De-aggregated Hazard of Freezing Rain Events

R. Erfani and Dr. L. Chouinard
Department of Civil Engineering and Applied Mechanics
Macdonald Engineering Building, 817 Sherbrooke Street West Montreal, Quebec H3A 2K6.
reza.erfani@mail.mcgill.ca

1   Introduction

Freezing precipitation and wind are the main climatic loads governing the design of electrical transmission lines. Wind loads are covered relatively comprehensively in the literature and codes; however, the uncertainty associated with ice loading is still large. Interest in this research area in Quebec has increased recently with the occurrence of events such as the 1998 ice storm.

A number of statistical approaches have been proposed to estimate ice loads which are based on the analysis of direct measurements (El-Fashney 2003 and Laflamme 1995) and empirical models (Makkonen 1998, Jones 2002, Jones 1998). These approaches are based on the premises that the characteristics of ice accumulation and wind speed are location dependant but that ice storms, and associated winds, can be modelled as events from a single population. However, there appears to be a great variability in freezing rain events, in terms of duration, quantity of precipitation, and wind speed. Rauber (2001) proposes a characterization of freezing rain events into seven archetypical patterns based on a historical review of ice storms that occurred in United States.

In this research the problem is studied by categorizing freezing rain events using various meteorological variables and statistical methods. The icing hazard for each type of storm is analyzed separately and combined to obtain icing hazard curves at various locations in Quebec. In this paper several aspects are discussed: the clustering procedure and the physical meaning of the results; and the de-aggregated extreme value analysis of total precipitation.

2   Objective

The objective of the work presented here is to improve estimates of atmospheric icing hazards by grouping freezing rainstorms into different groups using statistical methods. These groups are then retained only if they can be associated with plausible physical mechanisms. The groups are used to develop a de-aggregated hazard curve for total precipitation and associated wind speed. The procedure is then applied to develop de-aggregated hazard curves for the Montreal area.

3   Theory & Background

3.1   Archetypical Categories

Rauber (2001) compiled a list of 411 ice storms that occurred in the United States, east of the Rocky Mountains, during the period of 1970 to 1994. The data set includes storm duration, location where freezing rain was observed (Vermont, Pennsylvania etc.), and a category of archetypical weather patterns for each 12 hour period. The different patterns are discussed below. The freezing precipitation events were identified using NCDC (National Climatic Data Centre) monthly reports from 1970 to 1994 for the periods from October to April. Three-hour surface charts corresponding to freezing rain events were then compiled from NCDC microfilms and analyzed to categorize each weather system.

Weather systems are categorized into one of seven typical patterns (A through G) during ice storms for Eastern North America: Pattern A, 'Arctic Front/Anticyclone': An arctic front occurs ahead of arctic anticyclones, which develop in north western Canada and travel southeast. No particular cyclone is associated with the anticyclone. The warm air rises above the advancing cold air creating a temperature inversion. Owing to the absence of a strong low pressure cyclone, high winds and high precipitation would not be expected with this pattern;

Pattern B, 'Warm Front—Occlusion Sector of Cyclones': In this pattern, freezing precipitation occurs north of the warm front as a result of the frontal overrunning (warm air overrunning cold air). The presence of the cyclone carries warm moist air

from its south-eastern quadrant northward which favours higher levels of precipitation;

Pattern C, 'Cyclone/Anticyclone': Pattern C is essentially a combination of patterns A and B occurring simultaneously. The strong pressure gradient between the high and low pressure centers can create vigorous storm activity resulting in strong winds. For patterns A, B, and C precipitation typically occurs along a long narrow band. Rauber also reports that patterns A and C have the longest successive 12hr patterns. In consequence, the long duration storms are generally associated with these patterns;

Pattern D, 'Western Quadrant of Arctic High Pressure': The passage of an anticyclone can create a cyclone west to northwest of its center. The freezing precipitation will occur deep within the artic air mass in a circular area rather than a narrow band;

Pattern E, 'Cold Air Damming': This pattern develops when the arctic air mass moves well over the north Atlantic, warm Atlantic air rises above the cold air trapped against the east side of the Appalachian mountains;

Pattern F, 'Cold Air Damming with Atlantic Cyclone': An Atlantic cyclone can combine with Pattern E to create Pattern F. The presence of the cyclone has the effect of intensifying the pressure gradient and enhancing the amount of precipitation and wind speeds.

