, the zero-inflated Poisson (ZIP) model, and the negative binomial (NB) model, are considered. The development takes into account the underreporting due to population density and considers the annual average cloud-to-ground lightning flash (ACGLF) density alone, annual average thunderstorm days (ATD) alone, or both ACGLF density and ATD as the covariates. Estimation of the model parameters is carried out by using the maximum likelihood method and the Bayesian inference.

The analysis results indicate that the NB model for predicting the tornado occurrence rate is the preferred model for Canada and that the tornado occurrence is associated with large

overdispersion. According to the preferred model, the ratio of the estimated expected annual tornado occurrence for Canada to that of reported is about 2.38, which agrees with that suggested in the literature. This large ratio is partly ascribed to the underreporting in the regions with very low population densities in Canada. The results also suggest that the use of both ACGLF density and ATD as the covariates is preferred since these models are associated with the lowest AIC and BIC. The use of ATD alone as the covariate is worse than the use of ACGLF alone as the covariate. The suggested predicting model is the NB model with ACGLF density and ATD as the covariates (see the NB model in Table 3 and in Table 4).

In addition, statistical analysis carried out indicates that the predicting model for the tornado occurrence rate developed based on Bayesian inference is relatively insensitive to the assumed prior distributions and that the predicting model developed based on the MLM is practically the same as that based on Bayesian inference. This agreement is partly attributed to the use of the “non-informative” prior.

Data availability statement

Some or all data, models or code generated or used during the study are available from the corresponding author by request. These include the processed tornado catalogue used, the population density used, the estimated thunderstorm days, and estimated tornado occurrence rate for mapping.

References

- Anderson, C.J., Wikle, C.K., Zhou, Q., & Royle, J.A. (2007). Population influences on tornado reports in the United States. *Weather and forecasting*, 22(3), 571-579.
- Banik, S.S., Hong, H.P., & Kopp, G.A. (2007). Tornado hazard assessment for southern Ontario. *Canadian Journal of Civil Engineering*, 34(7): 830–842.

- Besag, J., & Kooperberg, C. (1995). On conditional and intrinsic autoregressions. *Biometrika*, 82(4), 733-746.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2), 192-225.
- Bissolli, P., Grieser, J., Dotzek, N., & Welsch, M. (2007). Tornadoes in Germany 1950–2003 and their relation to particular weather conditions. *Global and Planetary Change*, 57(1-2), 124-138.
- Branick, M. L., & Doswell III, C. A. (1992). An observation of the relationship between supercell structure and lightning ground-strike polarity. *Weather and forecasting*, 7(1), 143-149.
- Brooks, S., Gelman, A., Jones, G., & Meng, X. L. (Eds.). (2011). *Handbook of markov chain monte carlo*. CRC press.
- Burnham, K., & Anderson, D. (2002). *Model Selection and Multi-model Inference*. 2nd edn. Springer, New York.
- Burrows, W.R., & Kochtubajda, B. (2010). A decade of cloud-to-ground lightning in Canada: 1999–2008. Part 1: Flash density and occurrence. *Atmosphere-ocean*, 48(3), 177-194.
- Carey, L.D., & Rutledge, S.A. (2003). Characteristics of cloud-to-ground lightning in severe and non-severe storms over the central United States from 1989–1998. *Journal of Geophysical Research: Atmospheres*, 108(D15).
- Carey, L.D., & Buffalo, K.M. (2007). Environmental control of cloud-to-ground lightning polarity in severe storms. *Monthly weather review*, 135(4), 1327-1353.

- Cheng, V.Y., Arhonditsis, G.B., Sills, D.M., Auld, H., Shephard, M.W., Gough, W.A., & Klaassen, J. (2013). Probability of tornado occurrence across Canada. *Journal of Climate*, 26(23), 9415-9428.
- Cheng, V.Y., Arhonditsis, G.B., Sills, D.M., Gough, W.A., & Auld, H. (2016). Predicting the climatology of tornado occurrences in North America with a Bayesian hierarchical modelling framework. *Journal of Climate*, 29(5), 1899-1917.
- Elsner, J.B., & Widen, H.M. (2014). Predicting spring tornado activity in the central Great Plains by 1 March. *Monthly Weather Review*, 142(1), 259-267.
- Etkin, D., Brun, S.E., Shabbar, A., & Joe, P. (2001). Tornado climatology of Canada revisited: tornado activity during different phases of ENSO. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 21(8), 915-938.
- Gilks, W.R., Richardson, S., & Spiegelhalter, D.J. (Eds.), (1996). *Markov Chain Monte Carlo in Practice*, Chapman and Hall, New York.
- Heinze, G., Wallisch, C., & Dunkler, D. (2018). Variable selection—A review and recommendations for the practicing statistician. *Biometrical Journal*, 60(3), pp.431-449.
- Huffines, G.R., & Orville, R.E. (1999). Lightning ground flash density and thunderstorm duration in the continental United States: 1989–96. *Journal of Applied Meteorology*, 38(7), 1013-1019.
- Jagger, T.H., Elsner, J.B., & Widen, H.M. (2015). A statistical model for regional tornado climate studies. *PloS one*, 10(8).
- Johnston, K., Ver Hoef, J.M., Krivoruchko, K., & Lucas, N. (2003). ArcGIS9, Using ArcGIS geostatistical analyst, Redlands, CA: Environmental Systems Research Institute.

- King, P. (1997). On the absence of population bias in the tornado climatology of southwestern Ontario. *Weather and Forecasting*, 12(4), 939-946.
- Knapp, D.I. (1994). Using cloud-to-ground lightning data to identify tornadic thunderstorm signatures and nowcast severe weather. *Natl. Wea. Dig*, 19(2), 35-42.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1), 1-14.
- Lee, P. M. (2012). *Bayesian statistics: An introduction (4th ed.)*. West Sussex, UK: Wiley.
- MacGorman, D.R. & Burgess, D.W. (1994). Positive cloud-to-ground lightning in tornadic storms and hailstorms. *Monthly Weather Review*, 122(8), 1671-1697.
- Knapp, D.I., 1994. Using cloud-to-ground lightning data to identify tornadic thunderstorm signatures and nowcast severe weather. *Natl. Wea. Dig*, 19(2), pp.35-42.
- Newark, M.J. (1984). Canadian tornadoes, 1950–1979. *Atmosphere-Ocean*, 22(3), 343-353.
- Perez, A.H., Wicker, L.J., & Orville, R.E. (1997). Characteristics of cloud-to-ground lightning associated with violent tornadoes. *Weather and Forecasting*, 12(3), 428-437.
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Reap, R.M. & MacGorman, D.R. (1989). Cloud-to-ground lightning: Climatological characteristics and relationships to model fields, radar observations, and severe local storms. *Monthly Weather Review*, 117(3), 518-535.
- Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media, New York.
- Shephard, M.W., Morris, R., Burrows, W.R., & Welsh, L. (2013). A high-resolution Canadian lightning climatology. *Atmosphere-ocean*, 51(1), 50-59.

- Sills, D.M., Scriver, S.J. & King, P.W.S. (2004). The tornadoes in Ontario project (TOP). In Preprints from the 22nd AMS Conference on Severe Local Storms, Hyannis, Mass, *American Meteorological Society*, CD-ROM B (Vol. 7).
- Silverman, B.W. (1986). *Density estimation for statistics and data analysis* (Vol. 26). CRC press.
- Tan, L. (2008). Assessment of tornado hazard maps for southern Ontario, M.E.Sc. Thesis, the University of Western Ontario, London, Ontario, Canada.
- Tan, L., & Hong, H.P. (2010). Influence of spatial inhomogeneity of tornado occurrence on estimated tornado hazard. *Canadian Journal of Civil Engineering*, 37(2), 279-289.
- Tippett, M.K., Sobel, A.H., & Camargo, S.J. (2012). Association of US tornado occurrence with monthly environmental parameters. *Geophysical research letters*, 39(2).
- Twisdale, L.A., & Dunn, W.L. (1983). Probabilistic analysis of tornado wind risks. *Journal of Structural Engineering*, 109(2), 468-488.
- Wen, Y.K., & Chu, S.L. (1973). Tornado risks and design wind speed. *Journal of the Structural Division*, 99(12), 2409-2421.
- Wikle, C. K., & Anderson, C.J. (2003). Climatological analysis of tornado report counts using a hierarchical Bayesian spatiotemporal model. *Journal of Geophysical Research: Atmospheres*, 108(D24).
- Yarbrough Jr, J.W., & Meentemeyer, V. (1978). Seasonal and regional variation in the correlation of thunderstorm days with tornado frequency. *Journal of Applied Meteorology*, 17(11), 1741-1746.

Chapter 3

3 Development of a simple equivalent tornado wind profile for structural design and evaluation

3.1 Introduction

In structural design codes, the wind velocity or wind velocity pressure is given at 10 m height above the ground surface by considering a specified annual probability of exceedance or return period. This is a practical and useful approach in standardizing wind load for synoptic winds since the along height wind profile for the synoptic winds can be modeled using the power-law or logarithmic law (Simiu and Scanlan 1996), and the design wind pressure at any height can be calculated based on the recommended wind profile and the specified return period value of wind velocity at 10 m height.

For tornadoes, the along height wind profile varies with tornado intensity and other characteristics (Kuo 1971; Wen and Chu 1973; Wen 1975; Twisdale 1978; Twisdale et al. 1981, Wurman and Gill 2000; Hangan and Kim 2008; Refan and Hangan 2016; Honerkamp et al. 2020). Therefore, the use of a return period value of tornado wind velocity at 10 m height alone to define the tornado wind hazard without a corresponding wind profile may have questionable value. Moreover, the tornado occurrence rate is spatially varying (Wikle and Anderson 2003; Anderson et al. 2007; Tan and Hong 2010; Cheng et al. 2013, 2016). The spatially varying tornado occurrence rate influences the estimated rate of tornado striking structures of different footprints and affects the estimated return period value of tornado wind velocity at 10 m height (Twisdale et al. 1981; Twisdale and Dunn 1983; Banik et al. 2007, 2008; Tan and Hong 2010).

The spatially varying tornado activities for Canada is given in the commentary to the National Building Code of Canada (NRC 2015), but no guideline on the return period value of tornado wind velocity or along height wind profile is given. This does not facilitate the tasks of checking and evaluating structures subjected to the tornado wind velocity hazard. A comparison of the wind loads on the low-rise building due to tornado wind and synoptic wind indicates (Roueche et al. 2020) that the magnitudes of the peak tornado-induced pressures are reasonably similar to straight-line wind-induced pressures if the wind velocity

for the two types of winds is equal. This observation facilitates the possible implementation of tornado wind design in building codes for tornado hazard-prone regions, especially if the tornado hazard could be defined using quantiles of tornado wind velocity at 10 m height and an equivalent wind profile.

The quantiles or return period value of tornado wind velocity at a site can be estimated based on probabilistic spatially varying tornado occurrence model, tornado path characteristics, and tornado wind field model. The earlier systematic engineering development of tornado catalogue and quantitative tornado hazard assessment for Canada was presented by Newark (1984, 1991). It indicates that, on average, there is a tornado on Fujita-scale of F3 in every five years in southwestern Ontario, Canada. By using historical tornado catalogue and applying the adaptive Gaussian kernel smoothing technique, Tan and Hong (2010) evaluated the spatially varying tornado occurrence rate and striking rate for southern Ontario. An updated tornado catalogue from 1980 to 2009 for Canada was released by the Environment and Climate Change Canada (ECCC) (<https://open.canada.ca/data/en/dataset/fd3355a7-ae34-4df7-b477-07306182db69>). By using this catalogue, Cheng et al. (2013) estimated the spatially varying tornado occurrence rate for Canada based on a Bayesian hierarchical modelling framework. They considered the Poisson model and the zero-inflated Poisson (ZIP) model (Lambert 1992) for the modelling. Most importantly, they considered the reporting bias due to population density as well as the annual average cloud-to-ground lightning flash (ACGLF) density to develop their model. The use of ACGLF as an explanatory variable for the model development is justified since ACGLF is related to the tornado occurrence and tornado characteristics (Reap and MacGorman 1989; Branick and Doswell 1992; Knapp 1994; Carey and Buffalo 2007). It is noted that in the literature, the negative binomial (NB) model was also used to model the tornado occurrence (Elsner and Widen 2014; Jagger et al. 2015). One may also include the annual average thunderstorm days (ATD) as a covariate to estimate the tornado occurrence rate (Yarbrough and Meentemeyer 1978; Bissolli et al. 2007). By using the same tornado catalogue from ECCC, but considering ACGLF and ATD as the explanatory variables, Huang et al. (2020) estimated the tornado occurrence rate for Canada by applying the Bayesian hierarchical modelling technique and the maximum likelihood method. Their analysis results indicate that the use of the NB model is preferable to the use of the ZIP

model and Poisson model. The use of both ACGLF and ATD as the explanatory variables is preferable to use ACGLF alone as an explanatory variable to develop a tornado occurrence model. The preference is judged based on Akaike information criterion (AIC) and Bayesian information criterion (BIC) (Burnham and Anderson 2002). However, it is unknown whether such conclusions hold if only a tornado-prone region with relatively consistent and significant population density is considered.

It should be noted that the tornado catalogue released by ECCC includes events up to 2009, and statistics of tornado activities from 2010 to 2019 are not included. The tornado events from 2010 to 2019 could be gathered from relatively well-documented internet resources to update the catalogue for tornado occurrence modelling. The use of additional statistical data on tornado activities could also influence the estimated tornado occurrence rate.

Although several tornado wind field models are available in the literature, as mentioned earlier, only the probabilistic model parameters for the tornado wind field given in Twisdale et al. (1981) are well-calibrated and available in the literature. The model given by Twisdale et al. (1981) was successfully employed to assess the tornado wind velocity hazard for a site, a structural footprint, and a line system in Banik et al. (2007, 2008). It was also used to map the tornado wind velocity hazard for southern Ontario in Tan and Hong (2010) by considering spatially varying tornado occurrence.

The main objectives of the present study are to: a) estimate the return period value of the tornado wind velocity hazard along the height; b) develop an equivalent along height tornado wind profile that can be used to evaluate the bending moment and shear force for line-like structures and, c) to carry out disaggregation analysis to identify scenario tornado events based on selected exceedance probability. Although the analysis is focused on southern Ontario, Canada, the framework used to develop the equivalent tornado wind profile could be applicable to other regions. For the estimation and development, the tornado catalogue from 1980 to 2009 given by ECCC, as well as a newly assembled tornado catalogue from 1980 to 2019 for the region, are considered. The occurrence rate modelling is carried out based on the adaptive Gaussian kernel smooth, the Bayesian hierarchical modelling technique, and the maximum likelihood method. The probabilistic tornado wind

field model given in Twisdale et al. (1981) is adopted for the tornado wind velocity hazard modelling. In the following, first, the tornado catalogue, ACGLF, ATD, and population density used for the estimation of the tornado occurrence for southern Ontario are presented. The statistical analysis is then carried out to develop the tornado occurrence model and to estimate the tornado striking rate. The striking rate and the probabilistic tornado wind field model are employed to develop the equivalent tornado wind profile and obtain disaggregation results, which could facilitate the structural design and checking tasks and emergency preparedness planning.

3.2 Tornado occurrence rate and striking rate modelling

3.2.1 Tornado catalogue, explanatory variables, and population density for southern Ontario

As mentioned earlier, the released tornado catalogue by ECCC is for tornado events that occurred from 1980 to 2009 in Canada. The catalogue contains both confirmed and probable reported tornadoes; the catalogue does not include events earlier than 1980 because of data quality control. In addition, there is likely underreporting for the established tornado catalogue because of low population density in some regions in Canada (Newark 1991; King 1997). The tornado events in this catalogue for southern Ontario are extracted to form a catalogue that is referred to as TC80-09 for simplicity. The locations of the events in TC80-09 are presented in Figure 3.1a. To supplement this tornado catalogue with reported tornado activities from 2010 to 2019, a search of literature and internet resources is carried out, indicating that Wikipedia website (https://en.wikipedia.org/wiki/List_of_Canadian_tornadoes_and_tornado_outbreaks#cite_ref-ehfs_1-0; and https://en.wikipedia.org/wiki/List_of_21st-century_Canadian_tornadoes_and_tornado_outbreaks) provides relatively well-documented tornado events for Canada. By using this information and TC80-09, a new tornado catalogue, which is referred to as TC80-19, is formed for southern Ontario. The events listed in TC80-19 are presented in Figure 3.1b, where the tornado intensity of 7 events from 2010 to 2019 is not reported. In general, the spatial trends of tornado occurrence are similar for TC80-09 or TC80-19. Note that the reporting of the tornado intensity in Canada prior

to 2013 was based on the Fujita scale (F-scale). Since 2013, the adopted intensity scale for reporting was changed to enhanced Fujita scale (EF-scale) (Marshall et al. 2004) since 2013.

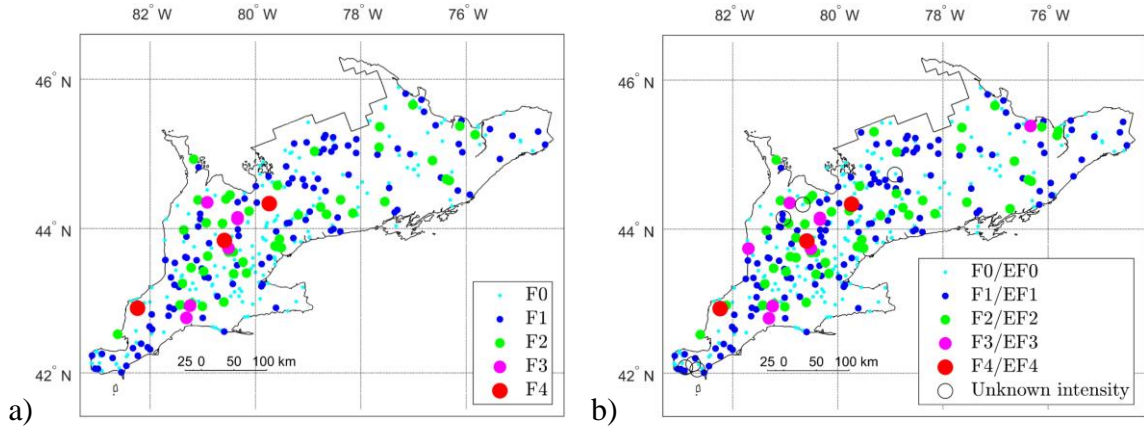


Figure 3.1. Tornadoes occurred in southern Ontario: a) For events from 1980 to 2009, b) for events from 1980 to 2019.

A more quantitative comparison of the reported tornadoes in these catalogues is shown in Table 3.1, according to F-scale (F_i) and EF-scale (EF_i). The table shows that the average number of the tornado for the region is about 12.4 based on TC80-09 and 11.4 based on TC80-19. This difference could be attributed to several factors, including that TC80-09 contains tornado events classified as confirmed and probable events. By following the detrending, and error and bias correction used in Banik et al. (2007) and Tan and Hong (2010) and without making a distinction between F_i and EF_i , the estimated probability mass function for the tornado intensity, $P(F_i)$, is also shown in Table 3.1. Since the tornado intensity for 7 events that occurred from 2010 to 2019 is unknown, and the consideration of F_i equal to EF_i to estimate $P(F_i)$ may be questionable, only $P(F_i)$ corresponding to TC80-09 is used in the remaining part of the present study. The estimated $P(F_i)$ indicates that the tornado intensity for about 73% and 91% of events is less than F_2 and F_3 , respectively.

Table 3.1. Reported tornadoes based on their intensity and the updated frequency distribution of tornado intensity.

F scale	Based on TC80-09		Based on TC80-19	
	Number	$P(F_i)$	Number	$P(F_i)$
F_0/EF_0	206	0.4248	246	0.4211

F_1/EF_1	115	0.3033	140	0.3044
F_2/EF_2	42	0.1802	51	0.1824
F_3/EF_3	7	0.0679	9	0.0690
F_4/EF_4	3	0.0198	3	0.0195
F_5/EF_5	0	0.0040	0	0.0035
Unknown intensity			7	
Total	373	1	456	1

Besides the development of tornado catalogues, the population density and the possible explanatory variables (i.e., ACGLF and ATD) are estimated for the region. To establish the density of ACGLF, the cloud-to-ground (CG) lightning flash data from the Canadian Lightning Detection Network (Burrows and Kochtubajda 2010; Shephard et al. 2013) for the period of 1999 to 2009 are used. The estimated ACGLF density for the region is shown in Figure 3.2a, where the squared grid system with the size of the cell equal to $4 \times 4 \text{ km}^2$ shown in the figure is used throughout the present study. A visual inspection of the spatial trends of the density presented in Figure 3.2a and the reported tornado events shown in Figure 3.1 indicates that there are similarities.

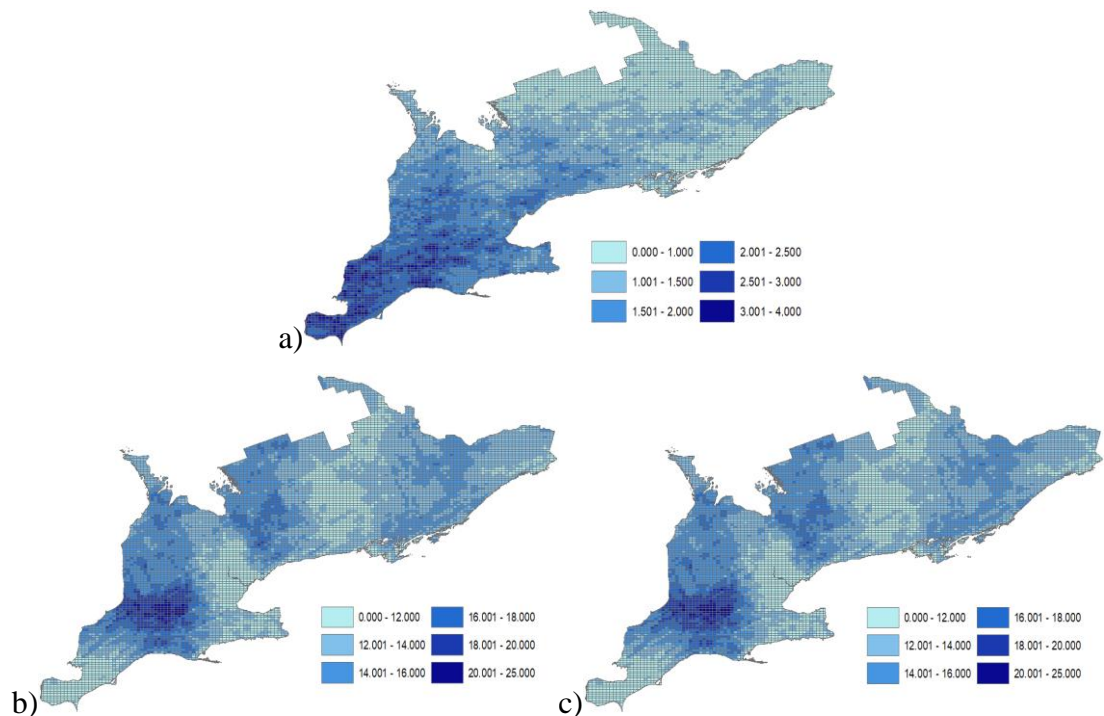


Figure 3.2. Estimated ACGLF, ATD: a) estimated ACGLF (per year per km^2) using data from 1999 to 2009, b) Estimated ATD based on the identifier used in ECCC DLY04

database from 1980 to 2009, and c) Estimated ATD based on the identifier used in ECCC DLY04 database from 1980 to 2017.

For the estimation of the thunderstorm days, it was noted that a thunderstorm identifier is used for the daily wind record stored in ECCC DLY04 digital archive (see http://www.climate.weatheroffice.gc.ca/prods_servs/documentation_index_e.html#hly01). The identifier is used to indicate if the recorded wind at a station is from a thunderstorm or non-thunderstorm event. By using the wind records from the available stations within the region, the number of thunderstorm days is counted by using the thunderstorm identifier, and ATD at each station is then calculated. A spatially smoothed ATD obtained based on the calculated ATD at each meteorological station is presented in Figure 3.2b based on wind record from 1980 to 2009 and Figure 3.2c based on wind records from 1980 to 2017 (ideally, data from 1980 to 2019 should be used for consistency with TC80-19, but the data from 2018 and 2019 are unavailable to this project). For the plot, the ordinary kriging (Johnston et al. 2003) is used for spatial smoothing. A comparison of the results depicted in Figures 3.1, 3.2b, and 3.2c indicates that there are similarities in their spatial patterns, suggesting that the ATD could be valuable as an explanatory to develop the tornado occurrence model.

For the population density estimate, the Canadian 2001 and 2011 census data are obtained from

<https://www12.statcan.gc.ca/english/census01/products/standard/popdwell/tables.cfm>.

The census is taken every ten years in Canada. The 2001 and 2011 census data are taken near the middle of the observation periods for the catalogues TC80-09 and TC80-19, respectively. The population data which is given by counties are assumed to be uniformly distributed within each census subdivision. Based on this assumption, the spatially distributed population density is estimated and shown in Figure 3.3, indicating that the population density is greater than 5 (people/km²) for the majority of locations.

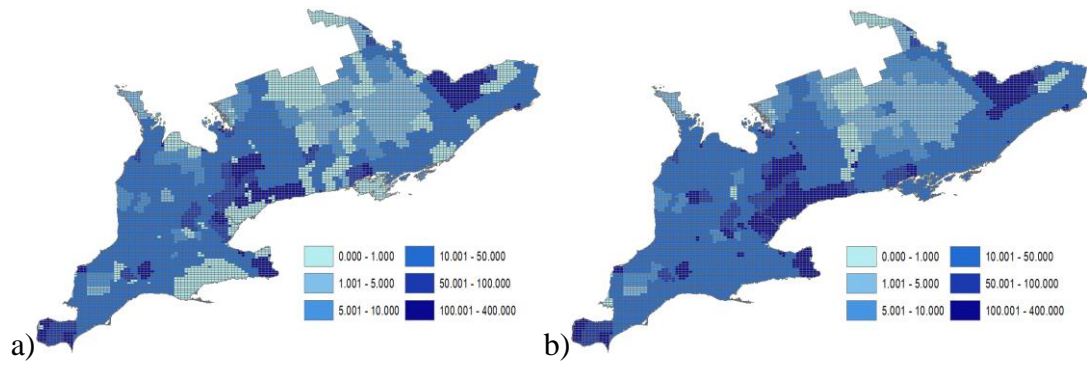


Figure 3.3. Estimated population density (number of people/km²): a) population density based on 2001 Canadian census data, e) population density based on 2011 Canadian census data.

3.2.2 Estimation of tornado occurrence rate for southern Ontario

By following Tan and Hong (2010) and applying the two-dimensional adaptive Gaussian kernel estimation technique (Silverman 1986), the estimated tornado occurrence rate is presented in Figures 3.4a and 3.4b for the tornado events depicted in Figure 3.1a and 3.1b, respectively. The spatial trends of the obtained smoothed tornado occurrence rate in Figures 3.4a and 3.4b are very similar. The inhomogeneity of tornado occurrence for southern Ontario is recognized by many researchers, including King et al. (2003). They suggested that the lake breeze boundary generated convection may be the dominant mechanism resulting in a preferred southwest-to-northeast pattern of tornado activity.

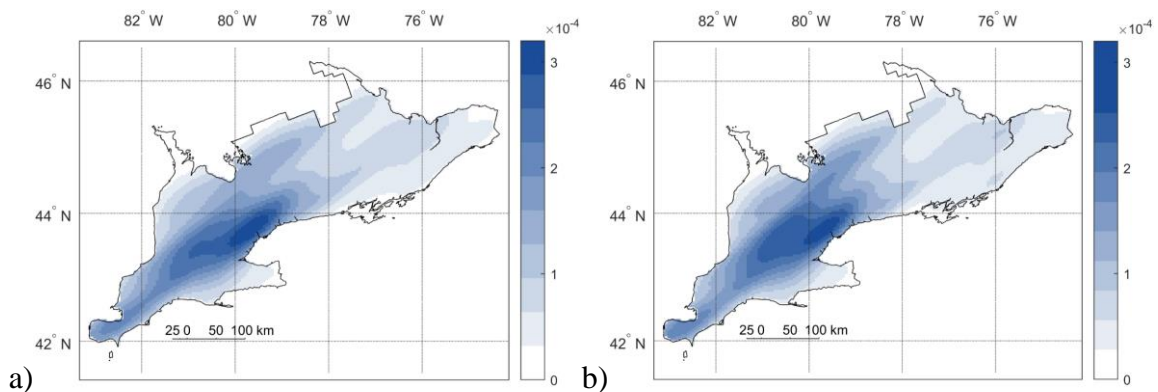


Figure 3.4. Estimated spatially varying tornado occurrence rate (number/(year \times km²)) by using adaptive Gaussian kernel smoothing: a) based on TC80-09, and b) based on TC80-

To further investigate the tornado occurrence rate for the region, it is considered that the occurrence could be modeled using the Poisson, ZIP, or NB models (Wikle and Anderson 2003; Anderson et al. 2007; Cheng et al. 2013, 2016; Elsner and Widen 2014; Jagger et al. 2015). Since our preliminary results based on AIC, BIC, and the estimates from the maximum likelihood method (MLM) indicate that the use of the NB model is preferable, only the NB model is given below by applying MLM and Bayesian hierarchical modelling technique.

By adopting the NB model, the probability distribution of the number of tornado events N_{tj} at the j -th cell is given by (Elsner and Widen 2014; Jagger et al. 2015),

$$P(N_{tj} = n) = \frac{\Gamma(n + rt)}{n! \Gamma(rt)} p_{NBj}^r (1 - p_{NBj})^n, \quad (3.1)$$

where r is a model parameter, and p_{NB} is the probability of non-occurrence for the j -th cell. The mean of N_{tj} , μ_j , equals $rt(1 - p_{NBj}) / p_{NBj}$, and the variance equals $rt(1 - p_{NBj}) / p_{NBj}^2$. The variance can also be expressed as $\mu_j + \mu_j^2 / r$ by simple algebraic manipulations. This indicates that the NB model could be adequate for stochastic point processes with overdispersion; the overdispersion decreases as r increases. In order to relate N_{tj} and the reported tornado events Y_{tj} , it is considered that an occurred tornado within the j -th cell has a probability p_{tj} of been reported (i.e., reporting probability). This results in that Y_{tj} is represented by the NB model as well, except that the model parameter p_{NBj} is replaced with p_{NBBj} , where

$$p_{NBBj} = p_{NBj} / (p_{tj} + p_{NBj}(1 - p_{tj})). \quad (3.2)$$

and the mean of Y_{tj} for the considered observation period, denoted as $t\eta_j$, equals $rt(1 - p_{NBBj}) / p_{NBBj}$, which can be re-written as,

$$p_{NBBj} = r / (r + \eta_j), \quad (3.3)$$

Following Wikle and Anderson (2003), Anderson et al. (2007), and Cheng et al. (2013), it is assumed that η_j (number of tornado/(year \times km²)) could be modeled using,

$$\ln \eta_j = -\beta / \gamma_j + \alpha_0 + \sum_{k=1}^{n_k} \alpha_k z_{kj} + \varepsilon_j, \quad (3.4)$$

where γ_j is the population density for the j -th cell; $\ln p_{ij} = -\beta / \gamma_j$ (Wikle and Anderson 2003; Anderson et al. 2007) is used to represent reporting probability with model parameter β ; the remaining terms on the right-hand side of the equation are used to represent the occurrence rate (Cheng et al. 2013) with model parameters α_k and the explanatory variables z_{kj} for $k = 0, \dots, n_k$; and ε_j is the residual which is modelled using the conditional autoregressive model (Besag 1974; Besag and Kooperber 1995). More specifically, ε_j is normally distributed $N(\zeta_j, 1/(\delta^2 n_{A_j}))$, where $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the normal distribution with the first and second arguments representing the mean (vector) and variance (covariance matrix), n_{A_j} is the number of adjacent cells of the j -th cell, and $\zeta_j = (1/n_{A_j}) \sum_k \varepsilon_k$ with k denotes all the adjacent cells of the j -th cell, δ^2 is a model parameter. The coefficients α_k in Eq. (3.4) are considered to be independent of the cell (i.e., spatial location) because of the scare of tornado occurrence data.

It can be shown that the likelihood function for the considered probabilistic model of Y_{ij} , \mathbf{L} , is can be written as,

$$\mathbf{L} = \prod_{j=1}^{M_c} \left[\frac{\Gamma(y_j + rt)}{y_j! \Gamma(rt)} \left(\frac{r}{r + \eta_j(\boldsymbol{\theta})} \right)^{rt} \left(\frac{\eta_j(\boldsymbol{\theta})}{r + \eta_j(\boldsymbol{\theta})} \right)^{y_j} \right]. \quad (3.5)$$

where M_c is the total number of cells for the considered region, $\boldsymbol{\theta} = (\alpha_0, \dots, \alpha_k, \beta)$, and the notation $\eta_j(\boldsymbol{\theta})$ is used instead of η_j to emphasize its dependency on the model parameters.

For the numerical analysis, the considered cases are listed in Table 3.2. These cases are based on combinations of two tornado catalogues TC80-09 and TC80-19, and the use of ACGLF or ATD or both ACGLF and ATD as explanatory variables. For each case, the MLM is applied, and AIC and BIC are calculated and presented in Table 3.2 as well. Since, for the same set of data, a model with the lowest AIC (or BIC) among the considered models is preferable (Burnham and Anderson 2002), the results presented in the table

indicate that the model with ACGLF as an explanatory variable is preferred (i.e., Case A1 and B1). The differences in the calculated AIC and BIC by using ACGLF or both ACGLF and ATD as explanatory variables are not very large. This observed behavior differs from that observed based on the analysis result carried out for Canada, where the use of ACGLF and ATD as explanatory variables is preferred (Huang et al. 2020), and the preference is followed by using ACGLF as an explanatory variable. This emphasizes that the preferred model is data-driven and influenced by the uniformity of population density.

Based on the above observations, the application of Bayesian hierarchical modelling to estimate model parameters is carried out only for Cases A1 and B1. For the estimation, the Markov chain Monte Carlo (MCMC) technique (Gilks et al. 1996; Brooks et al. 2011) is employed. The WinBUGS software (<http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml>) is used to implement the Metropolis-Hasting algorithm to carry out MCMC. The Brooks-Gelman-Rubin diagnostics are used for convergence assessment. Moreover, the prior probability distribution used for the analysis are $r \in \text{Gamma}(1, 0.01)$, $\alpha_0 \in N(0, 10^4)$, $\alpha_1 \in N(0, 10^4)$, $\alpha_2 \in N(0, 10^4)$, $\beta \in N(0, 10^4)$, and $\delta^2 \in \text{Gamma}(1, 0.01)$, where $\text{Gamma}(\square, \square)$ denotes the gamma distribution with the first and second arguments representing the shape and rate parameters. The assignment of the prior probability distribution is based on the “non-informative” prior consideration. The obtained model parameters are shown in Table 3.3.

The β value shown in the table is much smaller than 0.025 that is obtained in Huang et al. (2020) if the NB model is used for Canada. This discrepancy can be explained by noting that the population density in southern Ontario is consistently higher and relatively uniform than in many regions in Canada. The small β value shown in Table 3.3 also indicates that the consideration of population bias to develop the tornado occurrence model is not important for southern Ontario. Also, the average number of the tornado occurrence rate (per year) for the considered region is estimated based on the developed model. The calculated average values are shown in Table 3.3; they compare favorably with the estimated value directly from TC80-09 and from TC80-19 (which are 12.4 and 11.4). The predicted occurrence rate based on the developed models shown in Table 3.3 is presented

in Figure 3.5. The rate mimics well the spatial inhomogeneity of the tornado occurrence that is depicted in Figure 3.1 based on the catalogues. A comparison of the tornado occurrence rate shown in Figures 3.4 and 3.5 indicates that the rate presented in Figure 3.4 is smoother than that shown in Figure 3.5. The southwest-to-northeast pattern shown in Figure 3.5 is less apparent than that shown in Figure 3.4.

Table 3.2. Cases considered for the statistical analysis of tornado occurrence (see Eq. (3.4). z_{1j} is used to represent ACGLF, and z_{2j} is used to represent ATD).

Case	Tornado catalogue	Explanatory variable	AIC	BIC	Tornado catalogue	Case	Explanatory variable	AIC	BIC
A1	TC80-09	ACGLF	3021.5	3042.5	TC80-19	B1	ACGLF	3021.5	3042.5
A2		ATD	3089.2	3110.1		B2	ATD	3089.2	3110.1
A3		ACGLF & ATD	3022.0	3049.9		B3	ACGLF & ATD	3022.0	3049.9

Table 3.3. Estimated model parameters for Cases A1 and B1 based on Bayesian hierarchical modelling technique (the first and second entries represent the mean and standard deviation of the model parameters).

Case	r	α_0	α_1	β	Predicted rate
A1	0.397; 0.098	-4.060; 0.151	0.658; 0.078	0.001; 0.001	12.85
B1	0.362; 0.072	-3.777; 0.138	0.619; 0.072	0.001; 0.001	11.95

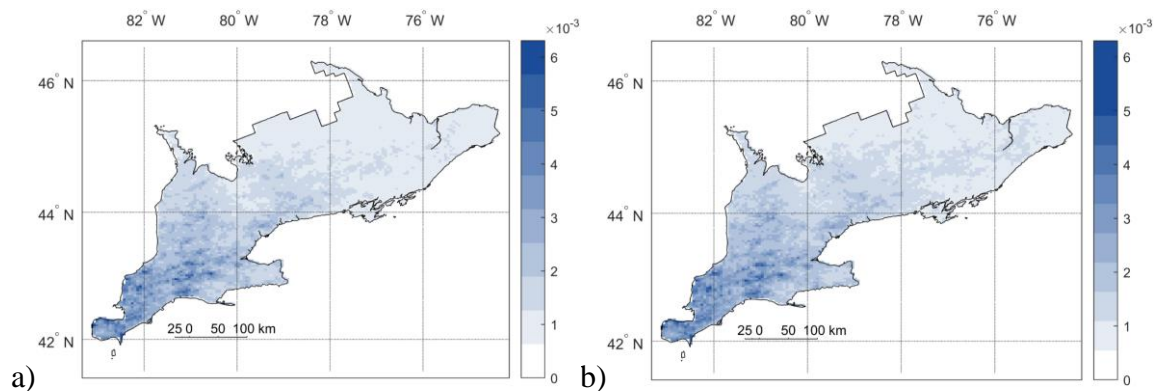


Figure 3.5. Estimated spatially varying tornado occurrence rate (number/(year \times km²)) based on the NB model with the model parameters shown in Table 3.3: a) based on Case A1, and b) based on Case B1.

3.2.3 Characteristics of tornado path and striking rate

The tornado occurrence rate, the tornado intensity distribution, and tornado path characteristics affect the rate that tornado striking a structure and the estimated tornado wind velocity hazard mapping. The tornado path characteristics for the considered region were already discussed extensively in Banik et al. (2007) and Tan and Hong (2010) for southern Ontario (see appendix B). This includes the probability distributions of tornado path direction, length, and width. The path direction is described by the probability mass function for eight directions, with one direction oriented towards the north and each separated by 45 degrees. The tornado path length and tornado path width for a given tornado intensity are modeled using the truncated Weibull distribution. In addition to these probability distributions, the actual tornado intensity, F_{Aj} , varies randomly along its path for a given tornado with reported or classified intensity F_i . Twisdale et al. (1981) provided a probability distribution model defining the percentage of the path length of striking intensity F_{Aj} conditioned on F_i . By using these probability distributions, Tan and Hong (2010) provided a detailed step by step simulation procedure to estimate the annual average tornado striking rate, γ_{Aj} , defined as a site that is hit by γ_{Aj} times per year with the actual intensity F_{Aj} . Their procedure, which is illustrated in Figure 3.6, is used in the following to evaluate γ_{Aj} .

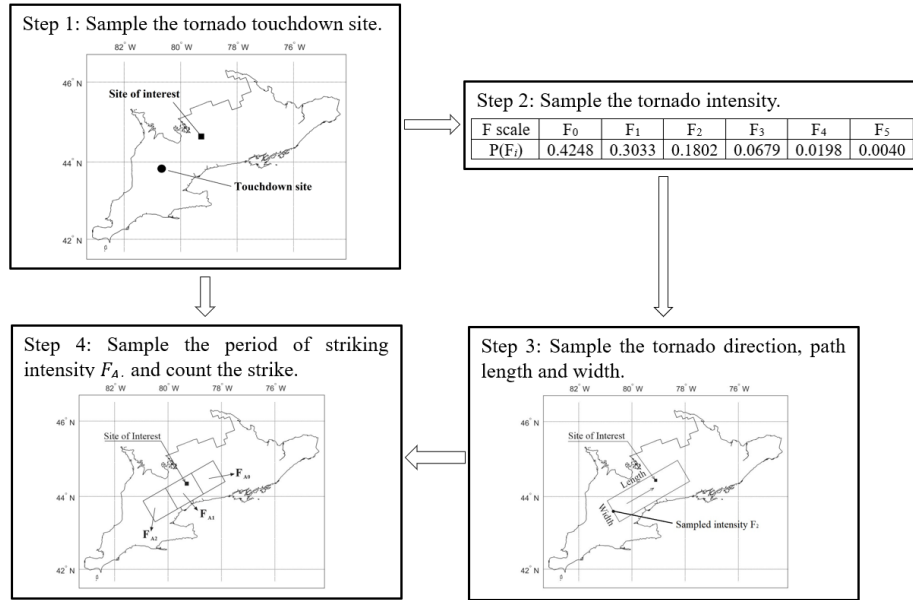
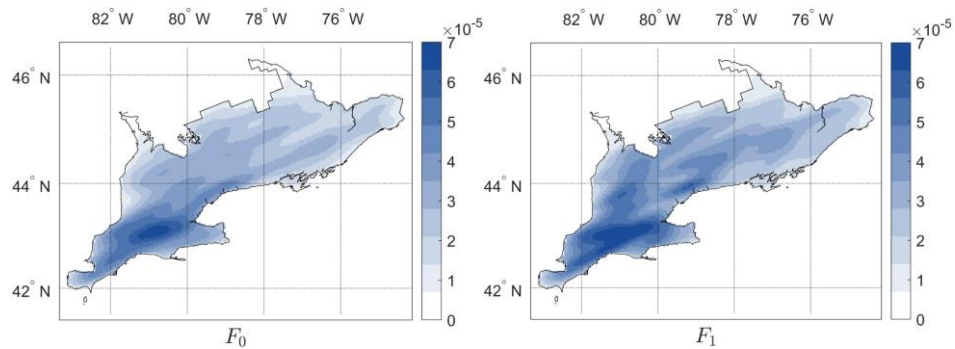


Figure 3.6. Steps of calculating the striking rates for different F_{Aj} .

For example, by considering the tornado occurrence rate depicted in Figure 3.5a, and carrying out the simulation, the obtained tornado striking rate is shown in Figure 3.7. For the analysis, 100,000 years of tornado activities are simulated. The estimated striking rate shows a clear southwest-to-northeast pattern. A similar analysis is carried out by using the occurrence rate shown in Figures 3.4 and 3.5a. As the obtained spatial trends of γ_{Aj} are the same as those shown in Figure 3.7, they are not plotted. Since the results presented in Figure 3.7 are associated with the highest annual occurrence rate for the region (i.e., 12.8 events per year as depicted in Table 3.3), as a conservative measure, only these striking rates are used in subsequent numerical analysis.



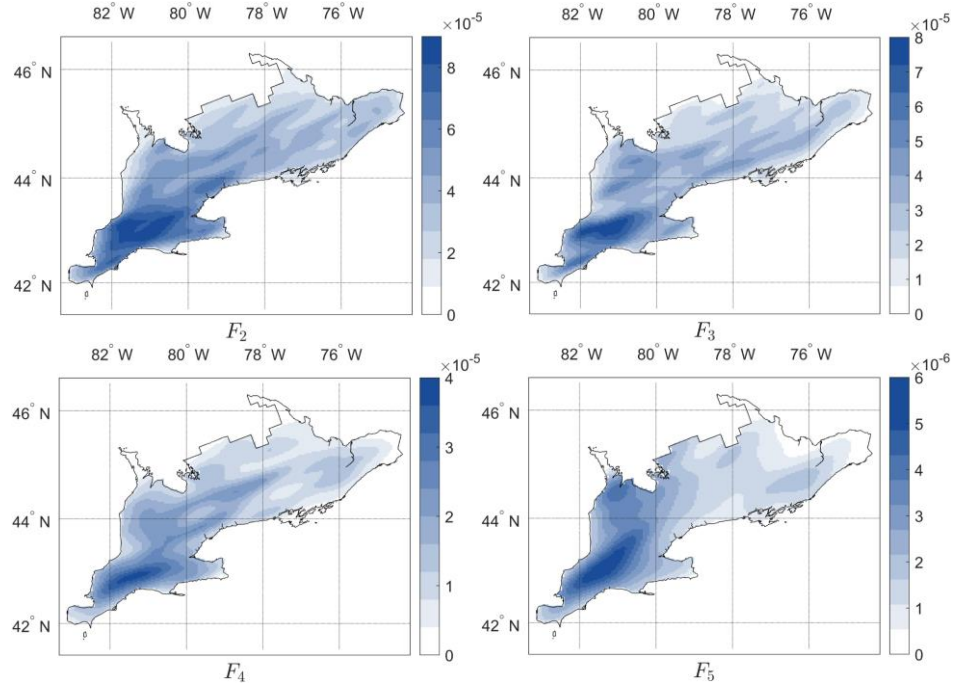


Figure 3.7. Spatially varying tornado striking rate with intensity F_{A_j} by using the tornado occurrence rate shown in Figure 3.5a.

3.3 Estimation of tornado wind velocity hazard

3.3.1 Assessment of probability distribution of tornado wind velocity

Consider that a site experiences a tornado with the actual tornado intensity F_{A_j} . Since the parameters of the wind field, as well as the position of the site within a tornado wind field, are uncertain, the tornado horizontal wind velocity, $V(z)$, at a height z (m) above the ground surface at the site of interest is uncertain. Let $P(V(z) < v | F_{A_j})$ denote the probability distribution of the tornado wind velocity for a point structure at z m height above the ground surface for the considered site, conditioned on that the site is hit by the tornado with F_{A_j} .

The procedure to assess $P(V(z) < v | F_{A_j})$ was already given in Tan and Hong (2010) for $z = 10$ m by using the probabilistic wind field model given in Twisdale et al. (1981) (see also Dunn and Twisdale (1979)). This procedure that is adopted for the following numerical

analysis is illustrated in Figure 3.8. The procedure involves simulating the wind field based on a set of model parameters defined by their corresponding probability distribution models (Dunn and Twisdale 1979; Twisdale et al. 1981) and evaluating the maximum wind velocity (or wind actions) experienced at the site of interest. By using this procedure, the estimated $P(V(z) < v | F_{Aj})$ is carried out for z ranging from 0 to 80 m with an increment of 1 m (the evaluation of probability distribution of bending moment and shear force for a line-like structure is discussed in the next section). A simulation cycle of 30000 is employed for assessing this conditional distribution. The obtained distributions are stored in a database, referred to as DIST-database, to be used to map tornado wind velocity hazard for a specified return period, which is elaborated shortly. The obtained $P(V(z) < v | F_{Aj})$ is illustrated in Figures 3.9a to 3.9d for a few selected z values, showing that $P(V(z) < v | F_{Aj})$ is shifted to the right as F_{Aj} increases, which is expected. An inspection of the plots indicates that the distributions for different z are not identical and $P(V(z) < v | F_{Aj})$ for a given F_{Aj} is shifted to the right as z increases. The steepest increase occurs for z varies from 10 to 20 m.

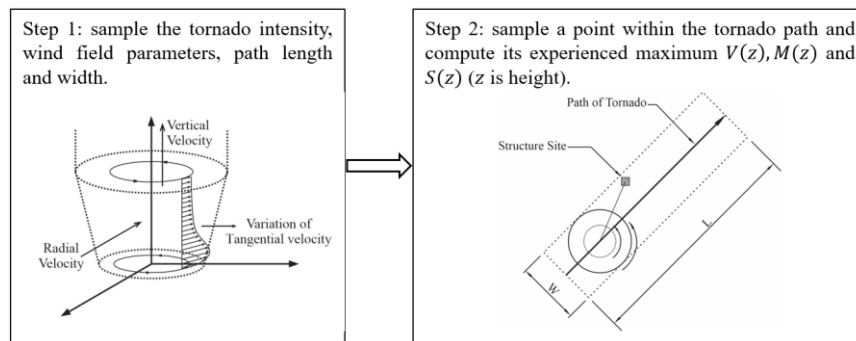


Figure 3.8. Steps to evaluate the probability distributions of wind velocity, and bending moment and shear force for a line-like structure.

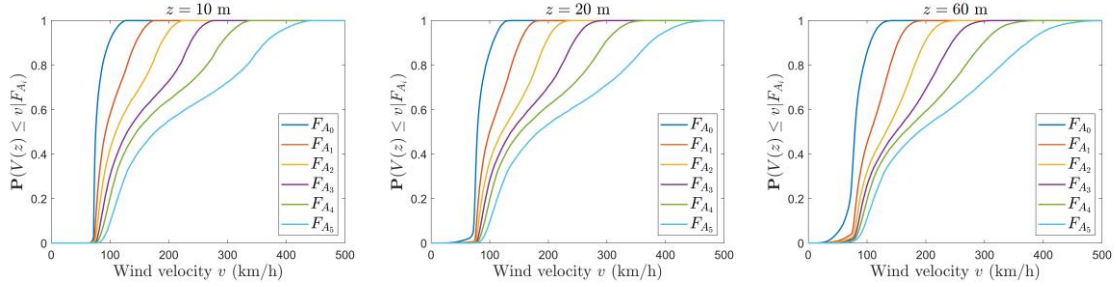


Figure 3.9. Assessed $P(V(z) < v | F_{A_j})$ for selected z values.

3.3.2 Assessment of probability distribution of bending moment and shear force

The estimated tornado wind velocity at the height z can be used as the basis to calculate the wind load effect for a point-like structure at z m height. However, for a line-like structure, the usefulness of the quantiles of $V(z)$ calculated by using $P(V(z) < v | F_{A_j})$ is unclear. This is because that the estimated quantiles at different height could correspond to different tornadoes, and that the application of the estimated quantiles to calculate the bending moment or shear force may be conservative for a line-like structure, which is idealized as a circular prismatic structure with a width of B and height H (m). To investigate the degree of conservatism, let $M(z, H)$ and $S(z, H)$ denote the maximum bending moment and of shear force at a height z for a line-like structure with a height of H , where

$$M(z, H) = \max_{\phi} \left(\frac{1}{2} C_D \rho_{air} B \int_z^H v^2(x, \phi) (x - z) dx \right) \quad (3.6)$$

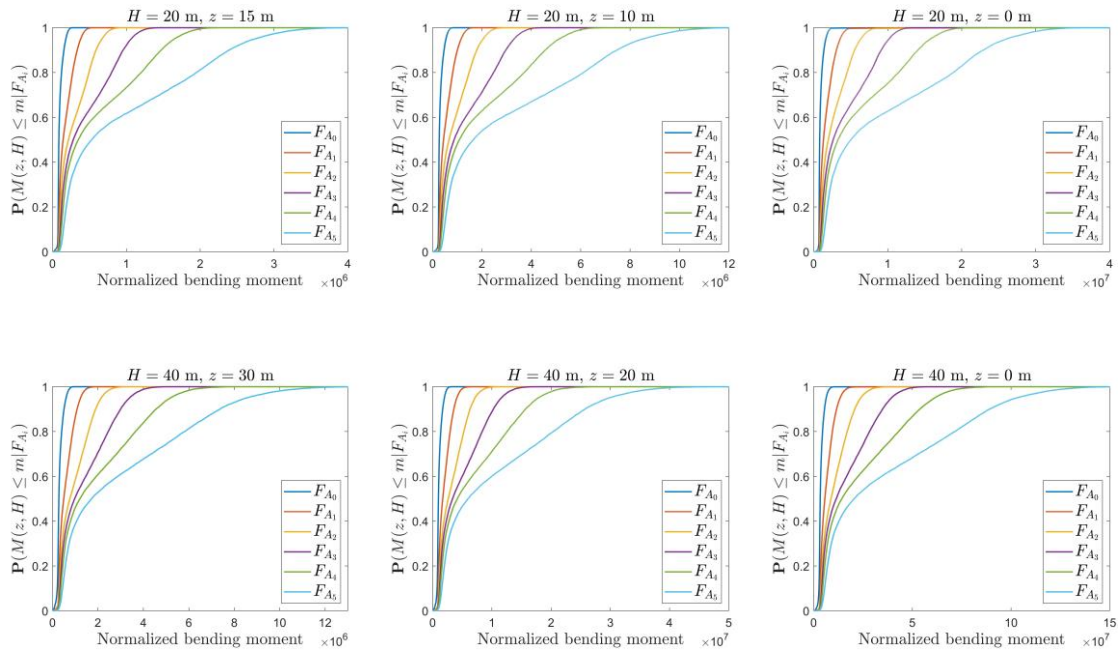
and

$$S(z, H) = \max_{\phi} \left(\frac{1}{2} C_D \rho_{air} B \int_z^H v^2(x, \phi) dx \right) \quad (3.7)$$

where $v(x, \phi)$ in these two equations denotes the tornado horizontal wind velocity along the same direction ϕ from the same tornado, C_D represents the drag coefficient, and ρ_{air} denotes the air density. Following the same steps in sampling tornado wind velocity, as shown in Figure 3.8, the values of $M(z, H)$ and $S(z, H)$ are simulated, except, in this case, an additional step in calculating $M(z, H)$ and $S(z, H)$ by using Eqs. (6) and (7) based on each simulated tornado wind field is required. Using the samples of $M(z, H)$ and $S(z, H)$, the probability

distribution of $M(z, H)$, $P\left(M(z, H) < m | F_{Aj}\right)$ and of $S(z, H)$, $P\left(S(z, H) < s | F_{Aj}\right)$, conditioned on that the line-like structure of height H experiences the tornado of intensity F_{Aj} , are assessed. For the assessment, H up to 80 m, covering the majority of line-like structures, is considered. For simplicity, all the sampled bending moments and shear forces are normalized with respect to $\frac{1}{2} C_D \rho_{air} B$. The assessed distributions are also stored in DIST-database that are used to evaluate an equivalent tornado wind profile in the following sections.