Pattern G, 'Cold Air Trapping': A Cyclone originating east of the Rockies can track east. Warm air circulating counter clockwise around the southern part of the cyclone can overrun the cold air of an Atlantic anticyclone trapped within the Appalachian Mountains.

3.2   Data

Two data are used in this work: NCEP reanalysis data (kalnay (1996)), and Environment Canada daily and hourly meteorological observations. Data from both Trudaeu (formerly Dorval) airport and St-Hubert airport are used. The period covered by the data set is from 1954 to 2004

NCEP reanalysis data was used to compile a data set on geopotential heights (specifically 1000 mb, 925 mb, and 500 mb), surface pressures, and wind speeds during ice storm events. Data is available on a grid of 2.5 degree latitude and 2.5 degree longitude, dating back to the 1948

3.3   Clustering Storms

A list of ice storms was developed by scanning the hourly observations for the occurrence of freezing rain at each airport. Individual storms were defined as each set of successive days of observation of freezing rain. Note that this definition does not correspond to the definition sometimes used by electric utilities where the notion of persistence of ice deposits on a surface is also considered. The total precipitation and average wind speed at Dorval and St-Hubert airports were compiled from the Environment Canada data for each storm.

Total precipitation is used as a measure of storm severity to simplify comparison of single population results with clustered results given the uncertainty and variability in observed ice accumulations. Analysis with an icing model would both reduce uncertainty and variability on estimates of ice accumulations.

The clustering process is a statistical procedure that groups individual storms into homogenous groups on the basis of common characteristics. For this purpose, each individual storm is characterized by maps of average anomalies for various storm characteristics derived from the NCEP reanalysis project. For each storm, an anomaly map is made for average sea level pressure (SLP). The anomaly is defined as the average value (at each grid point) during the storm minus the monthly average (calculated at each grid point for the period of 1954 to 2004). This method was chosen to study the storms following the work of Gyakum (2001). However the area used in the study is closer to that used by Rauber: from 250° to 310° W longitude, and 25° N lat to 60° N latitude.

Several multivariate statistical analysis methods can be used to perform cluster analysis or pattern analysis. Among these, Principal Component Analysis (PCA) is among one of the most popular in atmospheric sciences (Preisendorfer (1988). The purpose of PCA is to identify and rank by order of decreasing importance a set of uncorrelated linear combinations of the original variables that explain most of the variability in a data set. When applied to spatial data sets, the procedure will identify the dominant and recurrent spatial patterns for the variation of a physical characteristic.

Principal component analysis was performed on a data set consisting of the average anomaly values of SLP of all storms at 312 grid points. The first 10 principal components explain up to 90% of the variability which represent a compression of 3.2% of the original data set, and

are representative of the entire data set. The components are used to cluster the storms in homogeneous groups. The clustering is done using the k-means algorithm (Dillon (1984)). The k-means algorithm requires the definition of a metric that measures the variability between clusters and within clusters. In this case a Euclidean distance is used. Clusters are obtained by first specifying the desired number of clusters. Storms are then assigned to each cluster in order to minimize within cluster variability while maximizing between cluster variability. Various methods are proposed in the literature in order to optimize the number of clusters. In this case, objective methods were combined with the ability to physically interpret the clusters to determine their optimal number.

Following this procedure, the optimal number of clusters was determined to be 2 or 3 clusters. These clusters correspond to groups with distinct physical interpretations that are similar to Rauber's archetypical patterns. Increasing the number of clusters produces clusters that are poorly related to known patterns and results in poorly populated clusters.

3.4 De-aggregated Hazard Analysis

De-aggregated hazard analysis is routinely used in the assessment of natural hazards, and in particular for earthquake hazard analysis. The analysis is performed by analyzing the contributions from various sources of hazards separately and combining these individual contributions to determine the total hazards at a given location. The total hazards function is then used for defining design criteria as a function of the return period.