For selected values of H and z , the obtained distributions are illustrated in Figures 3.10 and 3.11. Figure 3.10 shows that $P\left(M(z, H) < m | F_{Aj}\right)$ depends on F_{Aj} , H and z , and it is shifted to the right as F_{Aj} and H increases, and z decreases, which is expected. Because the shape of distributions does not follow some commonly employed simple probability distribution models, no distribution fitting is carried out. These same observations are applicable to $P\left(S(z, H) < s | F_{Aj}\right)$ shown in Figure 3.11.



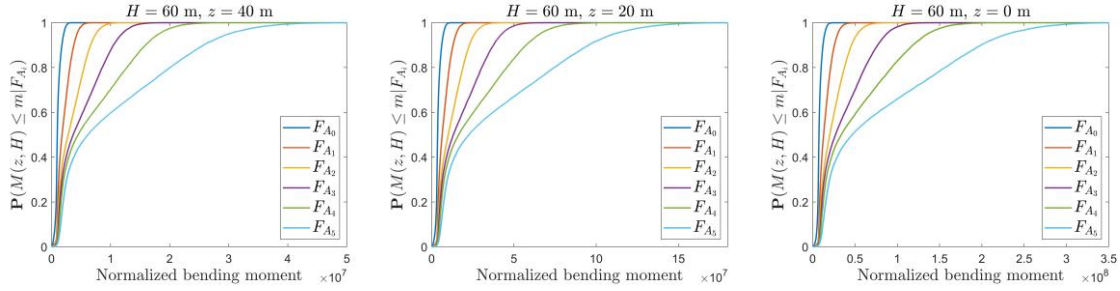


Figure 3.10. Estimated $P(M(z, H) < m | F_{A_j})$ the normalized bending moment equals

$$M(z, H) / (C_D \rho_{air} B / 2) (\text{km}^2 \times \text{m}^2 / \text{h}^2).$$

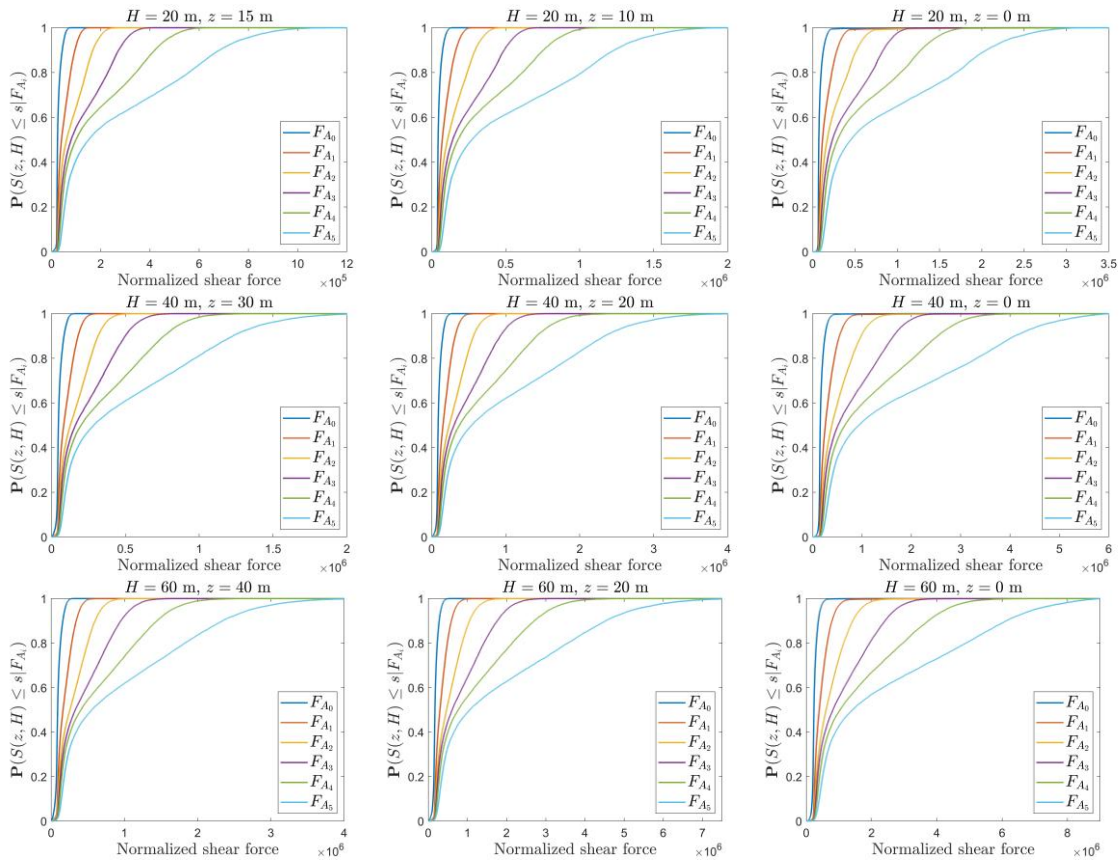


Figure 3.11. Estimated $P(S(z, H) < s | F_{A_j})$, where the normalized shear force equals

$$S(z, H) / (C_D \rho_{air} B / 2) (\text{km}^2 \times \text{m} / \text{h}^2).$$

3.4 Tornado wind hazard evaluation, disaggregation, and mapping

3.4.1 Quantile of tornado wind velocity along the height

By assuming that the estimated average annual striking rate γ_{Aj} is small, the probability that $V(z)$ is greater than v in a year, $P_T(V(z) > v)$, can be approximated using,

$$P(V(z) > v) \approx \sum_{j=0}^5 \gamma_{Aj} P(V(z) > v | F_{Aj}). \quad (3.8)$$

As an illustration, consider the site with latitude and longitude equal to (43.0338N and 81.1498W) representing London, Ontario. The annual striking rate for the site $[r_{A0}, r_{A1}, r_{A2}, r_{A3}, r_{A4}, r_{A5}]$ obtained from Figure 3.7 equals $[6.40, 6.86, 8.25, 5.15, 1.85, 0.51] \times 10^{-5}$. By using these rates and the pre-calculated distribution and stored $P(V(z) > v | F_{Aj})$, and solving Eq. (3.8), the obtained $(1-1/T)$ -quantile of $V(z)$, $V_T(z)$, (i.e., T -year return period value of $V(z)$) is shown in Figure 3.12 for $T = 5000, 7500, 10000, 20000, 50000$ and 100000 years. The bend at 10 m height shown in the curves plotted in Figure 3.10 is due to the adopted wind field model. This aspect will be further investigated using the results from the disaggregation analysis shortly.

Since the tabulation of $V_T(z)$ for a range of heights for the sites within a region may not be very practical for codified structural design implementation and application, it would be valuable to simply tabulate $V_T(10)$ and use it with an equivalent tornado wind profile $V_{TA}(z)$. As a practical measure, the following simple parametric form for $V_{TA}(z)$,

$$V_{TA}(z) = \begin{cases} V_T(10) & \text{for } z < 10 \text{ m} \\ V_T(10) \times \left(\frac{z}{10}\right)^{\max(0, \alpha_1 - \alpha_2 * z)} & \text{for } 10 \leq z < 80 \text{ m} \end{cases} \quad (3.9)$$

is used to fit the wind profile $V_T(z)$, where α_1 and α_2 are model parameters. The use of the quantile of wind velocity at 10 m height as the reference wind velocity is consistent with the treatment of synoptic winds for structural design. The fitted curves by using Eq. (3.9) for the cases shown in Figure 3.12 are also presented in the same figure, where the obtained (α_1, α_2) equals $(0.0580, 0.0003)$, $(0.0888, 0.0003)$, $(0.0910, 0.0004)$, $(0.0516, 0.0002)$,

(0.0266, 0.0001) and (0.0265, 0.0001) for $T = 5000, 7500, 10000, 20000, 50000,$ and 100000 years, respectively. This shows that α_1 and α_2 depend on T . The plot indicates that the simple model shown in Eq. (3.9) provides an adequate fit. For T ranging from 5000 to 10000 years, the relative differences between $V_T(80)$ and $V_T(10)$ vary but are less than about 15%. As T increases beyond 10000 years, this relative difference is less than 5%.

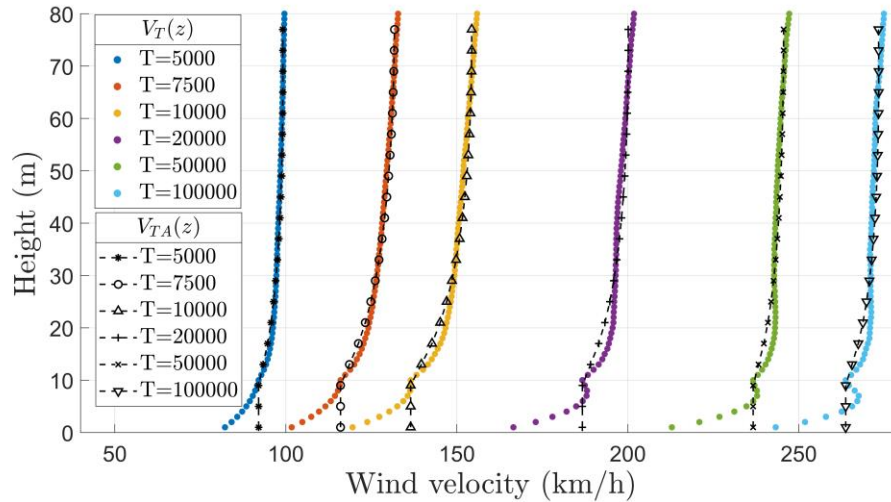


Figure 3.12. Estimated quantiles of $V(z)$, $V_T(z)$, (3-s gust wind velocity) for London, Ontario, and fitted Eq. (3.9) for z up to 80 m.

Note that based on Eq. (3.8), it can be shown that for two sites A and B, if γ_{Aj} for Site A equals ξ times of the annual striking rate for Site B, where ξ is a positive scaling factor, the T -year return period value of tornado wind velocity for Site A is approximately equal to the (T/ξ) -year return period value of the tornado wind velocity for Site B. This implies that α_1 and α_2 are likely to vary from location to location because they depend on T and the striking rate varies spatially. It also suggests that one may relate this dependency by using $V_T(10)$ so to establish simple predicting equations for α_1 and α_2 based on $V_T(10)$ for practical applications. The relationships between α_1 and $V_T(10)$ and between α_2 and $V_T(10)$ will be further explored in the following sections by using results from many sites.

3.4.2 Disaggregation of tornado wind velocity hazard

The disaggregation analysis is now frequently used to identify the scenario seismic events for a specified seismic hazard and risk level (Bazzurro and Cornell 1999; Hong and Goda

2006). The disaggregation analysis is used to identify possible events leading to a specified hazard level or exceeding a specified hazard level. The identified scenario events are valuable for emergency preparedness and for evaluating structural performance. In the following, the disaggregation concept is considered to find wind profiles that nearly results in the specified $V_T(z)$ at a specified height. Given the values of $V_T(z)$ and z for a site of interest, the analysis is carried out by:

- 1) Sample the occurrence of the tornado according to the flowchart shown in Figure 3.6;
- 2) If the segment of the tornado track with intensity F_{A_j} covers the site of interest, simulate the tornado wind field according to the steps shown in Figure 3.8 and calculate the wind velocity at the height z for the considered site.
- 3) If the difference between the calculated velocity $V(z)$ and $V_T(z)$ is within a given tolerance, save the identified tornado wind profile $V(z)$.
- 4) Repeat Steps 1) to 3) to identify a sufficient number of scenario wind profiles.

As a numerical example, we consider the site representing London, Ontario, again. By carrying out the disaggregation analysis for $V_T(10)$, $V_T(20)$, and $V_T(40)$ for $T = 5000$ and 10000 years, the identified wind profiles are shown in Figure 3.13. The figure indicates that different wind profiles could lead to the same $V_T(z)$. In particular, the disaggregation results for $V_T(10)$ indicate that the most identified wind profiles are associated with monotonic increasing wind profiles. This is especially the case for $T = 5000$ years. For $z = 20$ m, there are many identified wind profiles that are not monotonically increasing functions of height and are exhibiting a large wind velocity for z below 20 m. As z increases to 40 m, most of the identified wind profiles are monotonically increasing wind profiles if $T = 5000$ years. However, as T increases, some of the identified wind profiles exhibit a large wind velocity for height below 40 m. An inspection of the observed discontinuities in the wind profile present in the plots indicates that they are due to the adopted wind field model. The disaggregation results for given $V_T(z)$ also show that the two identified wind profiles for given $V_T(z)$ (and specified z and T) could lead to very different bending moment and shear force for a line-like structure. Therefore, use a single identified $V_T(10)$ to evaluate the wind load actions for the line-like structure may not be appropriate. Moreover, it can be shown that in all cases, the synoptic wind profile (defined by the power-law with an exponent of 0.11 and conditioned on the wind velocity at 10 m

height) envelops the wind velocity of those given by tornado wind profile for $z > 10$ m. Also, a rectangular wind profile within its values equals to $V_T(10)$ can underestimate the wind load effect for the line-like structure. In general, given the tornado wind velocity at 10 m height, the rectangular wind profile and the synoptic wind profile could be used as crude lower and upper bounds for the tornado wind profiles, respectively, if $z > 10$ m.

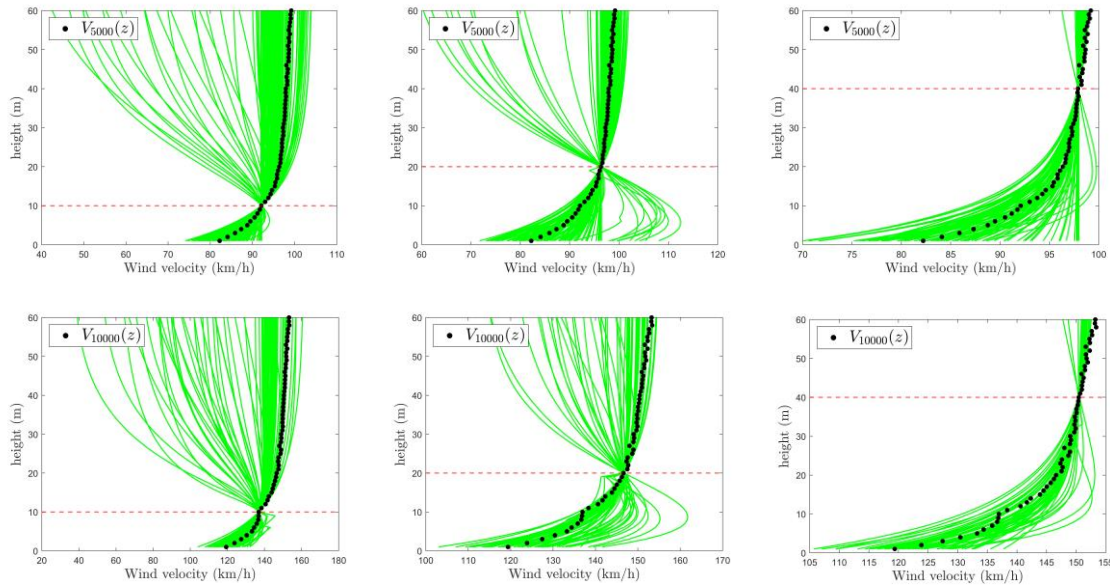


Figure 3.13. Identified wind profile based on disaggregation analysis for $T = 5000$ and 10000 years.

3.4.3 Tornado hazard mapping

The analysis carried out for the results presented in Figure 3.12 is repeated for other sites within southern Ontario, the estimated $V_T(10)$ is shown in Figure 3.14 for $T = 10000$, 50000 and 100000 years. The hazard mapping for $T = 5000$ and 7500 years are not presented. This is because, for many sites, the probability of tornado striking a site is less than 1/5000 or 1/7500 per year, resulting in that the tornado wind speed equals zero for an exceedance probability greater than or equal to 1/5000 or 1/7500, respectively. Figure 3.14 shows that the estimated $V_T(10)$ is spatially varying, and the spatial patterns resemble those shown in Figures 3.5 and 3.7 for the tornado occurrence rate and striking rate. It is noteworthy that the factored design wind load, according to the NBCC (2015) for southern Ontario, ranges from about 300- to 1900-year return period value of the annual maximum synoptic wind velocity. This corresponds to the hourly-mean wind velocity in the order of 110 km/h (i.e.,

about 160 km/h in terms of the 3-s gust mean wind velocity). It indicates that the tornado wind hazard dominates the synoptic wind hazard only for T at least greater than 10^4 years. It must be emphasized that this remark is only valid for a structure with a footprint represented by a point. For structures with footprints represented by a line or by an area, the tornado striking rate increases (Banik et al. 2008). This increase must be taken into account in comparing the wind velocity hazards due to tornado and synoptic winds.

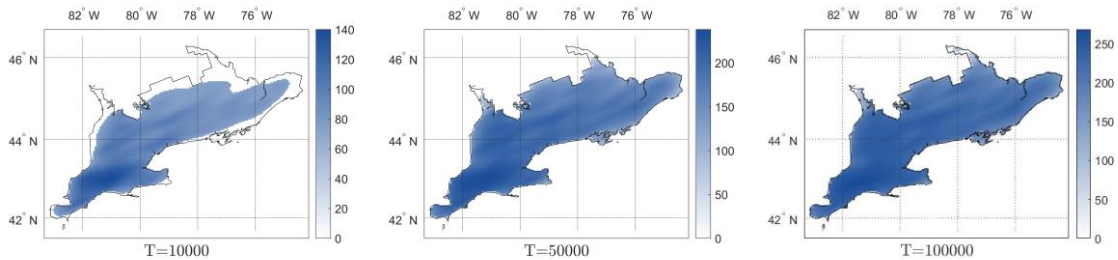


Figure 3.14. Tornado wind velocity, $V_T(10)$, (3-s gust wind velocity) map for southern Ontario.

The calculated model coefficients α_1 and α_2 for the wind profile defined by Eq. (3.9) are presented in Figures 3.15a and 3.15b, respectively, as functions of $V_T(10)$. The plots show that there are identifiable relations between α_1 and $V_T(10)$ and between α_2 and $V_T(10)$. α_1 attains the largest values for $V_T(10)$ within 115 to 135 km/h, implying a larger increase in $V_{TA}(z)$ as z increases. This increase is counteracted by the consideration of $\alpha_2 * z$, as shown in Eq. (3.9). Two simple parametric models are used to fit the samples shown in Figure 3.15,

$$\alpha_1(V_T(10)) = \begin{cases} -3.133 \times 10^{-2} + 1.137 \times 10^{-3} \times V_T(10), & 80 \leq V_T(10) \leq 125, \\ 0.204 - 7.424 \times 10^{-4} \times V_T(10), & 125 < V_T(10) \leq 225, \\ 0.0366, & 225 < V_T(10). \end{cases} \quad (3.10)$$

and,

$$\alpha_2(V_T(10)) = \begin{cases} 3.755 \times 10^{-5} + 5.622 \times 10^{-6} \times V_T(10), & 80 \leq V_T(10) \leq 140, \\ 1.636 \times 10^{-3} - 5.757 \times 10^{-6} \times V_T(10), & 140 < V_T(10) \leq 225, \\ 3.408 \times 10^{-4}, & 225 < V_T(10). \end{cases} \quad (3.11)$$

The use of α_1 and α_2 as well as $V_T(10)$, to define the tornado wind hazard largely simplifies

its potential implementation for codified structural design checking and tornado hazard and risk evaluation.

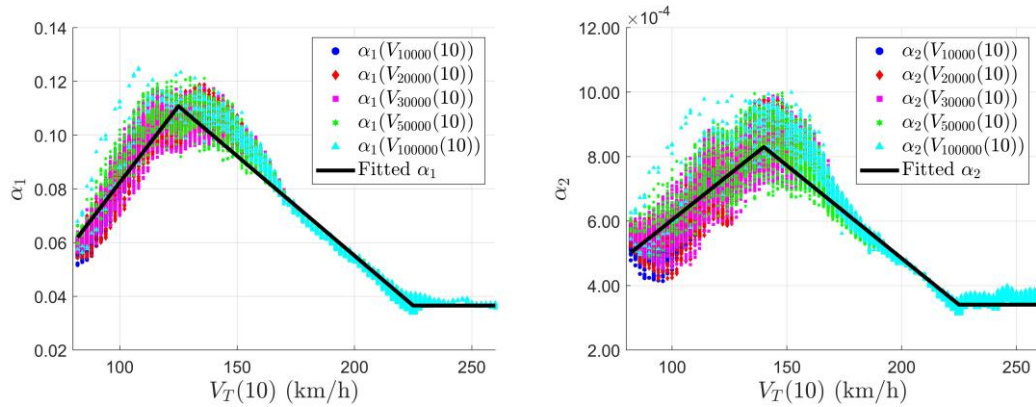


Figure 3.15. Variation of model parameters α_1 and α_2 as functions of $V_T(10)$.

3.5 Adequacy of wind profile for the line-like structure

The results presented in the previous section could be applied to a point structure at a height z . In the following, an assessment is carried out to investigate whether the developed wind profile could be used to provide an adequate estimate of quantiles of $M(z,H)$, $M_T(z,H)$, and quantile of $S(z,H)$, $S_T(z,H)$. Note that similar to estimate $V_T(z)$ for a site of interest, the estimation of $M_T(z,H)$ and $S_T(z,H)$ can be carried out by solving the exceedance distribution function of $M(z,H)$, $P(M(z,H) > m)$,

$$P(M(z,H) > m) \approx \sum_{j=0}^5 \gamma_{A_j} P(M(z,H) > m | F_{A_j}) \quad (3.12)$$

and the exceedance distribution function of $S(z,H)$, $P(S(z,H) > s)$,

$$P(S(z,H) > s) \approx \sum_{j=0}^5 \gamma_{A_j} P(S(z,H) > s | F_{A_j}). \quad (3.13)$$

Again, as an illustration, by carrying out the analysis, according to the steps shown in Figure 3.8, that is similar to that carried out for $V_T(z)$ presented in Figure 3.12 for London, Ontario, the obtained $M_T(z,H)$ and $S_T(z,H)$ for selected values of H and T are shown in Figure 3.16. Also, the bending moment and shear force are calculated by using the wind

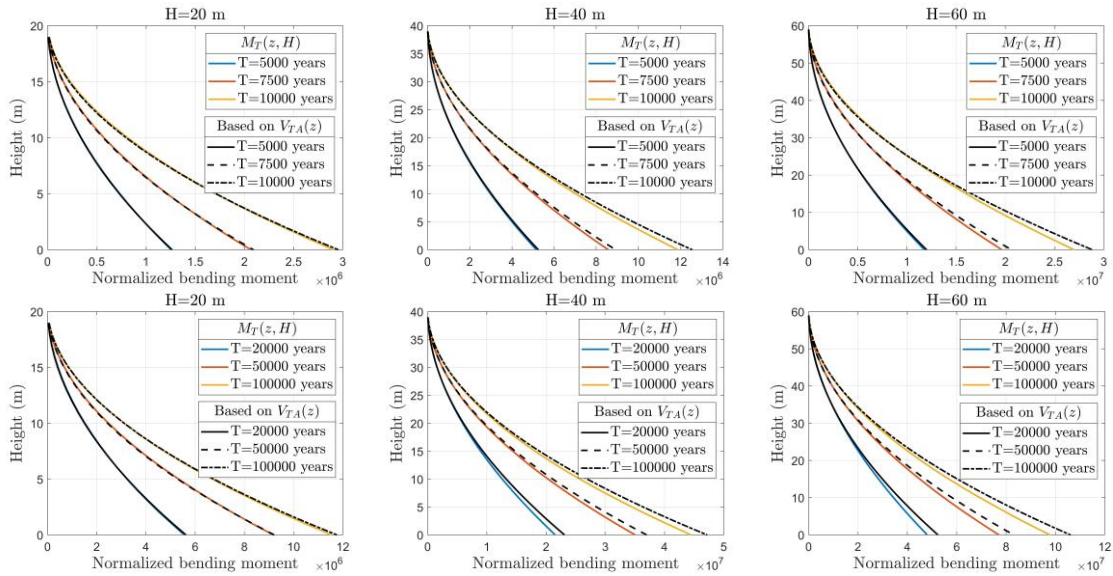
velocity $V_{TA}(z)$ (see Eq. (3.9)). The obtained results indicate that the calculated bending moment and shear force in such a manner overestimate $M_T(z,H)$ and $S_T(z,H)$ by about 20% to 35%. Therefore, it is suggested that the following wind profile is to be used for calculating the bending moment and shear force for line-like structure

$$V_{TA}(z) = R(V_T(10)) \times \begin{cases} V_T(10) & \text{for } z < 10 \text{ m} \\ V_T(10) \times \left(\frac{z}{10}\right)^{\max(0, \alpha_1 - \alpha_2 * z)} & \text{for } 10 \leq z < 80 \text{ m} \end{cases} \quad (3.14)$$

and

$$R(V_T(10)) = \min\{0.8 + 3.5 \times 10^{-4} \times V_T(10), 0.88\} \quad (3.15)$$

That is, a reduction factor of $R(V_T(10))$ is applied to Eq. (3.9) in defining the equivalent wind profile for the line-like structure in order to evaluate the bending moment and shear force. The calculated bending moment and shear force by using this wind profile are presented in Figure 3.16 and compared with $M_T(z,H)$ and $S_T(z,H)$. The comparison indicates that the use of the tornado wind profile shown in Eq. (3.14) to calculate the bending moment and shear force for the line-like structure is adequate, considering the uncertainties involved to assess the tornado wind velocity hazard.



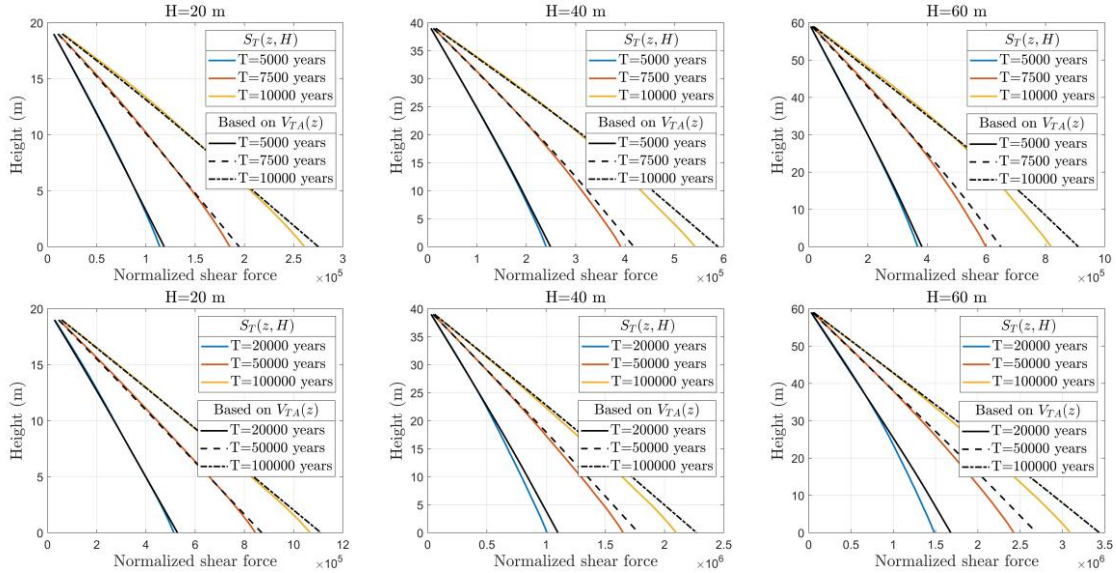


Figure 3.16. Estimated $M_T(z, H)$ and $S_T(z, H)$ for London Ontario.

3.6 Conclusions

A procedure for tornado occurrence modelling, wind hazard evaluation, disaggregation and mapping is presented. The procedure is applied to southern Ontario – a tornado-prone region in Canada. For the tornado occurrence modelling, the reporting bias due to population density and the use of the cloud-to-ground lightning flash density and thunderstorm days per year as the explanatory variables are considered. The statistical analysis results indicate that the negative binomial model is preferable to the Poisson model and the zero-inflated Poisson model for the modelling of tornado occurrence for southern Ontario.

The evaluated tornado wind hazard is used to develop along height tornado wind profile for specified return periods. It is shown that for a specified return period, the tornado hazard can be defined based on the T -year return period value of tornado wind speed at 10 m height, $V_T(10)$ and a simple parametric wind profile whose model parameters can be defined based on $V_T(10)$. This simplifies the potential implementation of the tornado wind velocity hazard for structural design checking and tornado hazard and risk evaluation. The disaggregation analysis results indicate that the identified scenario wind profiles could be

crudely bounded by using the rectangular wind profile and the synoptic wind profile, which are hinged at $V_T(10)$.

Also, it is shown that the proposed equivalent wind profile for the point-like structure could be used to evaluate the bending moment and shear force for the line-like structure if an additional reduction factor is considered (see Eq. (3.14) and (3.15)).

References

- Anderson, C.J., Wikle, C.K., Zhou, Q., & Royle, J.A. (2007). Population influences on tornado reports in the United States. *Weather and Forecasting*, 22(3), 571-579.
- Banik, S.S., Hong, H.P., & Kopp, G.A. (2007). Tornado hazard assessment for southern Ontario. *Canadian Journal of Civil Engineering*, 34(7), 830-842.
- Banik, S.S., Hong, H.P., & Kopp, G.A. (2008). Assessment of tornado hazard for spatially distributed systems in southern Ontario. *Journal of wind engineering and industrial aerodynamics*, 96(8-9), 1376-1389.
- Bazzurro, P., & Cornell, A. (1999). Disaggregation of seismic hazard. *Bulletin of the Seismological Society of America*, 89(2), 501-520.
- Besag, J., & Kooperberg, C. (1995). On conditional and intrinsic autoregressions. *Biometrika*, 82(4), 733-746.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2), 192-225.
- Bissolli, P., Grieser, J., Dotzek, N., & Welsch, M. (2007). Tornadoes in Germany 1950–2003 and their relation to particular weather conditions. *Global and Planetary Change*, 57(1-2), 124-138.

- Branick, M.L., & Doswell III, C.A. (1992). An observation of the relationship between supercell structure and lightning ground-strike polarity. *Weather and Forecasting*, 7(1), 143-149.
- Brooks, S., Gelman, A., Jones, G., & Meng, X. L. (Eds.). (2011). *Handbook of Markov chain Monte Carlo*. CRC press.
- Burnham, K., & Anderson, D. (2002). *Model Selection and Multi-model Inference*. 2nd edn. Springer, New York.
- Burrows, W.R., & Kochtubajda, B. (2010). A decade of cloud - to - ground lightning in Canada: 1999 - 2008. Part 1: Flash density and occurrence. *Atmosphere-ocean*, 48(3), 177-194.
- Carey, L.D., & Buffalo, K.M. (2007). Environmental control of cloud-to-ground lightning polarity in severe storms. *Monthly Weather Review*, 135(4), 1327-1353.
- Cheng, V.Y., Arhonditsis, G.B., Sills, D.M., Auld, H., Shephard, M.W., Gough, W.A., & Klaassen, J. (2013). Probability of tornado occurrence across Canada. *Journal of Climate*, 26(23), 9415-9428.
- Cheng, V.Y., Arhonditsis, G.B., Sills, D.M., Gough, W.A., & Auld, H. (2016). Predicting the climatology of tornado occurrences in North America with a Bayesian hierarchical modelling framework. *Journal of Climate*, 29(5), 1899-1917.
- Davenport, A.G. (1962). The response of slender, line-like structures to a gusty wind, *Proc. Inst. Civil Eng.* 23, 369-408.
- Dunn, W.L., & Twisdale, L.A. (1979). A synthesized windfield model for tornado missile transport. *Nuclear Engineering and Design*, 52(1), 135-144.
- Elsner, J. B., & Widen, H.M. (2014). Predicting spring tornado activity in the central Great Plains by 1 March. *Monthly Weather Review*, 142(1), 259-267.

- Gilks, W.R., & Richardson, S.S. and Spiegelhalter, D. (1996). Markov chain Monte Carlo in practice. London, UK: Chapman k Hall/CRC.
- Hangan, H., & Kim, J.D. (2008). Swirl ratio effects on tornado vortices in relation to the Fujita scale. *Wind and Structures*, 11(4), 291-302.
- Honerkamp, R., Yan, G., & Snyder, J.C. (2020). A review of the characteristics of tornadic wind fields through observations and simulations. *Journal of Wind Engineering and Industrial Aerodynamics*, 202, 104195.
- Hong, H.P., & Goda, K. (2006). A comparison of seismic-hazard and risk deaggregation. *Bulletin of the Seismological Society of America*, 96(6), 2021-2039.
- Huang, Q., Jiang, W.J., & Hong, H.P. (2020) Statistical assessment of spatial Tornado occurrence for Canada: modelling and estimation, Report to NRC.
- Jagger, T.H., Elsner, J. B., & Widen, H.M. (2015). A statistical model for regional tornado climate studies. *PloS one*, 10(8).
- Johnston, K., Ver Hoef, J.M., Krivoruchko, K., & Lucas, N. (2003). ArcGIS9, Using ArcGIS geostatistical analyst, Redlands, CA: Environmental Systems Research Institute.
- King, P.W., Leduc, M.J., Sills, D.M., Donaldson, N.R., Hudak, D.R., Joe, P., & Murphy, B.P. (2003). Lake breezes in southern Ontario and their relation to tornado climatology. *Weather and Forecasting*, 18(5), 795-807.
- King, P. (1997). On the absence of population bias in the tornado climatology of southwestern Ontario. *Weather and Forecasting*, 12(4), 939-946.
- Knapp, D.I. (1994). Using cloud-to-ground lightning data to identify tornadic thunderstorm signatures and nowcast severe weather. *Natl. Wea. Dig*, 19(2), 35-42.
- Kuo, H. L. (1971). Axisymmetric flows in the boundary layer of a maintained vortex. *Journal of the Atmospheric Sciences*, 28(1), 20-41.

- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1), 1-14.
- Marshall, T.P., McDonald, J. R., & Forbes, G. S. (2004, October). The enhanced Fujita (EF) scale. In *Preprints, 22nd Conf. on Severe Local Storms, Hyannis, MA, Amer. Meteor. Soc. B* (Vol. 3).
- NRC. 2015. National Building Code of Canada. Institute for Research in Construction, National Research Council of Canada, Ottawa, Ontario
- Newark, M.J. (1984). Canadian tornadoes, 1950–1979. *Atmosphere-Ocean*, 22(3), 343-353.
- Newark, M.J. (1991). A design basis tornado. *Canadian Journal of Civil Engineering*, 18(3), 521-524.
- Reap, R.M., & MacGorman, D.R. (1989). Cloud-to-ground lightning: Climatological characteristics and relationships to model fields, radar observations, and severe local storms. *Monthly Weather Review*, 117(3), 518-535.
- Refan, M., & Hangan, H. (2016). Characterization of tornado-like flow fields in a new model scale wind testing chamber. *Journal of Wind Engineering and Industrial Aerodynamics*, 151, 107-121.
- Roueche, D.B., Prevatt, D.O., & Haan, F.L. (2020). Tornado-induced and straight-line wind loads on a low-rise building with consideration of internal pressure. *Frontiers in Built Environment*, 6, 18.
- Shephard, M.W., Morris, R., Burrows, W.R., & Welsh, L. (2013). A high-resolution Canadian lightning climatology. *Atmosphere-Ocean*, 51(1), 50-59.
- Silverman, B.W. (1986). *Density estimation for statistics and data analysis* (Vol. 26). CRC press.

- Simiu, E., & Scanlan, R. H. (1996). Wind effects on structures: fundamentals and applications to design, third edition, Wiley-Interscience.
- Tan, L., & Hong, H.P. (2010). Influence of spatial inhomogeneity of tornado occurrence on estimated tornado hazard. *Canadian Journal of Civil Engineering*, 37(2), 279-289.
- Twisdale, L.A. (1978). Tornado data characterization and windspeed risk. *Journal of the Structural Division*, 104(10), 1611-1630.
- Twisdale, L.A., Dunn, W.L., Davis, T., & Horie, Y. (1981). Tornado missile simulation and design methodology. EPRI NP-2005.
- Twisdale, L.A., & Dunn, W.L. (1983). Probabilistic analysis of tornado wind risks. *Journal of Structural Engineering*, 109(2), 468-488.
- Wen, Y.K. (1975). Dynamic tornadic wind loads on tall buildings. *Journal of the Structural Division*, 101(1), 169-185.
- Wen, Y.K., & Chu, S.L. (1973). Tornado risks and design wind speed. *Journal of the Structural Division*, 99(12), 2409-2421.
- Wikle, C.K., & Anderson, C.J. (2003). Climatological analysis of tornado report counts using a hierarchical Bayesian spatiotemporal model. *Journal of Geophysical Research: Atmospheres*, 108(D24).
- Wurman, J., & Gill, S. (2000). Finescale radar observations of the Dimmitt, Texas (2 June 1995), tornado. *Monthly Weather Review*, 128(7), 2135-2164.
- Yarbrough Jr, J.W., & Meentemeyer, V. (1978). Seasonal and regional variation in the correlation of thunderstorm days with tornado frequency. *Journal of Applied Meteorology*, 17(11), 1741-1746.

Chapter 4

4 Tornado wind hazard mapping and tornado design wind profile for Canada

4.1 Introduction

A number of tornadoes are reported each year in Canada. Newark (1984, 1991) provided the first systematic Canadian tornado catalogue and a quantitative tornado hazard assessment for Canada. It was indicated that tornadoes of F4 on the Fujita scale have occurred in many regions of east of the Rocky Mountains and that there is little information available to assess the intensity of tornadoes that occurred in British Columbia. A tornado catalogue was released by the Environment and Climate Change Canada (ECCC) (<https://open.canada.ca/data/en/dataset/fd3355a7-ae34-4df7-b477-07306182db69>) (last accessed January 21, 2020). This catalogue, which is referred to as EC-T80-09 in the following, includes the tornadoes that occurred from 1980 to 2009. The rationale for including tornadoes that only occurred after 1979 is not given. Perhaps, it is due to that the information on the tornadoes that was archived after the development of Canada's regional severe weather offices around 1980 (Etkin et al. 2001) could have better quality control. A comparison of the EC-T80-09 with other available tornado catalogues for different regions (Sills et al. 2004; Patrick McCarthy (2012, personal communication)) indicates that EC-T80-09 contains events classified as confirmed and probable events. This observation is further evidenced by the comments made in Sills et al. (2012) (to their Figure 1). An inspection of EC-T80-09 indicates that the annual average reported (confirmed or probable) tornadoes in Canada is about 62. It is believed that the number of reported tornadoes is likely to be less than the actual number of occurrences (Newark 1984, 1991; King 1997) since the population density in a large area in Canada is less than five or six (people/km²). Valuable tornado damage surveys and damage analyses for some recent events were presented in Kopp et al. (2017).

The statistical assessment of spatial inhomogeneity of tornado occurrence in southern Ontario was presented in Tan (2008) by applying the Bayesian hierarchical modelling technique (Wikle and Anderson 2003), and in Tan and Hong (2010) by applying the

adaptive Gaussian kernel smoothing (GKS) technique (Silverman 1986). This inhomogeneity could be attributed to the lake breeze effect resulting in a preferred southwest-to-northeast pattern of tornadic activity (King et al. 2003). By using EC-T80-09, Cheng et al. (2013, 2016) investigated the tornado occurrence rate for Canada based on the Bayesian hierarchical modelling framework. For their analysis, the tornado occurrence is modeled by using the Poisson model and the zero-inflated Poisson (ZIP) model (Lambert 1992). Moreover, they considered the reporting bias due to population density and used the annual average cloud-to-ground lightning flash (ACGLF) density as an explanatory variable to estimate the tornado occurrence rate. The consideration of ACGLF as the explanatory variable is justified as the flash cloud-to-ground (CG) lightning density is positively correlated with the tornado occurrence (Reap and MacGorman 1989). Additional studies on the relation of the CG lightning and tornado characteristics include those given by Branick and Doswell (1992), Knapp (1994), and MacGorman and Burgess (1994), Perez et al. (1997), and Carey and Buffalo (2007).

Instead of the Poisson and ZIP models, Elsner and Widen (2014) and Jagger et al. (2015) applied the negative binomial (NB) model to represent the tornado occurrence. Both the NB model and ZIP model can cope with overdispersion. By using EC-T80-09, Huang et al. (2020) compared the performance of the Poisson, ZIP, and NB models to represent the tornado occurrence in Canada. Based on the Akaike information criterion (Akaike 1974), their statistical analysis results indicate that the NB model is preferable to the Poisson and ZIP models. For their analysis, they used ACGLF density as well as the annual average thunderstorm days (ATD) as the explanatory variables. The consideration of ATD as the explanatory variable is justified since the thunderstorm days and tornado occurrence are correlated (Yarbrough and Meentemeyer, 1978; Bissolli et al., 2007). The method of maximum likelihood and the Bayesian hierarchical modelling technique were employed to estimate the model parameters. It was observed that the model parameters estimated by using the two approaches are almost identical. This was attributed to the use of the “non-informative” prior and can be explained based on the relation between the maximum likelihood estimate and the maximum a posteriori probability estimate (Robert 2007; Lee 2012). This suggests that one can use the method of maximum likelihood to estimate model parameters, which simplifies the model development as compared to the use of the

Bayesian hierarchical modelling technique.

It should be noted that the models developed in Cheng et al. (2013, 2016) and Huang et al. (2020) used EC-T80-2009, which does not include tornado events that occurred from 2010 to 2019. Therefore, the potential benefit that may be gained by using additional recently reported tornadoes is unexploited.

A spatially varying tornado-prone map for Canada is implemented in the commentary to the National Building Code of Canada (NRC 2015). This map provided an overview of tornado hazard in terms of tornado occurrence rate. However, it cannot be used directly for checking and evaluating structural performance subjected to the tornado wind velocity hazard since no guideline on the return period value of tornado wind velocity is given. A comparison of the wind loads on the low-rise building due to tornadic wind and synoptic wind indicates that the magnitudes of the peak tornado-induced pressures are reasonably similar to straight-line wind-induced pressures (Roueche et al. 2020). This, at least, shows the feasibility of implementing tornado design wind pressure for a given tornado wind velocity.

Similar to the design for the synoptic winds, which is based on the return period value of wind velocity (e.g., Hong et al. 2014), the tornado wind velocity hazard should be estimated to provide a probabilistic basis for tornado wind design. The hazard can be evaluated by combining two probabilistic modules. One deals with the modelling of the spatially varying tornado occurrence and striking rate. The other describes the probabilistic tornado wind velocity field. The use of these modules to evaluate the return period value of tornado wind velocity at 10 m height above the ground surface was illustrated for southern Ontario in Banik et al. (2007, 2008) and Tan and Hong (2010). For the estimation, they adopted the probabilistic wind field model given in Twisdale et al. (1981). It is noteworthy that there are other wind field models available in the literature (Wen 1975; Twisdale 1978; Wurman and Gill 2000; Hangan and Kim 2008; Refan and Hangan 2016). Unfortunately, the probabilistic predicting equations of the model parameters for these models are unavailable.

Unlike the synoptic winds, there is no unique along height wind profile that could be used

for tornadoes of different intensity. This complicates the tornado wind hazard assessment since the estimation of the return period of the tornado wind velocity needs to be carried out at different heights above the ground surface. The consideration of the design of line-like structures subjected to the tornado winds further complicates the tornado wind velocity hazard assessment since, in such a case, an equivalent design tornado wind profile needs to be developed to simplify possible structural design code implementation. However, no along height design wind profile for horizontal tornado wind velocity was suggested in the literature. A conceptual development with very preliminary results focused on establishing an equivalent tornado design wind profile for southern Ontario was outlined in Hong and Huang (2020). This conceptual development is extended and explored in the present study for Canada that has varying tornado activities from region to region.

The main objectives of the present study are to a) model and map the spatially varying tornado occurrence for Canada, b) to map the tornado striking rate by considering the tornado path characteristics, c) to estimate quantiles of the tornado horizontal wind velocity and to map the wind velocity hazard at 10 m height above the ground surface, and d) to develop a simple equivalent tornado design wind profile by considering the tornado-induced bending moment and shear force for a line-like structure. In the next section, we first provide some details on the data and the procedure used to model the tornado occurrence rate. In Section 3, we described the assessed tornado striking rate and the probability distributions of tornado wind velocity for a point-like structure, and the bending moment and shear force for a line-like structure. The mapped tornado wind velocity hazard and the developed equivalent tornado design wind profile for point-like structure and line-like structure are then presented in Section 4. Finally, the concluding remarks and recommendations are presented in Section 5.

4.2 Developing tornado occurrence rate model

4.2.1 Tornado catalogue

As mentioned in the introduction, Newark (1984) established the first systematic Canadian tornado catalogue. The shortcomings of Newark's tornado catalogue include the

population bias. The most recently released EC-T80-09 can be downloaded from <https://open.canada.ca/data/en/dataset/fd3355a7-ae34-4df7-b477-07306182db69> (last accessed January 21, 2020). It includes 1839 tornado events that occurred from 1980 to 2009. For each event in the catalogue, Fujita-scale (i.e., F-scale), touchdown latitude-longitude, length, and motion direction are given if they are available. EC-T80-09 contains events classified as confirmed and probable events, as mentioned in the introduction. However, no identifier was given in the catalogue if an event belongs to confirmed or probable categories. The spatial distribution of the 1839 tornadoes events is shown in Figure 4.1a. The number of events according to each Fujita-scale, denoted as F_i , $i = 0, \dots, 5$, is also included in the figure. Furthermore, the estimated probability mass function for the tornado intensity, $P(F_i)$, is calculated and presented in the figure. For the calculation, the detrending, and the error and bias correction used in Banik et al. (2007) and Tan and Hong (2010) are employed. The figure shows that the intensity is smaller than F_2 for about 80% of events, and smaller than F_3 for about 94% of events. The plot confirms that substantial tornado activities occurred in the southwestern Ontario and Prairie provinces, which is consistent with the observation made by Newark (1984, 1991).

EC-T80-09 does not include tornado events that occurred from 2010 to 2019. An effort is made for the present study to update this tornado catalogue using the information gathered from different sources. It was noted that the internet sites within Wikipedia (https://en.wikipedia.org/wiki/List_of_Canadian_tornadoes_and_tornado_outbreaks#cite_ref-ehfs_1-0; and https://en.wikipedia.org/wiki/List_of_21st-century_Canadian_tornadoes_and_tornado_outbreaks) provide a well-organized account of tornadoes occurred in Canada in chronological order. Also, the tornado catalogue for Prairie and Northern region in Canada up to 2010, which is compiled by Patrick McCarthy (2012, personal communication), could be used to identify tornado events in 2010. By combining the EC-T80-09 and the information extracted from the internet sites and McCarthy's catalogue for tornadoes that occurred from 2010 to 2019, a new tornado catalogue is developed. This catalogue is referred to as T80-19 for simplicity in the following. In developing T80-19, it is noted that the reporting of the tornado intensity in Canada was changed from F-scale to enhanced Fujita scale (EF-scale) (Marshall et al. 2004) since 2013 and that the tornado intensities for some tornado events from 2010 to

2019 were not reported. The spatial distribution of the reported tornado events in T80-19 is presented in Figure 4.1b.

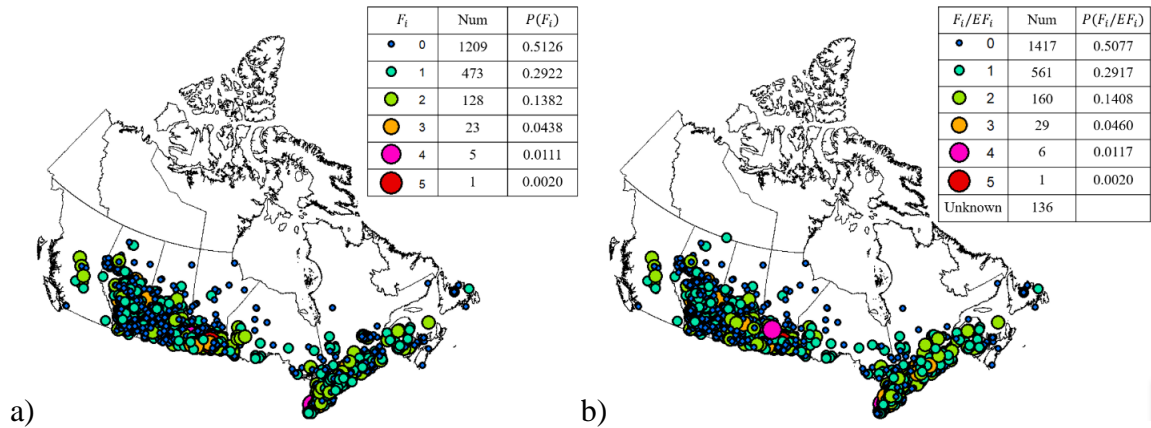


Figure 4.1. Spatial distribution of the reported tornado event: a) Reported events in EC-T80-09, and b) Reported events in T80-19.

A visual inspection of the results presented in Figures 4.1a and 4.1b indicates that the spatial trends in both plots are similar. The data show again that the intensity is smaller than F_2 for about 80% of the reported tornadoes, and smaller than F_3 for about 94% of the reported tornadoes. Furthermore, the annual tornado occurrence rate for Canada is 61.3 by considering events from 1980 to 2009 and 42.8 by considering events from 2010 to 2019. One potential explanation of the discrepancy between these rates is that the catalogue from 2010 to 2019 contains events that could all be classified as confirmed events while the catalogue from 1980 to 2009 includes confirmed and probable events. It is also possible that the catalogue from 2010 to 2019 is incomplete since the database is not issued by a weather agency office. The annual occurrence rate for Canada by using T80-19 equals 57.8, which is less than 61.3 obtained based on EC-T80-19.

There are differences between $P(F_i)$ developed based on EC-T80-09 and T80-19, as shown in Figure 4.1. As the tornado intensities for 136 of the events that occurred from 2010 to 2019 are unavailable, and the assumption of F_i equal to EF_i to assess $P(F_i)$ may be inadequate, $P(F_i)$ calculated based on EC-T80-09 is used in the following.

To better appreciate the differences in the spatial trends of the tornado occurrence rate based on the reported events, a calculation of the rate is carried out by applying GKS, as was done in Tan and Hong (2008) but for southern Ontario. The obtained results are presented in Figure 4.2, where the grid system shown in Figure 4.2, with each cell defined approximately by $30 \times 30 \text{ km}^2$, is used throughout the present study. It must be emphasized that possible reporting bias due to population density is not considered for the rate shown in Figure 4.2. Moreover, the tornado occurrence rate is not equal to the rate of the tornado striking a site of interest. This is because each tornado is associated with the tornado width, length, and track orientation, which will be discussed shortly. The figure indicates that the occurrence rates for southwestern Ontario and the southern border of Saskatchewan and Manitoba with the United States are much larger than that for other sites. The highest occurrence rate is about 2.3×10^{-4} (per year per km^2) based on EC-T80-09; it becomes 2.1×10^{-4} (per year per km^2) if T80-19 is employed.

4.2.2 Population density, cloud-to-ground lightning flash density and thunderstorm days

As mentioned in the introduction, it has been observed that the tornado reporting is biased by population density (Newark 1984; King 1997), and that ACGLF and ATD could be used as the explanatory variables for developing the tornado occurrence rate model.

Two sets of population data, one from 2001 and the other from 2011 Canadian census data (<https://www12.statcan.gc.ca/english/census01/oproducts/standard/popdwell/tables.cfm>) are used in the present study. Since the census data are given for each census subdivision, the density is estimated by assuming that the density within each census subdivision is uniform and then allocated in the considered grid cells, as shown in Figure 4.3, illustrating that the population is concentrated near the southern border.