In this application, the frequency and probability distribution functions for storms are defined using a peak-over-threshold for each cluster. Storm severity is described by the total precipitation during the storm. Then frequency versus intensity curves are obtained for each cluster and the sum of these frequencies is used to create a hazard curve. (Field)

The probability distribution function $f_i(h_o)$ is estimated for total precipitation or amount of freezing rain for each category of storm, i, as well as the corresponding annual occurrence rate, ν_i. The probability of a freezing rain event of intensity exceeding than a given value h_o is

$$1 - \int_0^{h_o} f_i(h_o)\,dh = 1 - F_i(h_o) \quad (1)$$

The annual occurrence rate is

$$\lambda_i(h_o) = \nu_i(1 - F_i(h_o)) \quad (2)$$

and the total annual rate of exceeding h_o, is obtained by summing up all the λ_i,

$$\lambda_{tot}(h_o) = \sum_{i=1}^{n} \lambda_i(h) = \sum_{i=1}^{n} \nu_i(1 - F_i(h_o)) \quad (3)$$

where n is the number of clusters. Plotting λ_i provides h_o makes disaggregated frequency-intensity curves.

The probability of exceeding h_o during a time period t can be estimated using a Poisson distribution..

$$P(h > h_o, t) = 1 - \exp(-\lambda_{tot} t) \quad (4)$$

Plotting P versus h_o provides a hazard curve for a given time period t.

Similarly, the corresponding return period is calculated as

$$T(h > h_0) = \frac{1}{P(h > h_o, t = 1)} = \frac{1}{1 - \exp(-\lambda_{tot})}$$
$$(5)$$

For small values of λ_{tot} the equation becomes

$$T = \frac{1}{\lambda_{tot}} \quad (6)$$

3.5 Distribution Function

The peaks over threshold method requires to specify a minimum threshold value for storm severity. The corresponding annual occurrence rate ν_i is simply equal to the total number of events meeting the criteria divided by the time period. The generalized Pareto distribution (GP) is often used in combination with the POT method (Jones 2002, Wang 1991, Picklands 1975). Here, both the GP and generalized extreme value (GEV) distribution are used in the analysis. El-fashney (2003) compared a number of extreme value distributions using POT and found that three parameter distributions performed well on average for ice accumulation during ice storms.

The parameters of the GP and GEV are estimated using L-moments (Hosking and Wallace 1997). The goodness of fit of the distributions is evaluated uses both maximum absolute error (MAE) and root mean square error (RMSE).

$$RMSE = \left[\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{(n-m)} \right]^{1/2} \quad (7)$$

$$MAE = \max|x_i - y_i| \quad (8)$$

where x_i are the observations and y_i are the predicted value using the probability distribution function with the corresponding exceedance probability using the Cunnane plotting position. The total number of observations is represented by n, and m is the number of parameters of the distribution function.

$$P_{i,n} = \frac{i - 0.4}{n + 0.2} \quad (9)$$

where i is the rank of an observation.

Thus,

$$y_i = F^{-1}(P_{i,n}) \quad (10)$$

4   Results

The results presented here are for storms compiled for the Montreal area. The storms are identified using the data from Trudeau airport. The values are averaged with the observations from St Hubert airport.

The results presented are for a 10mm threshold. Using a smaller threshold introduces too many events that have no relevance in terms of hazards. Eliminating smaller values generally results in better statistical fits, and greater distinction between mean precipitation values of clusters. However, using a larger threshold results in poor fits as the clusters have too few data. However, since high quantile estimates are usually insensitive to small values, efficient high quantile estimates can be obtained at reasonably high thresholds (Wang 1991).

4.1   Physical   Interpretation   of Patterns

Figures 4-1 and 4-2 show the average map of SLP anomalies for 2 clusters. Fig 4-1 indicates relatively high positive SLP anomalies over most of