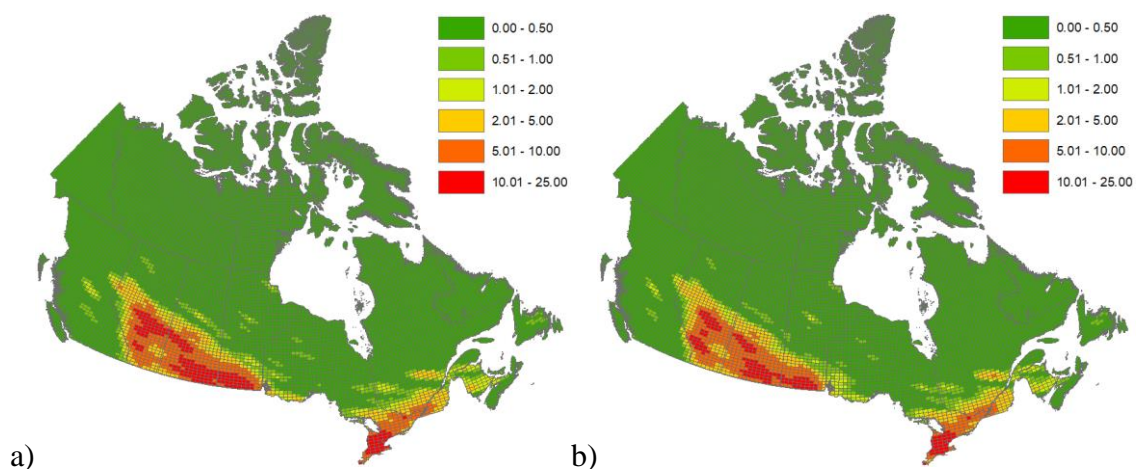


Figure 4.2. Spatially smoothed tornado occurrence rate (per $(\text{year} \times 10^5 \text{ km}^2)$) based on the tornado catalogues for the reported tornado events with locations depicted in Figure 4.1:

- a) Based on reported events depicted in Figure 4.1a, and b) Based on reported events depicted in Figure 4.1b.

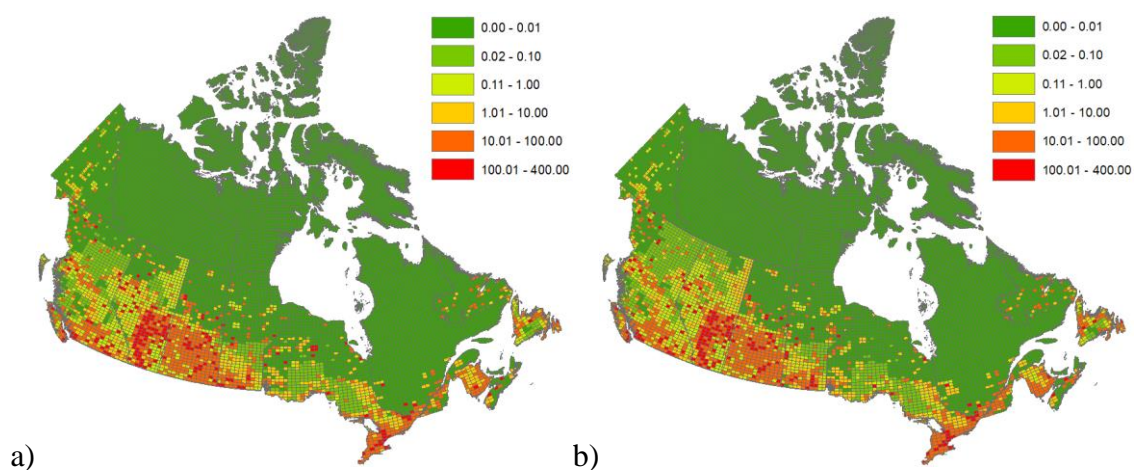


Figure 4.3. Spatially varying population density (number of people per km^2): a) Based on 2001 Canadian census data, and b) based on 2011 Canadian census data.

The CG lightning data from 1999 to 2009 that are available to the present study are obtained from the Canadian Lightning Detection Network. Details on the CG lightning data are described in Burrows and Kochtubajda (2010) and Shephard et al. (2013). The calculated ACGLF density is shown in Figure 4.4. An inspection of the plots presented in Figures 4.1

and 4.4 indicates that there are similarities in the spatial trends in the tornado occurrence and the ACGLF density. In particular, the ACGLF density in southwestern Ontario and the southern regions of Prairies provinces is greater than that in other regions.

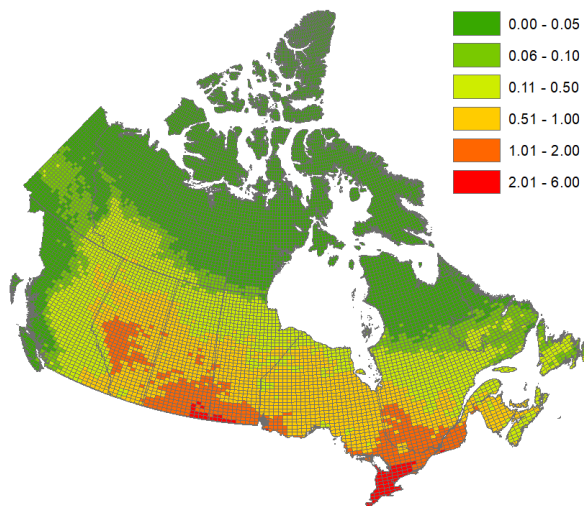


Figure 4.4. Estimated ACGLF density (per year per km²) using data from 1999 to 2009.

The records of the climatological elements in ECCC DLY04 digital archive (see http://www.climate.weatheroffice.gc.ca/prods_servs/documentation_index_e.html#hly01) are obtained and used to identify the thunderstorm days. In this database, an identifier for the daily maximum wind velocity recorded at each meteorological station is given. The identifier indicates if the recorded wind is caused by a thunderstorm. Stations in the database with at least 10 years of useable data are considered so to reduce statistical uncertainty and to have sufficient samples, and ATD is estimated. The stations, as well as the spatially interpolated values of the estimated ATD, are shown in Figure 4.5a by using the records from 1980 to 2009, and in Figure 4.5b by using the records from 1980 to 2017 (data from 2018 to 2019 is unavailable for the present study). The plotted ATD is obtained by spatially interpolating ATD obtained at the stations, where the ordinary kriging with nugget equal to zero (Johnston et al. 2003) is employed for the interpolation. A comparison of the results presented in Figures 4.5a and 4.5b indicates that the spatial trends of ATD are similar but not exactly the same due to statistical variability. A comparison of the reported historical tornadoes plotted in Figure 4.1, and the ATD shown in Figure 4.5 indicates that the spatial patterns of ATD and the tornado occurrence are similar,

suggesting that the ATD could be used as an explanatory variable for the tornado occurrence modelling.

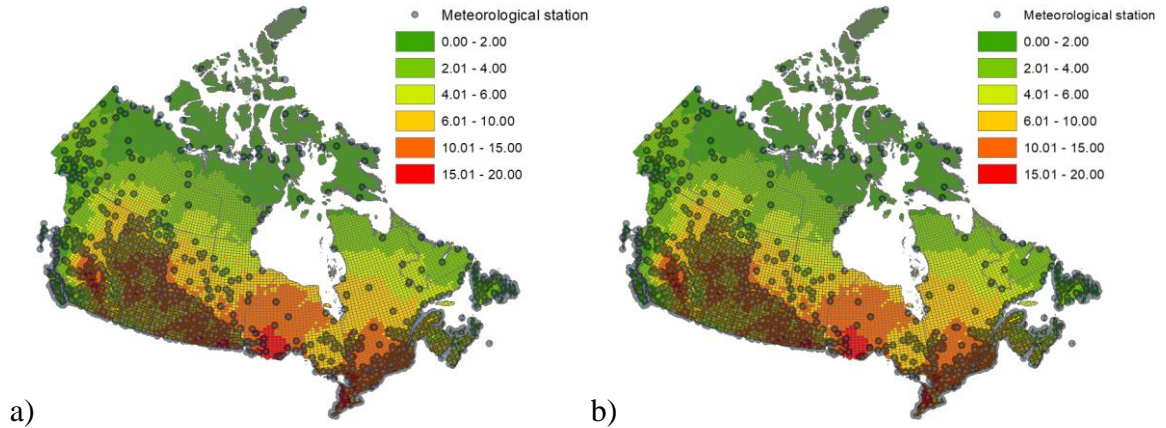


Figure 4.5. Estimated ATD (number of days/year): a) using records from 1980 to 2009 and b) using records from 1980 to 2017.

4.2.3 Stochastic modelling for tornado occurrence rate

In this section, three models used to model the spatial tornado occurrence are summarized by considering the possible under-reporting caused by population density, and the potential explanatory variables such as those shown in Figures 4.4 and 4.5. It is considered that the reported tornado occurrence in the j -th cell for the observation period of t years, Y_{tj} , is modeled using,

$$Y_{tj} = \sum_{i=1}^{N_{tj}} X_{tji}, \quad (4.1)$$

where N_{tj} is the actual number of the tornado that occurred within the period of t years, X_{tji} is independent and identically distributed random variable according to the Bernoulli distribution for the considered duration and fixed j . The parameter for the Bernoulli distribution representing the probability of reporting an occurred event within the j -th cell (i.e., $X_{tji} = 1$) is denoted by p_{tj} . The probability of no reporting (i.e., $X_{tji} = 0$) equals $1 - p_{tj}$. Since the reporting probability is considered to be a function of population density (Newark

1984; King 1997; Brooks et al. 2003; Anderson et al. 2007; Cheng et al. 2013), p_{ij} is a function of population density.

N_{ij} is frequently modeled by using a Poisson model (Wen 1975; Twisdale and Dunn 1983; Banik et al. 2007; Tan and Hong 2010; Cheng et al. 2013), the ZIP model (Lambert 1992; Wilke and Anderson 2003; Cheng et al. 2016) or the NB model (Elsner and Widen 2014; Jagger et al. 2015). The mathematical formulations of the models are summarized in Table 4.1.

The Poisson model is controlled by the occurrence rate for the j -th cell, denoted as λ_j (per year), which is modeled using (Wikle and Anderson 2003, Anderson et al. 2007 and Cheng et al. 2013),

$$\ln \lambda_j = \alpha_0 + \sum_{k=1}^{n_K} \alpha_k z_{kj}, \quad (4.2)$$

where α_k for $k = 0, \dots, n_K$, are model parameters, and n_K is the number of explanatory variables z_{kj} . As mentioned in the introduction, the use of ACGLF density alone as the explanatory variable was considered by Cheng et al. (2013). However, they did not consider the use of ATD or both the ACGLF density and ATD as the explanatory variables.

The ZIP model is controlled by λ_j and an additional zero tornado occurrence probability π_{ij} , and the NB model is controlled by the probability of non-occurrence denoted as p_{NB} and a positive model parameter r . These models are well-known and explained in Daley and Vere-Jones (2007) and Lambert (1992). It is considered that (Wikle and Anderson 2003),

$$p_{ij} = \exp(-\beta / \gamma_j), \quad (4.3)$$

and

$$\pi_{ij} = \exp(\xi) / (1 + \exp(\xi)), \quad (4.4)$$

where γ_j is the population density (number of people/km²) for the j -th cell and ξ is a model

parameter.

The maximum likelihood function \mathbf{L} for these three models based on the generalized linear model (which differs from the general linear model or the generalized least-squares method) and for given observed Y_{jt} is summarized in Table 4.1 (see also Huang et al. (2020)). In writing \mathbf{L} shown in Table 4.1, the fact that the occurrence rate for the reported tornado Y_{jt} in the j -th cell per unit time, $\eta_j(\boldsymbol{\theta})$, equals $p_{ij}\lambda_j$ is used, where $\boldsymbol{\theta}$ denotes a set of parameters used for the parametric model for p_{ij} , π_{ij} , and λ_j . Moreover, it is considered that the combination of the NB model for N_{ij} and Eq. (4.1) leads to Y_{ij} to be represented by the NB model but with P_{NBj} replaced by a new model parameter P_{NBBj} , that equals $r / (r + \eta_j(\boldsymbol{\theta}))$.

Given the reported tornado occurrence, the population density, ACGLF density, and ATD, the model parameters in $\boldsymbol{\theta}$ can be estimated by maximizing \mathbf{L} . To select a preferred model among all the considered model, one could consider the application of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) (Akaike 1974; Schwarz 1978; Burnham and Anderson 2002; Heinze et al. 2018) defined as,

$$\text{AIC} = 2p - 2 \ln \hat{\mathbf{L}} \quad (4.5)$$

and,

$$\text{BIC} = p \ln(n) - 2 \ln \hat{\mathbf{L}} \quad (4.6)$$

where p is the total number of model parameters to be estimated, n is the sample size, and $\hat{\mathbf{L}}$ is the maximized value of \mathbf{L} . The model with the lowest AIC (or BIC) is the preferred model.

Table 4.1. Models considered and the corresponding likelihood function

Model	Probability Distribution, $p(N_{ij} = n)$	Likelihood function, \mathbf{L}
-------	--	--------------------------------------

Poisson	$\frac{(\lambda_{ij}t)^n e^{-\lambda_{ij}t}}{n!}$	$\prod_{j=1}^{M_c} \frac{(\eta_j(\boldsymbol{\theta})t)^{y_j}}{y_j!} e^{-\eta_j(\boldsymbol{\theta})t}$
ZIP	$\begin{cases} \frac{(\lambda_{ij}t)^n e^{-\lambda_{ij}t}}{n!} & \text{with probability } 1-\pi_{ij} \\ 0 & \text{with probability } \pi_{ij} \end{cases}$	$\left[\frac{1}{1+\exp(\xi)} \right]^{M_c} \times \prod_{j=1}^{M_c} \left[\exp(\xi) I_{\{y_j=0\}}(y_j) + \frac{(\eta_j(\boldsymbol{\theta})t)^{y_j}}{y_j!} e^{-\eta_j(\boldsymbol{\theta})t} \right]$
NB	$\frac{\Gamma(n+rt)}{n!\Gamma(rt)} p_{NBj}^{r,t} (1-p_{NBj})^n$	$\prod_{j=1}^{M_c} \left[\frac{\Gamma(y_j+rt)}{y_j!\Gamma(rt)} \left(\frac{r}{r+\eta_j(\boldsymbol{\theta})} \right)^{rt} \left(\frac{\eta_j(\boldsymbol{\theta})}{r+\eta_j(\boldsymbol{\theta})} \right)^{y_j} \right]$

4.2.4 Estimating model parameters to predict tornado occurrence rate

Based on the data presented in the previous sections alone, it is difficult to recommend a single tornado catalogue for assessing tornado occurrence rate. This is because the catalogue from 1980 to 2009 contains events classified as confirmed or probable, and the duration of the catalogue is only 30 years. The catalogue from 1980 to 2019 is longer, but the data quality control cannot be verified. Therefore, as a parametric investigation, we considered both catalogues in the following.

Also, as part of parametric analysis, we consider four tornado occurrence models: the three parametric models (namely, the Poisson, ZIP, and NB models), and the non-parametric model (i.e., GKS). For the parametric models, we consider ACGLF density alone, ATD alone, and both the ACGLF density and ATD as the explanatory variables. The combinations of the catalogues, methods, and explanatory variables result in 20 cases, as listed in Table 4.2.

The estimated spatial varying tornado occurrence rate corresponding to Cases A1 and B1 by using GKS are already presented in Figures 4.2a and 4.2b, respectively. As mentioned earlier, in general, the rate shown in Figure 4.2a is greater than that shown in Figure 4.2b because of the differences in the catalogue as mentioned earlier.

Table 4.2. Combinations of cases to be considered for the statistical analysis of tornado occurrence rate (z_{1j} is used to represent ACGLF density and z_{2j} is used to represent ATD).

Case	Tornado catalogue	Model	Covariates	Case	Tornado catalogue	Model	Covariates
A1	EC-T80-09	GKS		B1	T80-19	GKS	
A2		Poisson model	ACGLF	B2		Poisson model	ACGLF
A3			ATD (Figure 4.5a)	B3			ATD (Figure 4.5b)
A4			ACGLF & ATD (Figure 4.5a)	B4			ACGLF & ATD (Figure 4.5b)
A5		ZIP model	ACGLF	B5		ZIP model	ACGLF
A6			ATD (Figure 4.5a)	B6			ATD
A7			ACGLF & ATD (Figure 4.5a)	B7			ACGLF & ATD (Figure 4.5b)
A8		NB model	ACGLF	B8		NB model	ACGLF
A9			ATD (Figure 4.5a)	B9			ATD (Figure 4.5b)
A10			ACGLF & ATD (Figure 4.5a)	B10			ACGLF & ATD (Figure 4.5b)

We apply the method of maximum likelihood (MLM) for Cases A2 to A10, and B2 to B10 listed in Table 4.2 with the likelihood function L given in Table 4.1. The application of MLM to estimate the model parameters is implemented in **R** programming language (R Core Team 2016) using the built-in function “glm” (i.e., generalized linear model). Using the estimated model parameters and the corresponding \hat{L} , the obtained AIC and BIC are presented in Table 4.3. Since a model with the lowest value of AIC or BIC is preferred among all the considered models, the table shows that the preferred model judged based on either AIC or BIC is the NB model with both the ACGLF and ATD as the explanatory variables. In general, the model with the ACGLF density and ATD as the explanatory variables outperforms the models with the ACGLF density alone or ATD alone as the explanatory variable. The NB model is preferable to the ZIP and Poisson models, while the Poisson model is least preferred. This is the case if EC-T80-09 or T80-19 is used. Also, the use of ACGLF and ATD as the explanatory variables is preferable to the use of ACGLF or ATD as the explanatory variable if the NB model is considered.

Table 4.3. Calculated AIC and BIC for the cases described in Table 4.2.

Case	AIC	BIC	Case	AIC	BIC
------	-----	-----	------	-----	-----

A2	6912	6934	B2	8112	8135
A3	7324	7346	B3	8667	8689
A4	6689	6718	B4	7793	7823
A5	6230	6260	B5	7186	7215
A6	6174	6203	B6	7164	7194
A7	6043	6080	B7	6963	7000
A8	5359	5382	B8	5927	5949
A9	5659	5681	B9	6244	6266
A10	5326	5355	B10	5886	5916

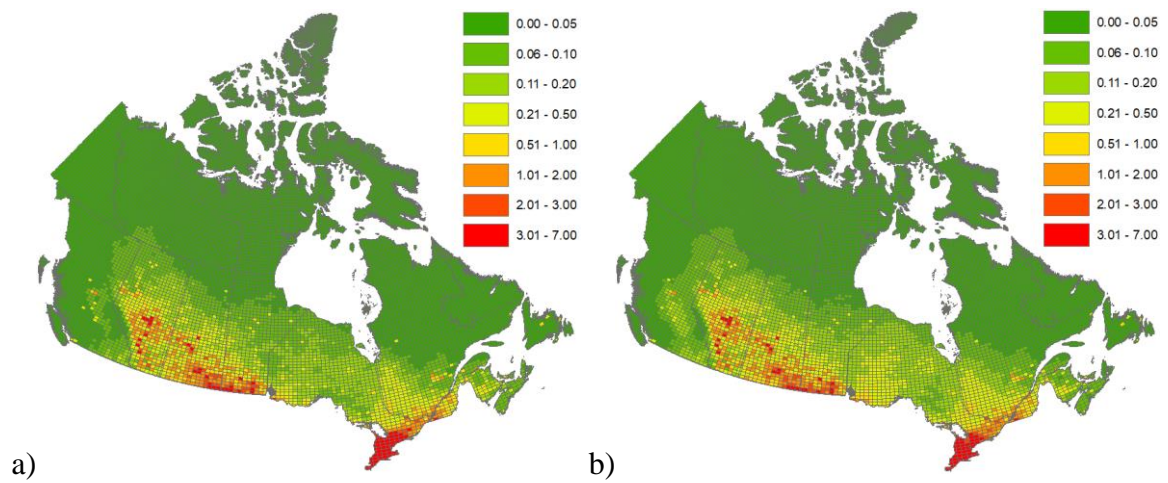
The estimated model parameters for the preferred model for each of the groups (As and Bs) are presented in Table 4.4. For comparison purposes, the preferred model by considering the ACGLF density alone as the explanatory variable is also presented in the table. A comparison of the predicted tornado occurrence rate based on the models listed in Table 4.4 is presented in Figure 4.6, and the predicted annual occurrence rate for Canada from the models (i.e., by integrating the predicted rate for all cells) is also included in Table 4.4. The plots presented in Figure 4.6 indicate that in all cases, the predicted spatial trends of the tornado activities are fairly consistent. The relative difference between the predicted annual occurrence rate for Canada is less than about 5% if the same catalogue but different explanatory variables are used. This relative difference becomes less than 6% if the same explanatory variables but different catalogues are used. The ratio of the predicted annual occurrence rate to the calculated rate directly from the considered catalogues ranges from 2.37 to 2.49, indicating that underreporting is severe. The large ratio is due to the consideration of the reporting bias correction due to population density. The calculated ratio is consistent with the value of 2.3 given in Cheng et al. (2013).

By weighing the estimated tornado occurrence rates depicted in Figures 4.6a to 4.6d equally, we estimated the boundaries for the regions, where the annual mean tornado occurrence rate equals 5×10^{-5} ($/(km^2 \times year)$) (i.e., annual mean tornado occurrence rate equals 10^{-5} ($/(km^2 \times year)$) for tornadoes with intensity between F_2 to F_5), the annual mean tornado occurrence rate equals 10^{-5} ($/(km^2 \times year)$), and the annual mean tornado occurrence rate equals 10^{-6} ($/(km^2 \times year)$). These boundaries are used to define the region prone to significant tornadoes, the region prone to tornadoes, and the region with rare tornado occurrence, as shown in Figure 4.7. The identified regions differ slightly from those given

in the commentary to the NBCC (NRC 2015).

Table 4.4. Estimated model parameters for selected models (the first and second entries represent the estimated mean and standard deviation, respectively).

Case	Model parameters					Predicted annual rate	Annual rate based on catalogue
	r	α_0	α_1	α_2	β		
A8	0.347; 0.021	-2.699; 0.061	2.399; 0.045		0.026; 0.001	152	61.3
A10	0.371; 0.023	-3.201; 0.102	1.834; 0.051	0.103; 0.011	0.025; 0.001	146	61.3
B8	0.344; 0.019	-2.708; 0.067	2.607; 0.050		0.024; 0.001	144	57.8
B10	0.367; 0.021	-3.239; 0.110	1.986; 0.056	0.111; 0.011	0.024; 0.001	137	57.8



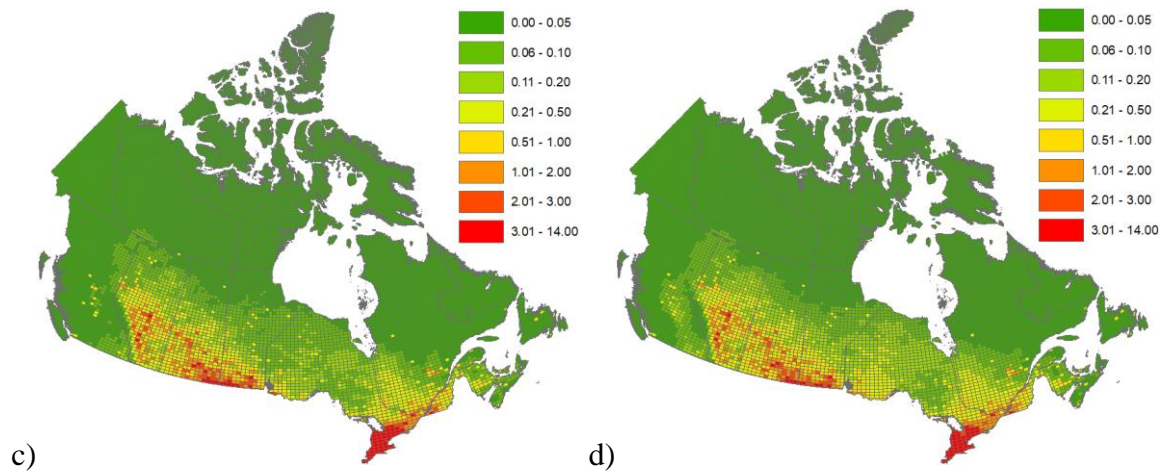


Figure 4.6. Predicted tornado occurrence rate (per year per 10⁴ km²) based on the NB model: a) using EC-T80-09 and ACGLF as the explanatory variable; b) using EC-T80-09 and ACGLF and ATD as the explanatory variables; c) using T80-19 and ACGLF as the explanatory variable; d) using T80-19 and ACGLF and ATD as the explanatory variables.



Figure 4.7. Identified contour lines with the annual mean tornado occurrence rate (per km^2 per year) equal to 5×10^{-5} , (I: Region prone to significant tornadoes), 10^{-5} (II: Region prone to tornadoes), and 10^{-6} (III: Region with rare tornado occurrence).

4.3 Tornado striking rate and database of probability distributions of wind velocity and its effects

4.3.1 Analysis procedure

A tornado can be characterized by its path orientation, width, and length. The intensity of the tornado varies along its path. Probabilistic models for the tornado width and length are given in Banik et al. (2007) and Tan and Hong (2010) for southern Ontario based on the data from southern Ontario and its neighboring regions in the United States. These probabilistic models are used in the following due to data scarcity in other regions in Canada. For the path orientation, it is assumed that it is uniformly distributed along eight directions, with one direction oriented towards the east and each separated by 45 degrees. This practical simplifying assumption may not be entirely correct for a region in Canada such as southern Ontario (Banik et al. 2007; Tan and Hong 2010), it is justified since there are insufficient data to develop site-dependent probability distribution of path orientation. Furthermore, the actual tornado intensity, F_{Aj} , varies randomly along its path for a given

tornado with reported or classified intensity F_i . A probabilistic model defining the percentage of the path length of striking intensity F_{Aj} conditioned on the occurrence of the tornado with intensity F_i was given in Twisdale et al. (1981), which is adopted for the numerical analysis to be presented in the following.

Based on the above consideration, we assess the tornado striking a site with the actual intensity F_{Aj} by using the simulation procedure described in Tan and Hong (2010). The procedure requires simulating the location of the occurrence of the tornado (i.e., touchdown point), its path orientation, length, and width, and the intensity along its path length. If the sampled path covers the site of interest with the actual intensity F_{Aj} , a striking event with F_{Aj} is counted. By repeating this simulation procedure for a sufficient number of years, the total number of tornadoes striking the site of interest with F_{Aj} is counted, and the (average) annual striking rate γ_{Aj} to the site conditioned on F_{Aj} is obtained.

Consider that a point-like or line-like structure is located at the site of interest. If a point-like structure at a height z (m) above the ground surface is located at the site, let $V(z)$ denote the maximum wind speed experienced by the point-like structure. If a line-like structure of height H (m) is located at the site, let $M(z,H)$ and $S(z,H)$ denote the maximum bending moment and maximum shear at height z for the line-like structure, respectively. For the evaluation of the probability distributions of $V(z)$, $M(z,H)$ and $S(z,H)$ conditioned on F_{Aj} , we simulate the model parameters for the adopted wind field model (Twisdale et al. 1981), including the tornado translation velocity, to define the sampled wind field (Twisdale 1978; Twisdale et al. 1981; Banik et al. 2007; Tan and Hong 2010). Based on the sampled wind field and the translation velocity, we calculate $V(z)$, $M(z,H)$ and $S(z,H)$, where,

$$M(z, H) = \max_{\theta} \left(\frac{1}{2} C_D \rho_{air} B \int_z^H v^2(x|\theta)(x-z) dx \right) \quad (4.7)$$

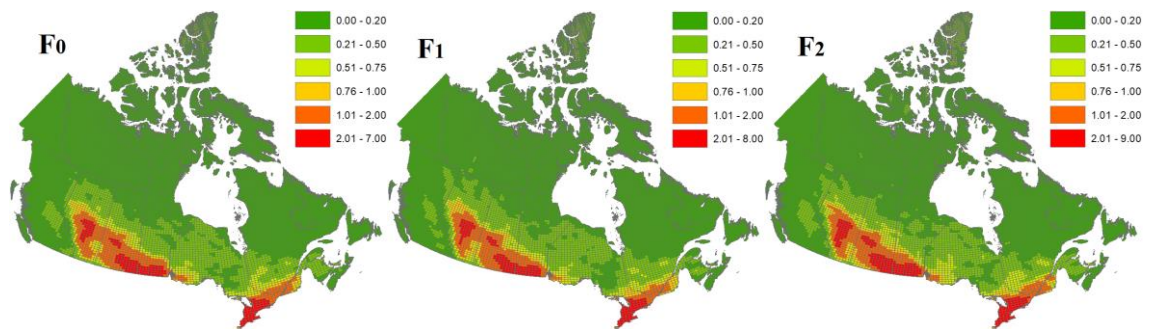
and

$$S(z, H) = \max_{\theta} \left(\frac{1}{2} C_D \rho_{air} B \int_z^H v^2(x|\theta) dx \right) \quad (4.8)$$

in which $v(x|\theta)$ is the wind velocity at height x above the ground surface for a given orientation θ , C_D represents the drag coefficient, ρ_{air} denotes the air density, and B represents the width of the line-like structure that is idealized as a circular prismatic structure. For more details on the simulation steps, the reader is referred to Banik et al. (2007) and Tan and Hong (2010). The samples of $V(z)$, $M(z,H)$ and $S(z,H)$ by considering a large simulation cycle are then sorted and used to represent their probability distributions, denoted as $P(V(z) < v | F_{Aj})$, $P(M(z,H) < m | F_{Aj})$ and $P(S(z,H) < s | F_{Aj})$, respectively. These distributions are to be stored in a database (referred to as D-database for simplicity) and used for tornado hazard evaluation.

4.3.2 Estimated tornado striking rate and assessed probability distributions of $V(z)$, $M(z,H)$ and $S(z,H)$

Following the simulation procedures described in the previous section, first, the striking rate is estimated by using the tornado occurrence rate predicted based on the preferred model, as shown in Figure 4.6b (i.e., Case A10). The use of this case is based on the consideration that it is the preferred model, and it also serves almost as an upper bound for the predicted annual occurrence rate for Canada. For the analysis, the simulation is carried out for 30000 years of tornado activities. The estimated annual striking rate γ_{Aj} is presented in Figure 4.8. The spatial distribution of strikes reflects the spatial distribution of occurrence.



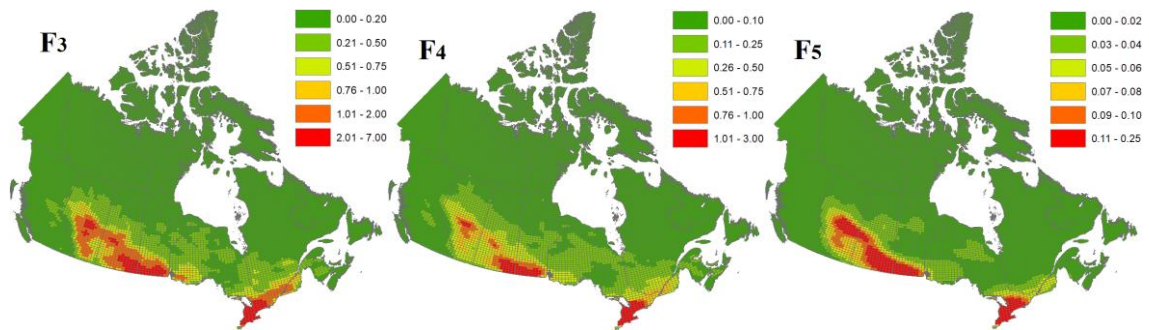
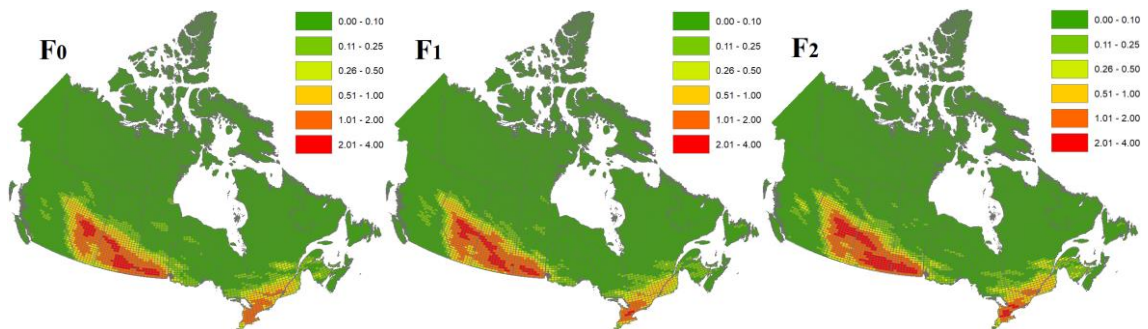


Figure 4.8. Spatially varying annual striking rate γ_{Aj} (number per 10^5 years) using occurrence rate shown in Figure 4.6b.

As part of sensitivity analysis, we repeat the estimation of γ_{Aj} by considering the tornado occurrence rate shown in Figure 4.2b, which is obtained by applying GKS and using T80-019. The results of this case serve as a “lower bound” case since no reporting bias correction is considered in this case, and the use of T80-19 leads to the lowest annual occurrence rate for Canada. The obtained γ_{Aj} is shown in Figure 4.9. A comparison of the results presented in Figures 4.8 and 4.9 indicates that the spatial trends of striking by using different occurrence models are quite similar. The striking rates obtained by using the occurrence rate shown in Figure 4.6b are, on average larger than those obtained by using the occurrence rate depicted in Figure 2b. The ratio of the overall striking rate shown in Figure 4.8 to that shown in Figure 4.9 is about 2.3, which is consistent with the overall ratio of occurrence rate between the results presented in Figure 4.6b to that shown in Figure 4.2b.



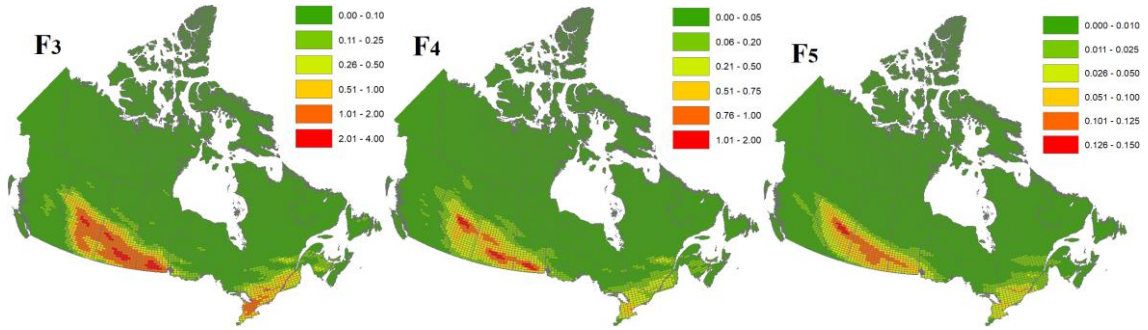


Figure 4.9. Spatially varying annual striking rate γ_{Aj} (number per 10^5 years) using occurrence rate shown in Figure 4.2b.

By using the procedure described in the previous section, the assessment of $P(V(z) < v | F_{Aj})$, $P(M(z, H) < m | F_{Aj})$ and $P(S(z, H) < s | F_{Aj})$ is carried out by considering H up to 80 m and z varying from 0 to H . The obtained distributions are stored in D-database that are to be used to map tornado wind hazard and to develop equivalent design tornado wind profile in the next section. Typical probability distributions are illustrated in Figure 4.10, where the wind velocity is expressed in terms of the 3-second gust mean wind speed. The bending moment and shear force shown in Figure 10 are normalized with respect to $(C_D \rho_{air} B / 2)$, where the normalized bending moment and the normalized shear force have units of $\text{km}^2 \times \text{m}^2 / \text{h}^2$ and $\text{km}^2 \times \text{m} / \text{h}^2$, respectively. Note that a factor of 1.52 (Durst 1960) could be used to multiply the 3-second gust mean wind velocity to obtain the hourly-mean wind velocity, which is implemented in the NBCC (NRC 2015). The 3-second gust mean wind velocity is used throughout the remaining part of the present study.

The plots shown in Figure 4.10 indicate that they vary irregularly as compared to commonly used probability distributions such as the normal lognormal, Gamma, Weibull, and Gumbel distributions. When the actual intensity F_{Aj} increases, the distribution is shifted to the right, which is as expected. Also, a comparison of the plots indicates that for given values of H and F_{Aj} , $P(M(z, H) < m | F_{Aj})$ and $P(S(z, H) < s | F_{Aj})$ is shifted towards the right as z decreases, which is as expected.

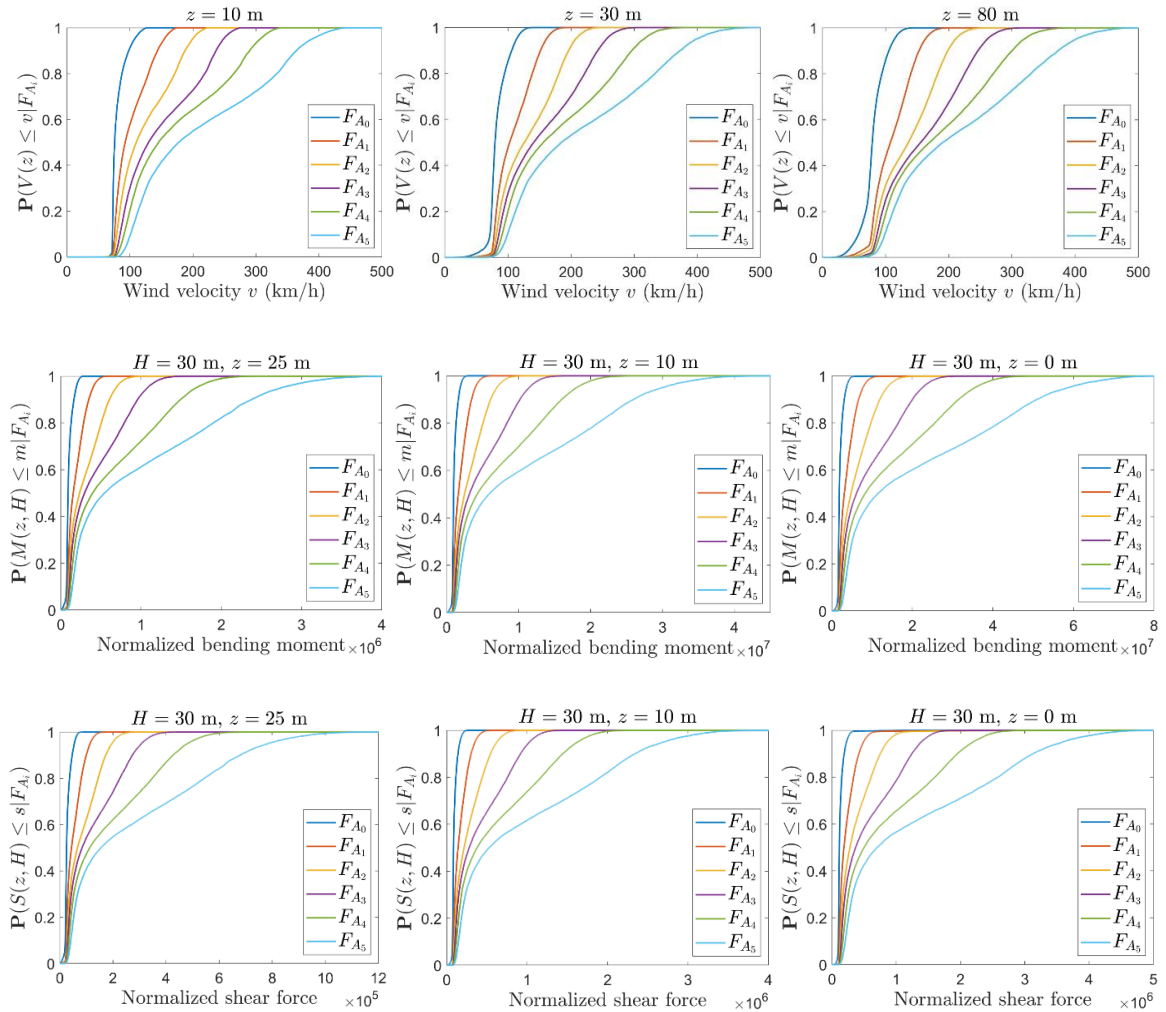


Figure 4.10. Estimated $P(V(z) < v | F_{A_j})$, $P(M(z, H) < m | F_{A_j})$ and $P(S(z, H) < s | F_{A_j})$.

4.4 Tornado hazard mapping and tornado wind profile

4.4.1 Tornado hazard assessment for a single site

Before carrying out the tornado hazard mapping, consider the estimation of the $(1-1/T)$ -quantile (or T -year return period value) of $V(z)$, $V_T(z)$, for a single site. Given the annual average striking rate γ_{A_j} which is small, the probability that $V(z)$ is greater than v in a year, $P(V(z) > v)$, can be approximated by,

$$P(V(z) > v) \approx \sum_{j=0}^5 \gamma_{A_j} P(V(z) > v | F_{A_j}) \quad (4.9)$$

Given the values of $[r_{A0}, r_{A1}, r_{A2}, r_{A3}, r_{A4}, r_{A5}]$ such as those shown in Figures 4.8 and 4.9, and the conditional probability distribution $P(V(z) > v | F_{A_j})$, which are already pre-calculated and stored in D-database, $V_T(z)$ can be calculated by iteratively solving Eq. (4.9) for a given exceedance probability of $1/T$. As an illustration, consider a site representing Toronto with the latitude and longitude equal to 43.65°N and 79.38°W . The value of $[r_{A0}, r_{A1}, r_{A2}, r_{A3}, r_{A4}, r_{A5}]$ obtained from Figure 4.8 equals $[5.32, 5.82, 6.05, 4.20, 2.16, 0.21] \times 10^{-5}$. By using these striking rates and solving Eq. (4.9), the obtained $V_T(z)$ is depicted in Figure 4.11a for z ranging from 0 to 80 m and $T = 5000, 7500, 10000, 50000,$ and 100000 years. A similar analysis is carried out for Winnipeg, where $[r_{A0}, r_{A1}, r_{A2}, r_{A3}, r_{A4}, r_{A5}]$ obtained from Figure 4.8 equals $[3.36, 2.96, 3.43, 2.87, 1.18, 0.11] \times 10^{-5}$. The obtained results are also shown in Figure 4.11b. Note that the value of $V_T(z)$ for $T = 5000$ years is not presented in Figure 4.11b. This is because the calculated $V_T(z)$ for $T = 5000$ years equals zero. In general, according to, $V_T(z)$ equals zero if T is smaller than the inverse of the sum of the striking rates for a site.

The non-smooth along-height varying wind velocity near $z = 10$ m shown in Figure 4.11 is attributed to the adopted wind field model. For comparison purposes, the power-law wind profile with an exponent equal to 0.11 for synoptic winds is also presented in the plots in Figure 4.11. For the plot, the wind speed at 10 m height is based on the 500-year return period value of the annual maximum (synoptic) wind velocity for the considered sites, which are calculated using the statistics in Hong et al. (2014). The figures show that in all cases, the rate of increase in wind velocity along the height for the quantiles of tornado wind velocity is slower than that for synoptic wind velocity. The shape of the tornado wind profile depends on T . The differences in the shape of the wind profile for the same T , but different locations, are expected since the striking rates for different sites differ. A further comparison of the probability distributions of the annual maximum synoptic wind velocity and the tornado wind velocity is presented in Figures 4.11c for Toronto and Figure 4.11d

for Winnipeg. Figures 4.11c and 4.11d emphasize that the wind velocity hazard is governed by tornado winds for a site (defined by a point) only if the exceedance probability is less than 2×10^{-5} .

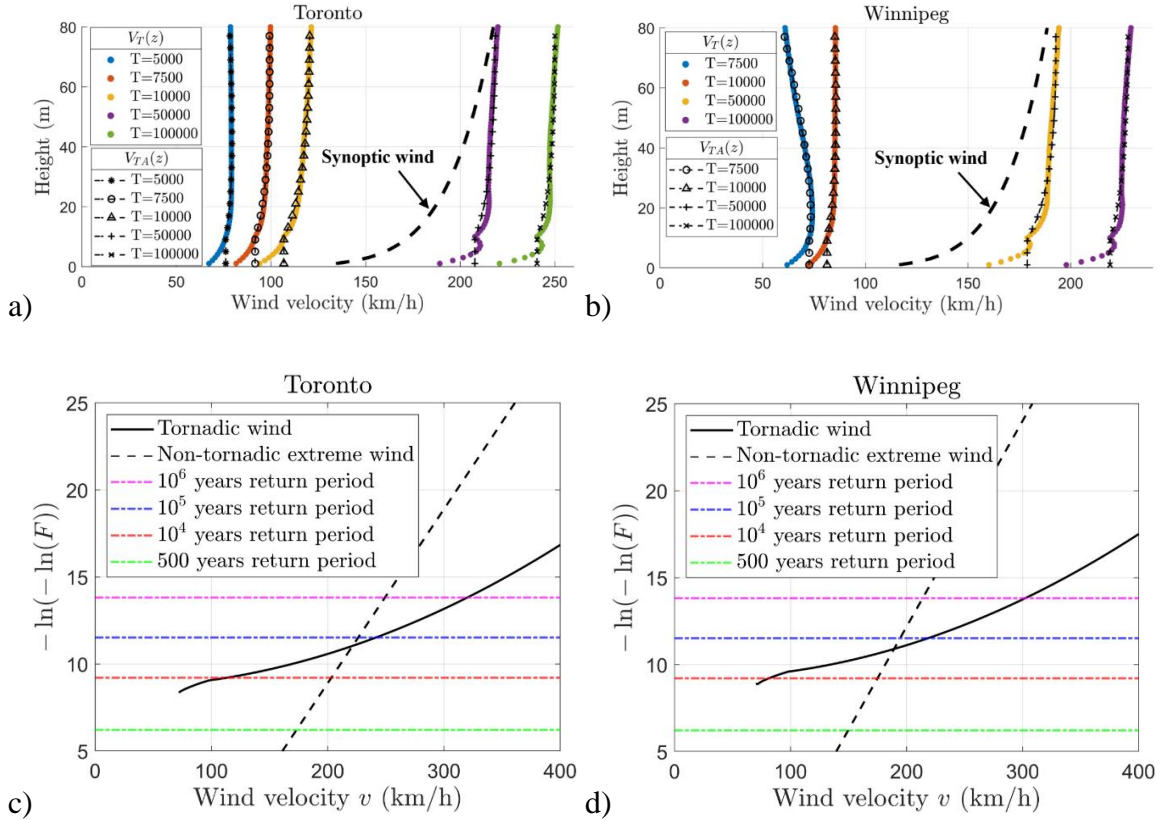


Figure 4.11. Estimated $V_T(z)$, for Toronto and Winnipeg and comparison of the probability distributions of annual maximum synoptic wind velocity and of tornado wind velocity (the wind speed is presented in terms of the 3-second gust mean wind velocity (km/h)).

It is of interest to note that if the striking rate for a site named as Site-1, γ_{1Aj} , and for a site named as Site-2, γ_{2Aj} , are small, and they are proportional (i.e., $\gamma_{1Aj}/\gamma_{2Aj} = c$), the application of Eq. (4.9) results in that $P(V(z) > v)$ for Site-2 equals c times that for Site-1. This implies that the T -year return period value for Site-1 approximately equals (cT) -year return period value for Site-2. This and the fact that the shape of tornado wind profiles shown in Figures 11a and 11b depends on $V_T(10)$ suggest that we could adopt the following simple parametric form to approximate $V_T(z)$,

$$V_{TA}(z) = \begin{cases} V_T(10) & \text{for } z < 10 \text{ m} \\ V_T(10) \times \left(\frac{z}{10}\right)^{\alpha_1 - \alpha_2 z} & \text{for } 10 \leq z < 80 \text{ m} \end{cases} \quad (4.10)$$

with the model parameters α_1 and α_2 to be determined through the regression analysis.

The fitted Eq. (4.10) to the results presented in Figures 4.11a and 4.11b is also included in the same figure. The comparison presented in Figure 4.11 indicates that the fit is adequate. The obtained α_1 and α_2 as functions of $V_T(10)$ are presented in Figure 4.12; in the same figure, the values of α_1 and α_2 calculated by considering additional return periods are also shown. The figure indicates that α_1 and α_2 could be represented as functions of $V_T(10)$. These suggest that Eq. (4.10) can be used advantageously to define the tornado wind velocity hazard at a site based on $V_T(10)$, α_1 , and α_2 .

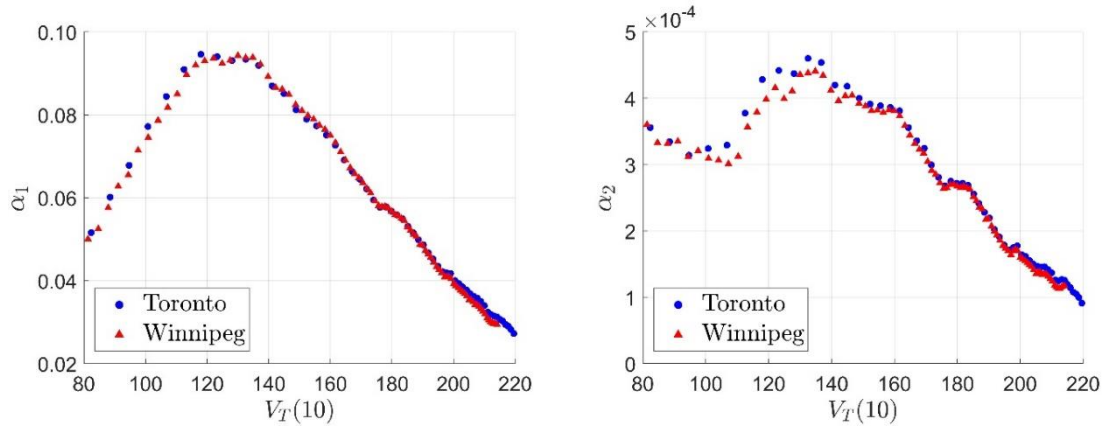


Figure 4.12. Values of α_1 and α_2 as functions of $V_T(10)$ for Toronto and Winnipeg.

However, it is unknown whether the fitted model shown in Eq. (4.10) could lead to satisfactory approximation to the $(1-1/T)$ -quantile of $M(z,H)$, $M_T(z,H)$, and $(1-1/T)$ -quantile of $S(z,H)$, $S_T(z,H)$ for a line-like structure. To investigate this, we note that $M_T(z,H)$ and $S_T(z,H)$ can be calculated by solving,

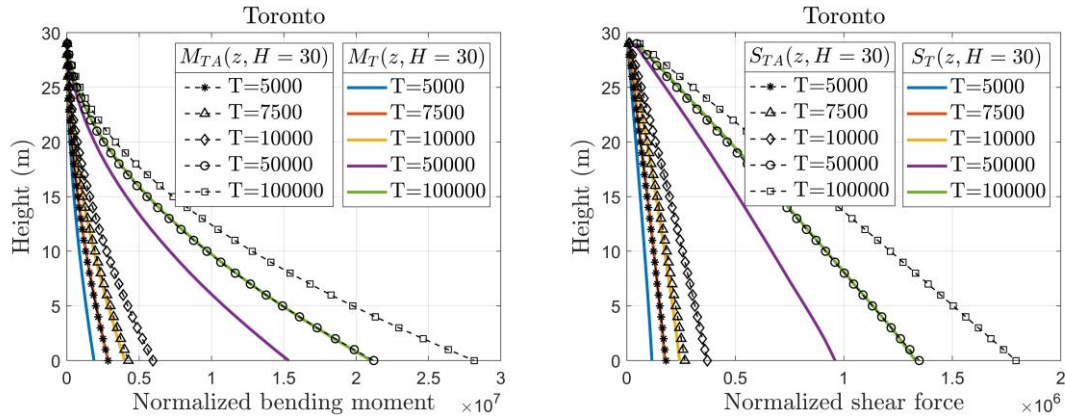
$$P(M(z,H) > m) \approx \sum_{j=0}^5 \gamma_{A_j} P(M(z,H) > m | F_{A_j}) \quad (4.11)$$

and,

$$P(S(z, H) > s) \approx \sum_{j=0}^5 \gamma_{Aj} P(S(z, H) > s | F_{Aj}). \quad (4.12)$$

for the exceedance probability equal to $1/T$.

Again, as an illustrative example, consider the same sites representing Toronto and Winnipeg. By solving Eqs. (4.11) and (4.12), the obtained $M_T(z, H)$ and $S_T(z, H)$ for $H = 30$, and $T = 5000, 7500, 10000, 50000$ and 100000 years are shown in Figure 4.13. In the same figure, we show the calculated bending moment and shear force for the line-like structure by using $V_T(z)$ defined by Eq. (4.10) (denoted as $M_{TA}(z, H)$ and $S_{TA}(z, H)$). The comparison shown in the figure indicates that, in general, $M_{TA}(z, H)$ and $S_{TA}(z, H)$ overestimate $M_T(z, H)$ and $S_T(z, H)$. The overestimation increases as z decreases. The maximum overestimation depends on site, height, and return period. It also depends on if the bending moment or shear is considered.