Ontario, Quebec, and particularly over the Maritime Provinces. Positive values indicate high pressure systems, and negative values indicate low pressures systems. The spatial extent of the positive values and the relatively low negative anomalies (-4 hPa) is characteristic of a Type A pattern. Figure 4-2 is reversed. The lack of high positive anomalies and the presence of high negative anomalies suggest a Type B pattern. Table 4-1 shows the averages and standard deviations for the total precipitation and maximum wind speeds for the storms associated with each cluster. Cluster 1 (containing 58 storms) has a higher average precipitation than cluster 2 (containing 91 storms); however the standard deviation is quite high. Note that the 1998 ice storm is included in the first cluster.
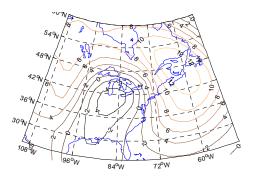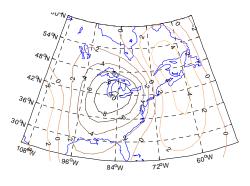

Figure 4-1: SLP (hPa) Anomaly Cluster 1/2

Figure 4-2: SLP (hPa) Anomaly Cluster 2/2

Table 4-1: Cluster Averages 10mm threshold

Cluster	No.	Mean Precip. (mm)	Std. Dev. Precip (mm)	Mean Max Wind. (m/s)	Std. Dev. Max Wind. (m/s)
Single	149	21.5	11.5	37.4	12.3
1/2	58	23.4	15.8	38.2	13.0
2/2	91	20.3	7.4	36.8	11.9
1/3	44	24.3	15.8	37.9	11.7
2/3	49	21.4	7.6	37.9	13.3
3/3	56	19.3	9.9	36.5	11.9

Figures 4-3 to 4-5 show the average cluster map of the SLP anomalies for 3 clusters. Cluster 1/3 has a high positive anomaly centered over north eastern Quebec, and a moderate negative anomaly. The strong pressure gradient suggests that many pattern C storms form this cluster. Cluster 2/3 is similar to an average anomaly map of pattern B. Cluster 3/3 is similar to cluster 1, but shows a weaker pressure gradient. There is greater difference in the values of average total precipitation for 3 clusters. Cluster 2 exhibits the greatest precipitation, greater than the single population average. This is consistent with the expectation that more vigorous storms would be observed in a pattern C synoptic system. The 1998 ice storm is included in this cluster. The third cluster has a lower average precipitation. This is again consistent with there being a weaker pressure gradient than for the other clusters.
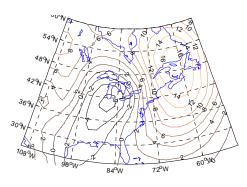

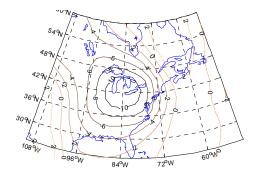**Figure 4-3 : SLP (hPa) Anomaly Cluster 1/3**

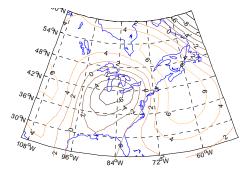**Figure 4-4: SLP (hPa) Anomaly Cluster 2/3**

Figure 4-5: SLP (hPa) Anomaly Cluster 3/3

4.2 Extreme Value Analysis

Figure 4-6 shows both the MAE error and RMSE for both GEV and GP for a single population. When comparing the values for single populations, the GEV distribution produces better fits. For this data, the GEV distribution performs quite well. This is not always the case. Similar plots for Quebec City, indicate that the GEV does not provide the best fit Erfani (2007).

Figure 4-7 presents a similar plot for the MAE of 3 clusters and the single population. The GP plot for the single population has the highest MAE until a threshold of 20mm. At a threshold value of 6mm and above, at least one of the distribution functions (GP or GEV) for a cluster performs better than the single population GEV function. The clusters are named according to their maximum precipitation value. The worst fits are for the cluster which contains the 94mm (93.5mm) accumulation that occurred during the 1998 ice storm. Extremely good fits are obtained for the other clusters. Plots for the RMSE are not so favourable (Figure 4-8); nevertheless the errors are below 3mm. Owing to the fewer number of observations in each cluster this is expected. However the fit of the larger values is good.
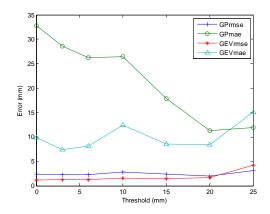
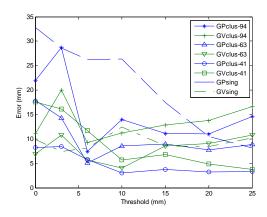**Figure 4-6: RMSE and MAE vs. Threshold for GP and GEV and a Single Distribution.**