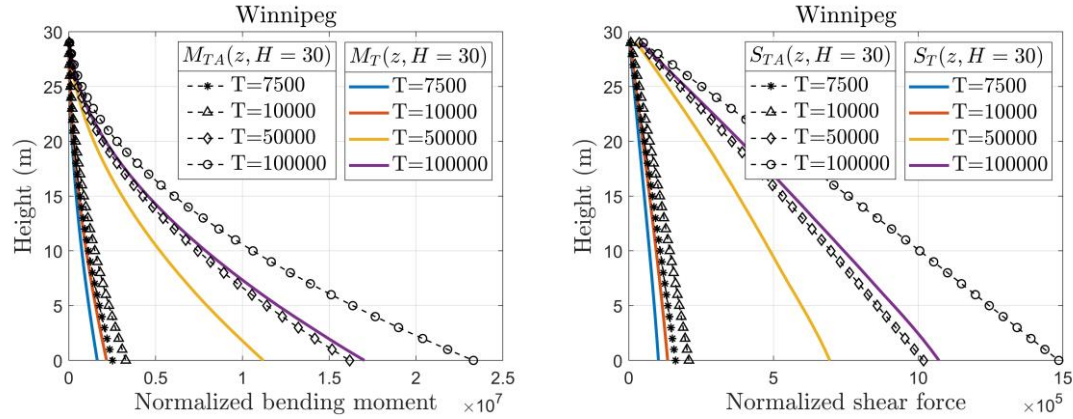


Figure 4.13. The comparison of $M_T(z, H)$ and $M_{TA}(z, H)$ and of $S_T(z, H)$ and $S_{TA}(z, H)$.

This analysis is repeated for the line-like structure with $H = 20, 40, 50, 60,$ and 80 m. In all cases, the observed maximum overestimation is less than about 15% for the bending moment and about 17% for the shear force for the two considered sites. Therefore, Eq. (4.10) can be used as an equivalent tornado design wind profile but with an additional reduction factor for a line-like structure, at least for the two considered sites. The required reduction factor will be discussed again shortly.

4.4.2 Tornado hazard mapping

In the previous section, it was pointed out that the tornado wind velocity hazard for a site can be defined based on Eq. (4.10) with $V_T(10)$, α_1 , and α_2 as parameters. The analysis that is carried out for Toronto and Winnipeg is repeated for each of the points on the grid system in Figure 4.3. The obtained $V_T(10)$ is presented in Figures 4.14a to 4.14c for $T = 10,000, 50,000,$ and $100,000$ years. For comparison purposes, we included the 500-year return period value of the annual maximum wind velocity estimated based on the wind records at meteorological stations (Hong et al. 2014).

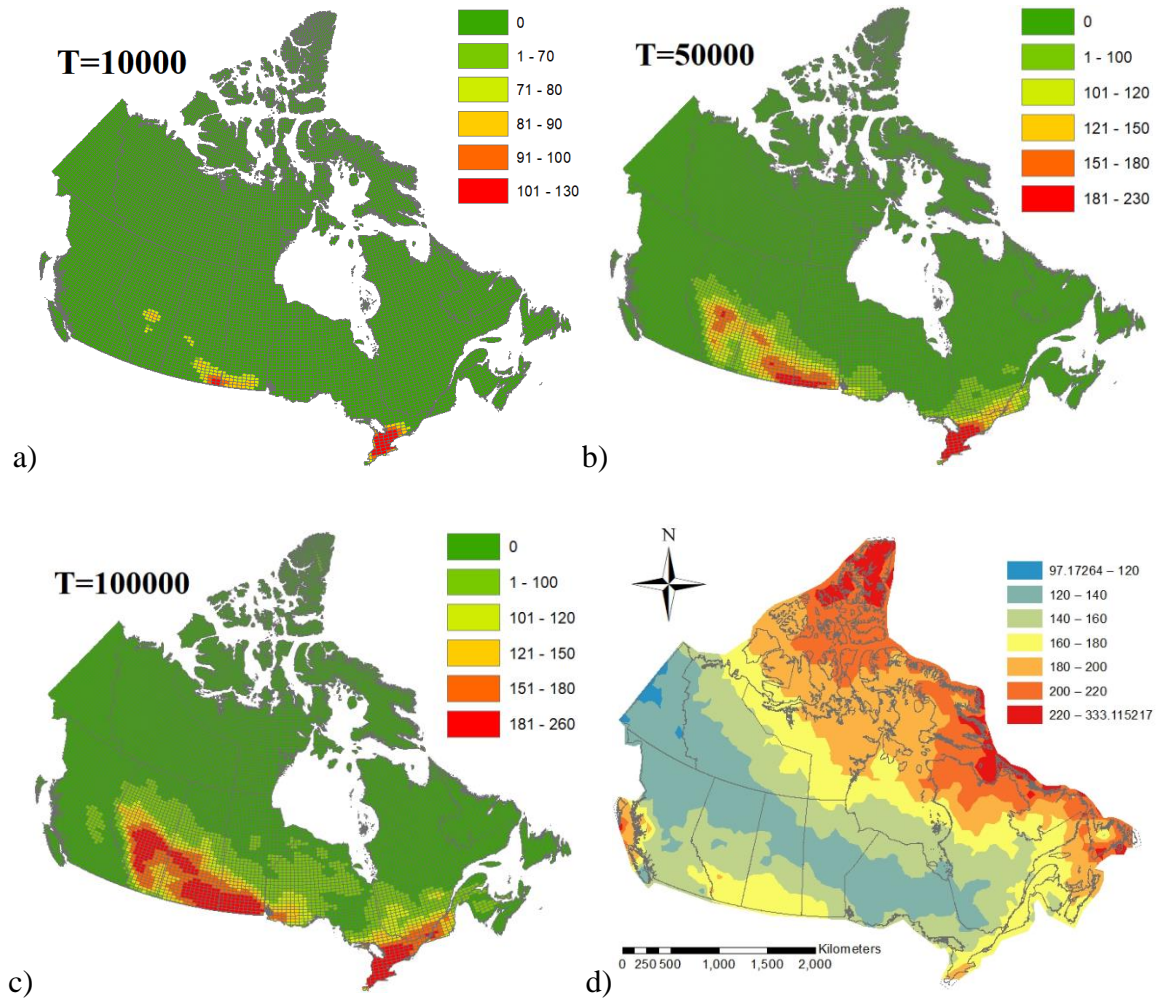


Figure 4.14. Estimated wind hazard maps: a) to c) Tornado wind hazard maps by considering the tornado occurrence rate shown in Figure 4.6b (i.e., striking rates shown in Figure 4.8) and for $T = 10,000$, $50,000$, and $100,000$ years, respectively; d) synoptic wind hazard map for $T = 500$ years.

In presenting the results in Figure 4.14a to 4.14c, $V_T(10)$ equal to zero is assigned if the total striking rate is smaller than the considered exceedance probability (i.e., $1/T$), as mentioned earlier. These plots indicate that the estimated $V_T(10)$ is spatially varying, reflecting the spatially varying striking rates shown in Figure 4.8. The comparison of the results presented in the figure indicates that in most cases, the tornado wind velocity for T less than 10,000 years is less than the 500-year return period value of the annual maximum synoptic wind velocity for all locations.

The results depicted in Figures 4.14a to 4.14c are obtained by using the tornado occurrence rate model shown in Figure 4.6b (i.e., striking rates shown in Figure 4.8), which serves as an upper bound, as mentioned earlier. By using the tornado occurrence rate shown in Figure 4.2b (i.e., striking rates shown in Figure 4.9) as a lower bound, the mapping of tornado wind velocity hazard is repeated. The obtained maps are presented in Figure 4.15. The figure only shows the values of $V_T(10)$ for $T = 50,000$ and $100,000$ years. The values of $V_T(10)$ for $T = 10,000$ years are not presented since, in this case, $V_T(10)$ is smaller than 100 km/h for almost all places. The spatial trends of the tornado wind velocity hazard maps shown in Figure 4.15 are consistent with those presented in Figure 4.14, especially for regions with a population density greater than about five people/km². In other words, for regions prone to significant tornadoes (see Figure 4.7), the differences between the estimated tornado wind velocity hazards are small by using the rates shown in Figure 4.2b and 4.6b.

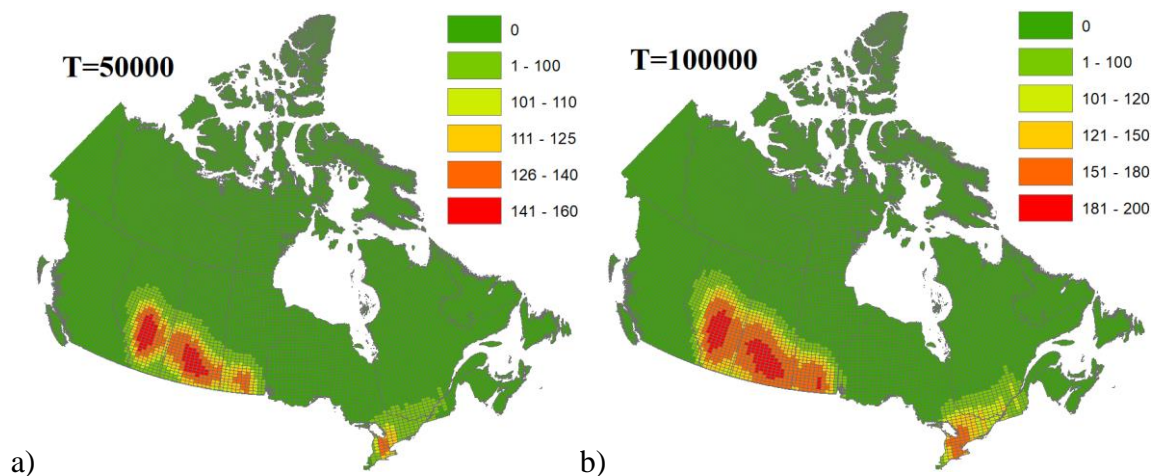


Figure 4.15. Tornado wind velocity hazard maps based on the occurrence rate model shown in Figure 4.2b (i.e., the striking rate shown in Figure 4.9): a) for $T = 50,000$ years, and b) for $T = 100,000$ years.

It must be emphasized that the comparison of synoptic and tornado wind hazards is based on a structure that has a footprint represented by a point. The synoptic wind hazard map could be assumed to be applicable for structures of different footprint sizes. However, this

is not the case for the tornado wind hazard maps. For example, for a structure with footprint represented by a line of 1 km, and 10 km the results presented in Banik et al. (2007, 2008) imply that the striking rate (at least part of the structure is hit by a tornado) is increased (very approximately) by about 5 to 15 times and by 20 to 200 times, respectively. This indicates that the tornado wind hazard depicted in Figures 4.14a to 4.14c and Figures 4.15a to 4.15c could be used for this type of structures except that T presented in the figures should be divided by about 10 for a line of 1 km and by about 100 for a line of 10 km. It indicates that the tornado wind velocity hazard for a line of 10 km could be greater than or comparable to the synoptic wind hazard for a return period of 500 years.

By considering a structure with a footprint represented by a circle of a diameter D equal to 100 m (representing an office building), the striking rates are estimated by using the occurrence rate shown in Figure 4.6b. The obtained rates are shown in Figure 4.16. The spatial trends of the striking rate are almost identical to that shown in Figure 4.8. However, the rate shown in Figure 4.16 is greater than that presented in Figure 4.8. To better appreciate the difference, the mean and standard deviation of the ratio of the striking rate of the structure with $D = 100$ m to that of a structure with a point footprint, denoted as R_{ST} , are evaluated for region prone to tornadoes (see Figure 4.7). The estimated statistics are presented in Table 4.5. The statistics indicate that, in general, the mean of R_{ST} ranges from 1.3 to 2.9. Therefore, for a structure with $D = 100$ m, the tornado wind hazard is the same as those shown in Figures 4.14 and 4.15 except that T depicted in the figures should be divided by a factor ranging from about 1.3 to 2.9.

A similar analysis is carried out but considering $D = 20$ (representing a house) and the striking rate shown in Figure 4.2b. Since the spatial trends are similar to those presented in Figure 4.16, the rates are not plotted. However, the statistics of R_{ST} are summarized in Table 4.5. In general, for a structure with $D = 20$ m, the tornado wind hazard shown in Figures 4.14 and 4.15 can be used except that T shown in the figures should be divided by a factor ranging from 1.1 to 1.6. Considering that the factored design wind load corresponds approximately the 500-year return period value of the annual maximum synoptic wind speed, and the factored design resistance is less than the 0.1-quantile of the resistance variable, the estimated annual probability of a small building being damaged by a tornado

is likely to be less than 10^{-5} . This estimate is consistent with the observation made by Allen (1992); and implies that it is usually uneconomic design small buildings to prevent tornado loss beyond that required by the code. However, this observation may not be extended for a building complex, where the failure of an individual building or component would like to the impaired functionality and damage to the entire system.

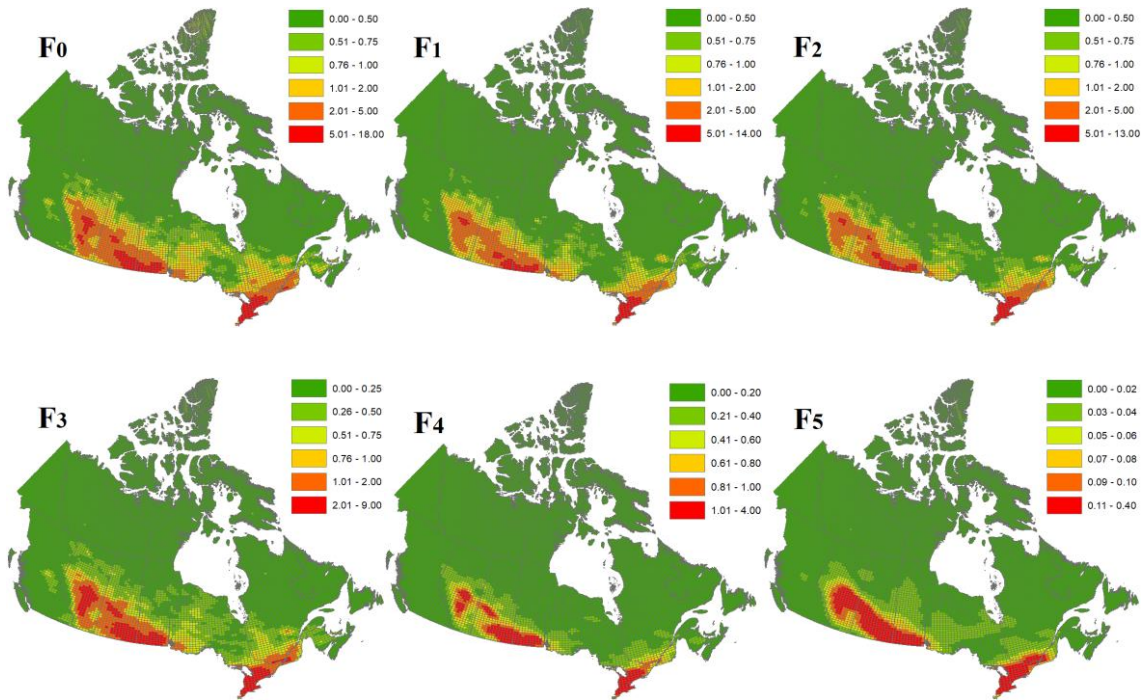


Figure 4.16. Spatially varying annual striking rate γ_{Aj} (number per 10^5 years) using occurrence rate shown in Figure 4.6b for a structure with footprint represented by a circle with a diameter of 100 m.

Table 4.5. Statistics of R_{ST} (The first and second entries represent the mean and the standard deviation).

Model for structural footprint & considering tornado occurrence model	F_{A_0}	F_{A_1}	F_{A_2}	F_{A_3}	F_{A_4}	F_{A_5}
Footprint with $D = 100$ m & striking rate shown in Figure 4.6b	2.89; 1.91	2.26; 1.52	1.87; 1.24	1.86; 1.37	1.54; 0.87	1.31; 0.45
Footprint with $D = 100$ m &	2.60;	1.94;	1.76;	1.66;	1.64;	1.43;

striking rate shown in Figure 4.2b	1.71	1.42	1.11	1.02	0.86	0.64
Footprint with $D = 20$ m & striking rate shown in Figure 4.6b	1.56; 0.91	1.51; 0.89	1.33; 0.86	1.27; 0.94	1.18; 0.66	1.05; 0.41
Footprint with $D = 20$ m & striking rate shown in Figure 4.2b	1.55; 0.78	1.39; 0.60	1.28; 0.64	1.21; 0.61	1.16; 0.80	1.08; 0.38

The parameters α_1 and α_2 for the equivalent wind profile was calculated and shown in Figure 4.12 as functions of $V_T(10)$ for two sites. It is unknown if the illustrated relation in Figure 4.12 is applicable to other sites in Canada. In order to propose relations between α_1 and $V_T(10)$, and between α_2 and $V_T(10)$. The analysis that is carried out for the results presented in Figure 4.12 is repeated for grid points covering Canada. The obtained results are presented in Figure 4.17, showing visible trends. Based on the values presented in Figure 4.17, the following two simple empirical equations are proposed for calculating α_1 and α_2 ,

$$\alpha_1 = \begin{cases} -0.032 + 0.114\delta, & \text{for } 80 \leq V_T(10) \leq 125 \\ 0.203 - 0.074\delta, & \text{for } 125 < V_T(10) \leq 225, \\ 0.0366, & \text{for } 225 < V_T(10) \end{cases} \quad (4.13)$$

and,

$$\alpha_2 = 10^{-4} \times \begin{cases} 0.55 + 5.52\delta, & \text{for } 80 \leq V_T(10) \leq 140 \\ 16 - 5.70\delta, & \text{for } 140 < V_T(10) \leq 225. \\ 3.43, & \text{for } 225 < V_T(10) \end{cases} \quad (4.14)$$

where $\delta = V_T(10)/100$. These values differ very slightly from those obtained if only southern Ontario is considered. The coefficients of these equations are obtained based on nonlinear regression analysis. The comparison of the proposed empirical equations and the data shown in Figure 4.17 indicates that the suggested empirical equation represents the trends adequately.

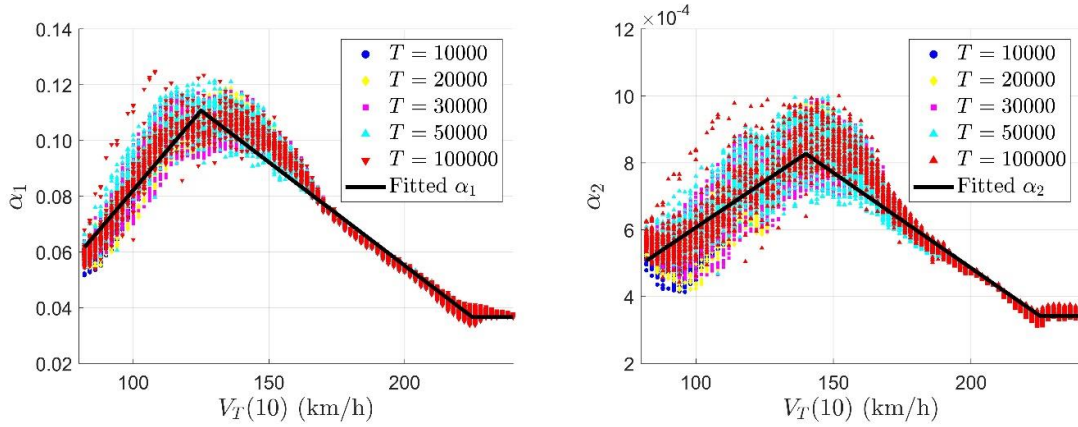


Figure 4.17. Values of α_1 and α_2 as functions of $V_T(10)$ and fitted simple empirical relations.

Besides defining the equivalent wind profile for a point-like structure, it was indicated in the previous sections, an equivalent wind profile to calculate the bending moment and shear force for a line-like structure subjected to tornado wind velocity hazard is required. It was indicated in the previous section that the direct use of the wind profile defined by Eq. (4.10) could overestimate the bending moment and shear force for the line-like structure, at least for Toronto and Winnipeg. To further investigate this overestimation, additional four sites (Ottawa, London, Regina, and Calgary) are considered, the required reduction factor for

$\sum_{\text{Over } z} (M_{TA}(z, H) - M_T(z, H))^2$ to be minimum and the required reduction factor for

$\sum_{\text{Over } z} (S_{TA}(z, H) - S_T(z, H))^2$ to be minimum are calculated for $H = 20, 30, 40$ and 60 m .

These calculated values are shown in Figure 4.18. It can be observed from the figure, in all cases, $\text{Min}(0.79 + 3.5 \times 10^{-4} V_{10}(T), 0.87)$ almost envelops the reduction factors. Therefore, an equivalent tornado wind profile for the line-like structure given below is recommended,

$$V_{TA}(z) = \text{Min}(0.79 + 3.5 \times 10^{-4} V_{10}(T), 0.87) \times \begin{cases} V_T(10), & \text{for } z < 10 \text{ m} \\ V_T(10) \times \left(\frac{z}{10}\right)^{\alpha_1 - \alpha_2 * z}, & \text{for } 10 \leq z < 80 \text{ m} \end{cases}, \quad (4.15)$$

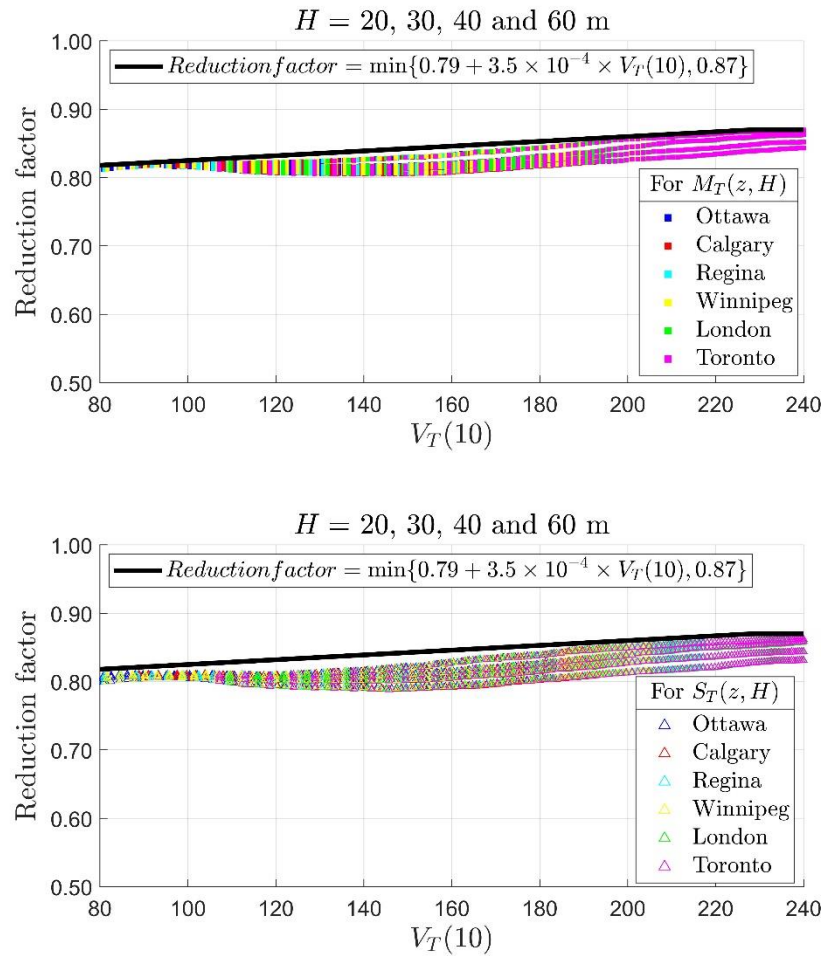


Figure 4.18. Calculated reduction factor if the equivalent tornado wind profile defined by Eqs. (10), (13) and (14) is used to calculate the bending moment and shear force: a) for bending moment, b) for shear force.

4.4.3 Discussions

The analysis framework presented in the present study for assessing the tornado occurrence rate, mapping tornado wind velocity hazard, and developing an equivalent tornado design wind profile can be applied for other H regions in the world. It can also be used if different tornado wind profiles are adopted. Although not elaborated, the extension of the analysis framework by considering combinations of different tornado occurrence rate models, tornado intensity models, and tornado wind field models is straightforward based on the

logic tree approach commonly used for seismic hazard mapping (McGuire 2004).

Due to the lack of sufficient high-quality data on the tornado path orientation, length, and width for Canada, the assumptions made in developing the maps shown in Figures 4.14 and 4.15 may not be entirely accurate. For example, there is a preferred tornado orientation for tornado events that occurred in southern Ontario. The probabilistic model for the tornado orientation could affect the accuracy of the estimated tornado hazard for infrastructure extended in many kilometers (e.g., transmission lines). Therefore, for a large spatially distributed infrastructure system, a more detailed and project-specific analysis may be required.

The wind field model parameters used for mapping the tornado wind velocity hazard for Canada presented in the present study are based on those given in Twisdale et al. (1981) and Banik et al. (2007). An update of the tornado path and wind field models by incorporating new tornado damage survey information, the tornado database available from the United States, and the wind field characteristics observed from full scale and laboratory tests should be carried out. The use of such updated models for tornado wind velocity hazard modelling can enhance our understanding of the tornado hazard for Canada.

4.5 Conclusions

Systematic modelling of tornado occurrence and assessment of tornado wind velocity hazard map are carried out for Canada based on two historical tornado catalogues. One of the catalogues covering the reported tornado from 1980 to 2009 was available in the literature, and the other covering reported events from 1980 to 2019 was developed in the present study. The Poisson model, zero-inflated Poisson model, and negative binomial model, as well as the adaptive Gaussian kernel smooth technique, are considered. When using the stochastic models, the tornado reporting bias due to population density is considered, and the cloud-to-ground lightning flash density and the thunderstorm days are used as the explanatory variables.

The comparison of the two catalogues for the reported tornadoes indicates that the annual average tornado occurrence for Canada has decreased by about 30% as compared to those reported from 1980 to 2009. This large discrepancy may be due to several different factors, including the use of various sources of the reported tornado activities.

Statistical analysis results indicate that the negative binomial model is preferable. The annual average tornado occurrence by considering the reporting bias is about 2.4 times that obtained without considering the bias. However, most of this increase is for regions with very low population density (i.e., population density much lower than five people per km²), and the increase is minimal for areas with a population density greater than about five people per km².

The tornado wind velocity hazard maps for Canada, in terms of wind velocity at 10m height above the ground surface and for a return period T , $V_T(10)$ (in terms of 3-second gust mean wind speed in km/h), are developed by considering an “upper bound” and a “lower bound” tornado occurrence model. In addition, a simple along height equivalent tornado design wind profile for point-like structures is developed. The equivalent wind profile is approximately represented by the power-law wind profile, but with the exponent varying with $V_T(10)$ and the height. It is shown that such an equivalent wind profile can be used to calculate the bending moment and shear force for (vertical) line-like structures by including a reduction factor ranging from about 0.85 for $V_T(10)$ varying from 80 to 240 km/h. The developed $V_T(10)$ and equivalent wind profile facilitate the implementation of tornado wind loading in the structural design codes and standards for structural design checking and risk modelling.

A comparison with the wind velocity hazard due to the synoptic and tornado winds indicates that, in general, the wind hazard is dominated by the synoptic winds for a structure with a footprint represented by a point if the considered return period is in the order of 500 years. However, the tornado winds could dominate the wind hazard as the length of the footprint of an infrastructure system or the area of the footprint of a structure increases. This indicates that the consideration of tornado winds is necessary for a spatially extended

building complex and infrastructure system that are critical for the safe operation of the society.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Allen, D.E., (1992) Discussion of “A design basis tornado” by M. J. Newark, *Can. J. Civ. Eng.*, 19, 361.
- Anderson, C.J., Wikle, C.K., Zhou, Q., & Royle, J.A. (2007). Population influences on tornado reports in the United States. *Weather and forecasting*, 22(3), 571-579.
- Banik, S.S., Hong, H.P., & Kopp, G.A. (2007). Tornado hazard assessment for southern Ontario. *Canadian Journal of Civil Engineering*, 34(7), 830-842.
- Banik, S.S., Hong, H.P., & Kopp, G.A. (2008). Assessment of tornado hazard for spatially distributed systems in southern Ontario. *Journal of wind engineering and industrial aerodynamics*, 96(8-9), 1376-1389.
- Bissolli, P., Grieser, J., Dotzek, N., & Welsch, M. (2007). Tornadoes in Germany 1950–2003 and their relation to particular weather conditions. *Global and Planetary Change*, 57(1-2), 124-138.
- Branick, M.L., & Doswell III, C.A. (1992). An observation of the relationship between supercell structure and lightning ground-strike polarity. *Weather and forecasting*, 7(1), 143-149.
- Brooks, S., Gelman, A., Jones, G., & Meng, X.L. (Eds.). (2011). *Handbook of markov chain monte carlo*. CRC press.
- Burnham, K., & Anderson, D. (2002). *Model Selection and Multi-model Inference*. 2nd edn. Springer, New York.

- Burrows, W.R., & Kochtubajda, B. (2010). A decade of cloud - to - ground lightning in Canada: 1999 - 2008. Part 1: Flash density and occurrence. *Atmosphere-ocean*, 48(3), 177-194.
- Carey, L.D., & Buffalo, K.M. (2007). Environmental control of cloud-to-ground lightning polarity in severe storms. *Monthly weather review*, 135(4), 1327-1353.
- Cheng, V.Y., Arhonditsis, G.B., Sills, D.M., Auld, H., Shephard, M.W., Gough, W.A., & Klaassen, J. (2013). Probability of tornado occurrence across Canada. *Journal of Climate*, 26(23), 9415-9428.
- Cheng, V.Y., Arhonditsis, G.B., Sills, D.M., Gough, W.A., & Auld, H. (2016). Predicting the climatology of tornado occurrences in North America with a Bayesian hierarchical modelling framework. *Journal of Climate*, 29(5), 1899-1917.
- Daley, D.J., & Vere-Jones, D. (2007). *An introduction to the theory of point processes: volume II: general theory and structure*. Springer Science & Business Media.
- Elsner, J.B., & Widen, H.M. (2014). Predicting spring tornado activity in the central Great Plains by 1 March. *Monthly Weather Review*, 142(1), 259-267.
- Etkin, D., Brun, S. E., Shabbar, A., & Joe, P. (2001). Tornado climatology of Canada revisited: tornado activity during different phases of ENSO. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 21(8), 915-938.
- Hangan, H., & Kim, J.D. (2008). Swirl ratio effects on tornado vortices in relation to the Fujita scale. *Wind and Structures*, 11(4), 291-302.
- Heinze, G., Wallisch, C., & Dunkler, D. (2018). Variable selection—A review and recommendations for the practicing statistician. *Biometrical Journal*, 60(3), pp.431-449.
- Hong, H.P., & Huang, Q. (2020). Simplified Hazard Modelling and Structural Reliability Analysis Considering Non-Synoptic Wind Systems (NSWS), in *Handbook of Non-Synoptic Wind Storms* edited by Horia Hangan and Ahsan Kareem, Oxford press.

- Hong, H.P., Mara, T.G., Morris, R., Li, S.H., & Ye, W. (2014). Basis for recommending an update of wind velocity pressures in Canadian design codes. *Canadian Journal of Civil Engineering*, 41(3), 206-221.
- Huang, Q., Jiang, W.J., & Hong, H.P. (2020) Statistical assessment of spatial Tornado occurrence for Canada: modelling and estimation, Report to NRC.
- Jagger, T. H., Elsner, J.B., & Widen, H.M. (2015). A statistical model for regional tornado climate studies. *PloS one*, 10(8).
- Johnston, K., Ver Hoef, J.M., Krivoruchko, K., & Lucas, N. (2003). ArcGIS9, Using ArcGIS geostatistical analyst, Redlands, CA: Environmental Systems Research Institute.
- King, P. (1997). On the absence of population bias in the tornado climatology of southwestern Ontario. *Weather and Forecasting*, 12(4), 939-946.
- King, P.W., Leduc, M.J., Sills, D.M., Donaldson, N.R., Hudak, D.R., Joe, P. & Murphy, B.P. (2003). Lake breezes in southern Ontario and their relation to tornado climatology. *Weather and forecasting*, 18(5), pp.795-807.
- Knapp, D.I. (1994). Using cloud-to-ground lightning data to identify tornadic thunderstorm signatures and nowcast severe weather. *Natl. Wea. Dig*, 19(2), 35-42.
- Kopp, G.A., Hong, E., Gavanski, E., Stedman, D., & Sills, D.M. (2017). Assessment of wind speeds based on damage observations from the Angus (Ontario) Tornado of 17 June 2014. *Canadian Journal of Civil Engineering*, 44(1), 37-47.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1), 1-14.
- Lee, P.M. (2012). *Bayesian statistics: An introduction* (4th ed.). West Sussex, UK: Wiley.
- MacGorman, D.R. & Burgess, D.W. (1994). Positive cloud-to-ground lightning in tornadic storms and hailstorms. *Monthly Weather Review*, 122(8), 1671-1697.

- Marshall, T.P., McDonald, J.R., & Forbes, G.S. (2004). The enhanced Fujita (EF) scale. In *Preprints, 22nd Conf. on Severe Local Storms, Hyannis, MA, Amer. Meteor. Soc. B* (Vol. 3).
- McGuire, R. K. (2004). *Seismic hazard and risk analysis*, EERI Monograph MNO-10, Earthquake Engineering Research Institute, Oakland, California
- NRC (2015). National Building Code of Canada. National Research Council Canada, NRC, Ottawa, Canada.
- Newark, M.J. (1984). Canadian tornadoes, 1950–1979. *Atmosphere-Ocean*, 22(3), 343-353.
- Newark, M.J. (1991). A design basis tornado. *Canadian Journal of Civil Engineering*, 18(3), 521-524.
- Perez, A.H., Wicker, L.J., & Orville, R.E. (1997). Characteristics of cloud-to-ground lightning associated with violent tornadoes. *Weather and Forecasting*, 12(3), 428-437.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Reap, R.M., & MacGorman, D.R. (1989). Cloud-to-ground lightning: Climatological characteristics and relationships to model fields, radar observations, and severe local storms. *Monthly Weather Review*, 117(3), 518-535.
- Refan, M., & Hangan, H. (2016). Characterization of tornado-like flow fields in a new model scale wind testing chamber. *Journal of Wind Engineering and Industrial Aerodynamics*, 151, 107-121.
- Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media, New York.
- Roueche, D.B., Prevatt, D.O., & Haan, F.L. (2020). Tornado-induced and straight-line wind loads on a low-rise building with consideration of internal pressure. *Frontiers in*

built environment, 6, 18.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461-464.

Shephard, M.W., Morris, R., Burrows, W.R., & Welsh, L. (2013). A high-resolution Canadian lightning climatology. *Atmosphere-ocean*, 51(1), 50-59.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis* (Vol. 26). CRC press.

Sills, D., Cheng, V., McCarthy, P., Rousseau, B., Waller, J., Elliott, L., Klaassen J., & Auld, H. (2012). Using tornado, lightning and population data to identify tornado prone areas in Canada. In *Preprints, 26th Conf. on Severe Local Storms, Nashville, TN, Amer. Meteor. Soc. P* (Vol. 59).

Sills, D.M., Scriver, S.J. & King, P.W.S. (2004). The tornadoes in Ontario project (TOP). In *Preprints from the 22nd AMS Conference on Severe Local Storms, Hyannis, Mass, American Meteorological Society, CD-ROM B* (Vol. 7).

Tan, L. (2008). Assessment of tornado hazard maps for southern Ontario, M.E.Sc. Thesis, the University of Western Ontario, London, Ontario, Canada.

Tan, L., & Hong, H.P. (2010). Influence of spatial inhomogeneity of tornado occurrence on estimated tornado hazard. *Canadian Journal of Civil Engineering*, 37(2), 279-289.

Twisdale, L.A. (1978). Tornado data characterization and windspeed risk. *Journal of the Structural Division*, 104(10), 1611-1630.

Twisdale, L.A., Dunn, W.L., Davis, T., & Horie, Y. (1981). Tornado missile simulation and design methodology. EPRI NP-2005.

Twisdale, L.A., & Dunn, W.L. (1983). Probabilistic analysis of tornado wind risks. *Journal of Structural Engineering*, 109(2), 468-488.

Wikle, C.K., & Anderson, C.J. (2003). Climatological analysis of tornado report counts

using a hierarchical Bayesian spatiotemporal model. *Journal of Geophysical Research: Atmospheres*, 108(D24).

Wurman, J., & Gill, S. (2000). Finescale radar observations of the Dimmitt, Texas (2 June 1995), tornado. *Monthly weather review*, 128(7), 2135-2164.

Yarbrough Jr, J.W., & Meentemeyer, V. (1978). Seasonal and regional variation in the correlation of thunderstorm days with tornado frequency. *Journal of Applied Meteorology*, 17(11), 1741-1746.

Chapter 5

5 Conclusions and future research

5.1 Conclusions

The contribution of the present study is threefold. First, a new predicting model is developed which takes into account the significant population bias effect and overdispersion feature of data. Second, a new wind hazard map over the southern Ontario (highly populated area) and whole Canada is obtained based on the newly developed spatial occurrence model. Third, a simple equivalent parametric along-height wind profile is developed based on the 10m-high wind speed, which greatly facilitates the evaluation of tornado-induced wind loading.

For the first point of contribution, a review of the existing methodologies is done comprehensively. Although three spatial point processes, i.e. Poisson, zero-inflated Poisson and negative binomial, have been investigated for US tornado data, only Poisson process was applied to Canada as per the existing literature. As such, the above-mentioned three models are revisited and applied to Canada. The original Bayesian hierarchical model is accommodated into a simple GLM framework and this facilitates the model parameter estimation using maximum likelihood method. It is found that among the three models the NB model is preferred and that utilizing both the ACGLF and ATD as covariates outperforms utilizing only ACGLF or ATD as covariate in the sense of achieving lower AIC or BIC. A side discovery is that for model parameter estimation, there is almost no difference between maximum likelihood method and Bayesian MCMC method if the latter uses “non-informative” prior distributions for the parameters. However, the simplicity of maximum likelihood method makes it a preferred choice in the practice.

For the second point of contribution, a new tornado occurrence model utilizing the NB model is first developed for the southern Ontario and whole Canada. With the characteristics of tornado summarized in the literature, a Monte-Carlo-simulation-based approach is adopted to obtain the striking rates of tornado at different sites (grid points). With those striking rates, the T-year return wind velocities at 10m high could be calculated

for the considered sites. To this end, a hazard map based on the T-year return wind velocity at 10m height is obtained. The T-year return wind velocity at 10m height plays an important role in characterizing the whole along-height wind profile and thus could be used as a proxy for the hazard assessment, which is investigated in detail in the third point of contribution.

For the third point of contribution, various T-year return wind profiles are first established for a couple of selected sites (tornado-prone cities). Then, a simple parametric form is found to fit the profiles with the wind velocity at 10m height as the key parameter. To make such a parametric form applicable to the whole country, the connection between the other embedded parameters and the wind velocity at 10m height is examined and established. However, it is found that applying the parametric wind profile would result in over-estimate of bending moment and shear force. To this end, a reduction factor, which is expressed as a function of the wind velocity at 10m height, is found to lower the discrepancy between the parametric-profile-induced and T-year return bending moments or shear forces. After being equipped with the reduction factor, it is shown that the bending moment and shear force which are calculated from the developed parametric wind profile could approximate the T-year return bending moment and shear force very well and in a conservative manner.

5.2 Future research

The methodologies and analysis framework developed and presented in this thesis could be applied to other regions in the world. The idea of “equivalent” wind profile could be generalized based on more safety requirements and professional engineering considerations.

As mentioned in Chapter 4, although not elaborated, the extension of the analysis framework by considering combinations of different tornado occurrence rate models, tornado intensity models, and tornado wind field models is straightforward based on the logic tree approach commonly used for seismic hazard mapping (McGuire 2004). For the selection of covariates or explanatory variables in the predicting model, it is understandable

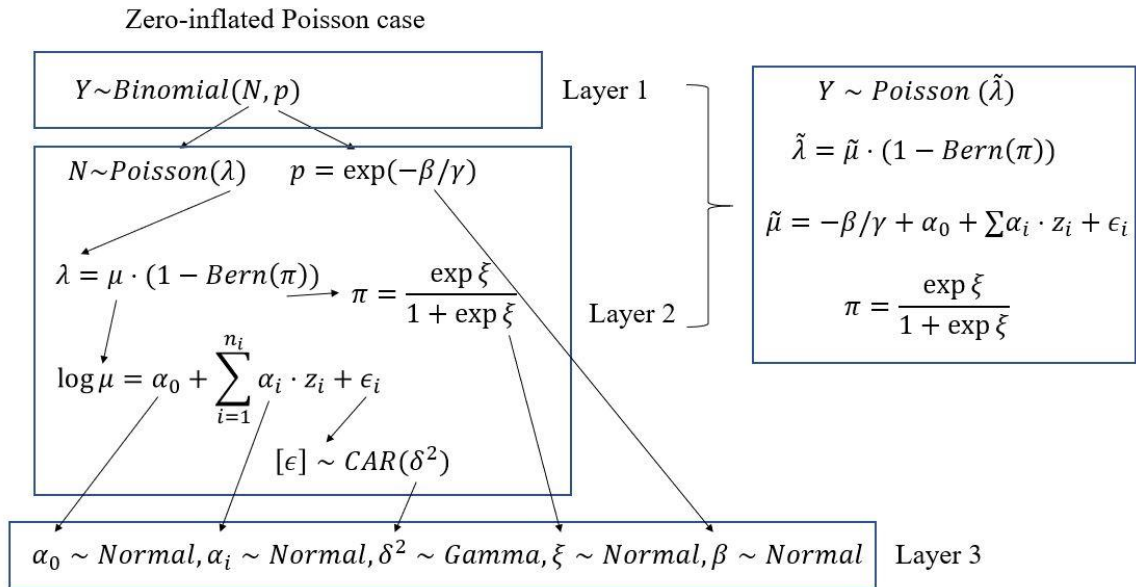
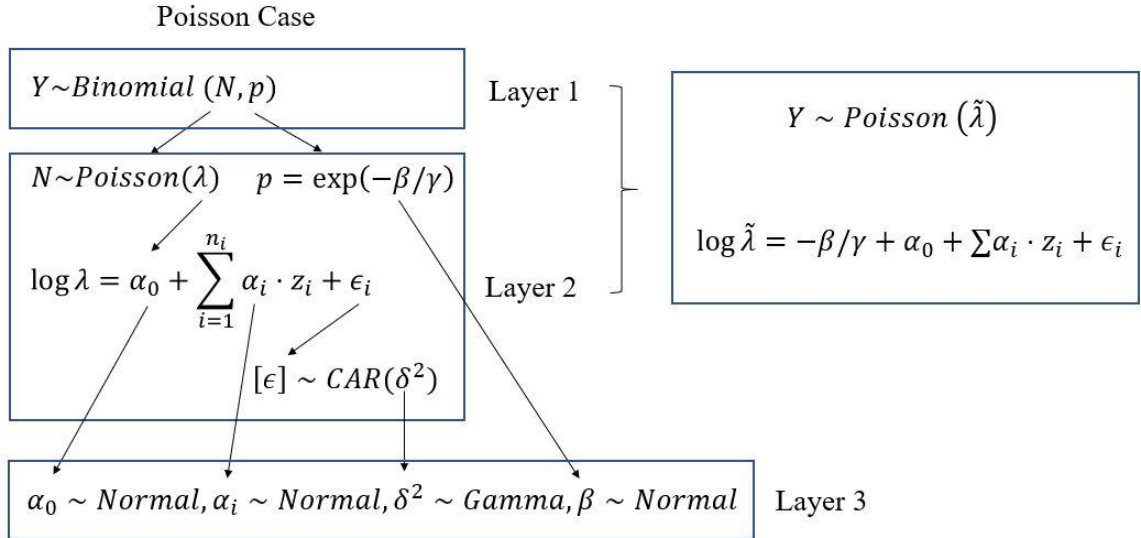
that a refined classification of the meteorological information could facilitate the prediction of occurrence of tornadoes with different intensities. It is also worth exploring other meteorological variables to improve the accuracy of prediction.

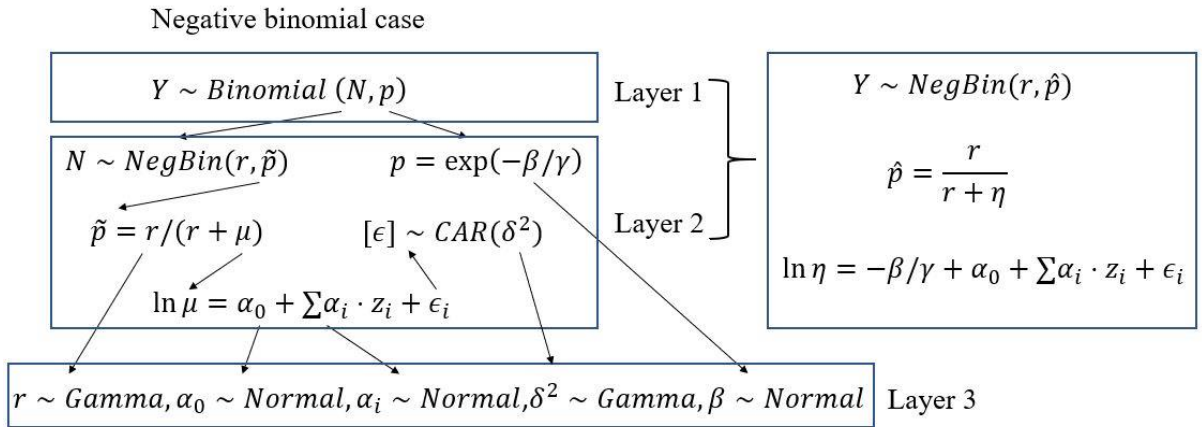
The lack of sufficient high-quality data on the tornado path orientation, length, and width for Canada indeed limits the current study of developing a universally applicable model for Canada. For example, it is observed that the tornado orientation in southern Ontario has a preference and this preference may vary from province to province based on the local environment. The probabilistic model for the tornado orientation could affect the accuracy of the estimated tornado hazard for infrastructure extended in many kilometers (e.g., transmission lines). Therefore, for large spatially distributed infrastructure system, a more detailed and project-specific analysis may be required.

In the present study, due to the limited resource, the wind field model as well as the parameters and their distributions from Twisdale et al. (1981) and Banik et al. (2007) are still utilized. There is a need for an update of the tornado path and wind field models by incorporating new tornado damage survey information, the tornado database available from the U.S., and the wind field characteristics observed from full scale and laboratory tests after years of research. The use of such updated models for tornado wind velocity hazard modelling can enhance our understanding of the tornado hazard for Canada.

Appendices

Appendix A: Structure of the Bayesian hierarchical model





Appendix B: Convergence issue with different probability detection model in MCMC

By adopting the Poisson model and the probability of detection $p_{ij} = \exp(-\beta / \exp(\gamma_j))$ as in Cheng et al. (2013) and set the length of each Markov chain to be 20000, the estimation of the model parameters that is carried out in Chapter 4 is repeated. The results are summarized in the following table and figures. Fig. A.1 shows that the generated markov chain fails to pass the BGR test as the ratio of within-chain and between-chain variabilities does not converge to one. Fig A.2 shows that the generated sample values for α_0 and α_1 are not stable. All these indicate that there is a convergence issue by adopting the setting of Cheng et al. (2013) in the present study.

Table A.1. The mean and standard deviation of model parameters for grid system whose grid size is $30 \times 30 \text{ km}^2$.

	β	α_0	α_1	σ^2
Mean	2.3707	-2.9045	1.7208	0.7835
SD	0.1313	0.2836	0.0978	0.2206

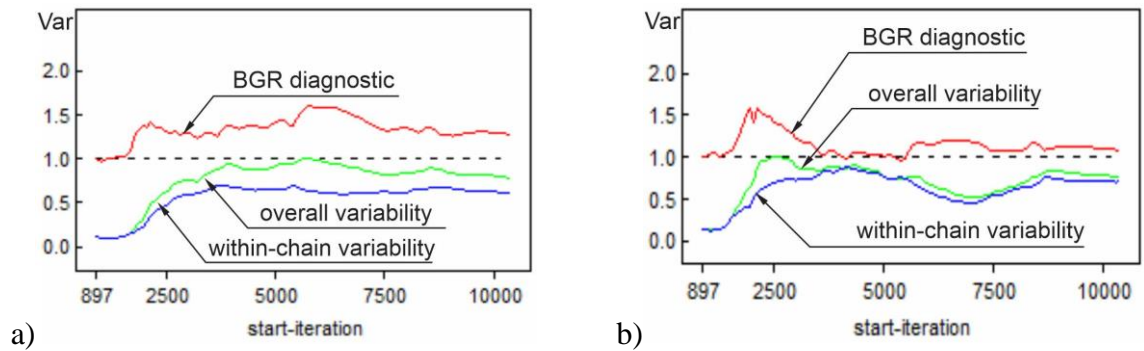
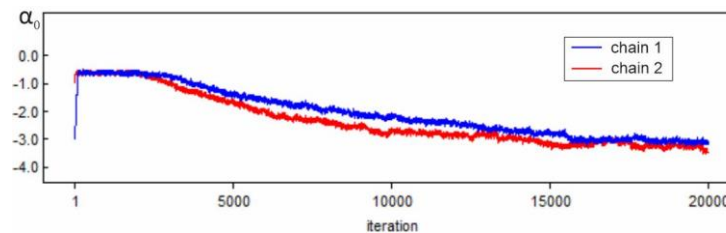


Figure A.1. The BGR diagnostics for Markov chain convergence: a) for α_0 ; b) for α_1 .



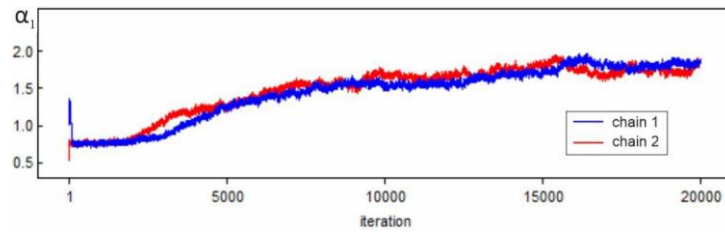


Figure A.2. The generated sample paths of α_1 and α_2 .

Appendix C: Tornado statistics for southern Ontario

This appendix contains probabilistic models adopted from the literature on the tornado path characteristics for southern Ontario.

Table B.1. Tornado path direction in southern Ontario (Banik et al., 2007).

Direction	N	NE	E	SE	S	SW	W	NW
Frequency	0.0422	0.378	0.4488	0.1084	0.0075	0.003	0.0015	0.0106

Table B.2. Distribution parameters for the tornado path length and width (Banik et al., 2007).

F scale	Length		Width	
	<i>a</i>	<i>b</i>	<i>a</i>	<i>B</i>
F0	0.9150	0.6442	0.0450	0.8812
F1	0.5015	0.6618	0.0229	0.9147
F2	0.2965	0.6563	0.0223	0.8124
F3	0.0786	0.8282	0.0022	1.1271
F4	0.0264	0.9865	0.0028	0.9591
F5	0.0005	1.8758	1.18E-10	3.6684

Table B.3. Intensity variations along tornado track length (Twisdale, 1981).

		Touch-down intensity					
		F0	F1	F2	F3	F4	F5
% length	F0	1.00	0.57	0.28	0.12	0.14	0.13
	F1		0.43	0.35	0.25	0.16	0.10
	F2			0.37	0.32	0.28	0.19
	F3				0.32	0.21	0.24
	F4					0.21	0.19
	F5						0.15

Curriculum Vitae

Name: Qian Huang

Post-secondary Education and Degrees: Shenzhen University
Shenzhen, China
2009-2013 B.Sc.

The University of Western Ontario
London, Ontario, Canada
2014-2015 M.Eng.

The University of Western Ontario
London, Ontario, Canada
2016-2020 Ph.D.

Related Work Experience Research Assistant
The University of Western Ontario
2016-2020

Teaching Assistant
The University of Western Ontario
2016-2020

Publications:

- [1] Hong, H.P., & Huang, Q. (2020). Simplified Hazard Modelling and Structural Reliability Analysis Considering Non-Synoptic Wind Systems (NSWS), in Handbook of Non-Synoptic Windstorms edited by Horia Hangan and Ahsan Kareem, Oxford press.
- [2] Huang, Q., Jiang, W.J., & Hong, H.P. (2020) Statistical assessment of spatial Tornado occurrence for Canada: modelling and estimation, submitted.
- [3] Huang, Q., Jiang, W.J., & Hong, H.P. (2020) Development of a simple equivalent tornado wind profile for structural design and evaluation, submitted.