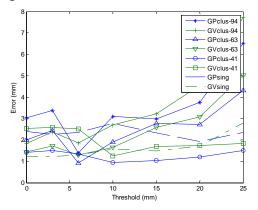**Figure 4-7: MAE vs. Threshold for GP and GEV Single Dist and Clusters**

Figure 4-8: RMSE vs. Threshold for GP and GEV Single Dist and Clusters

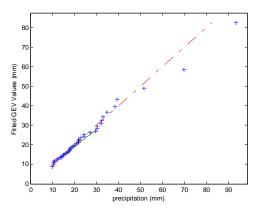**Figure 4-9: QQ plot Cluster 1/3 (max value 94mm) threshold=10mm, GEV.**

Figure 4-10: QQ plot Cluster 2/3 (max value 41mm) threshold=10mm, GP.

Figure 4-11: QQ plot Cluster 2/3 (max value 63mm) threshold=10mm, GEV.

Figures 4-9 to 4-11 show the QQ plots for each cluster at a 10mm threshold. The best fit for the 1998 ice storm is obtained for a 6mm threshold (Figure 4-12).

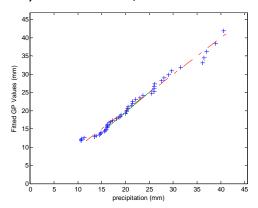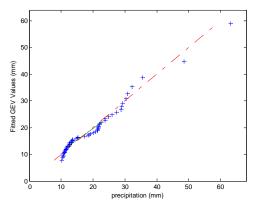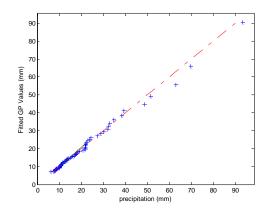**Figure 4-12: QQ plot Cluster 1/3 (max value 94mm) threshold=6mm, GP.**

Figure 4-13: Return Period vs. Threshold, for Best Fit De-aggregated Analysis and Single Distributions.

Figure 4-13 shows a plot of the estimated return periods at different thresholds for the GEV (GEVS) and GP (GPS) single distribution and the best fit results for the de-aggregated distributions (BF) with 3 clusters. The return periods are calculated for 94mm (as it represents the 1998 storm) and 75mm (as it corresponds to approximately the 50yr return period) of total precipitation. The GP distribution produces very high return periods at low thresholds, but converges to the results for the GEV at high thresholds. This result is found by others when using the GP distribution (Hosking and Wallis 1987), and is probably a result of the high errors at low thresholds. However the GEV distribution, in the POT approach, shows more consistent results. A similar result was found for data from Quebec City, Erfani (2007). The variation of return period as a function of threshold is similar in shape to the variation of MAE, showing the sensitivity of error on return period, and that a larger error overestimates the return period (in this case). The de-aggregated analysis produces the most stable results as a function of threshold and predicts return periods of approximately 100 years and 50 years for 94mm and 75mm respectively.
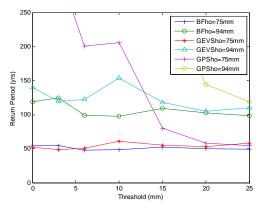
5   Conclusions

A procedure has been presented to perform a de-aggregated analysis of atmospheric icing hazards. The premise of the procedure is that ice storms can be the result of different meteorological conditions that produce storms with widely varying characteristics. In the latter case, it would be a mistake to treat icing observations as being from a single population.

The proposed procedure uses SLP data during ice storms to characterize storm patterns. Individual storms are then assigned to a homogenous group of storms using statistical clustering techniques. Hazard analysis is then performed by summing up the contributions for each homogenous group of storms.

The procedure was applied to the Montreal area and the following observations were made:

1) Using a POT approach combined to the GEV distribution produces optimal and robust results as a function of the threshold value.

2) The deaggregated analysis also introduced robustness in the estimation of the return period as a function of the threshold, especially at high return periods.

References

El-Fashny, K., (2003), "Modeling Ice Loads". Ph.D. Thesis, Dept. of Civil Engineering, McGill University.

Dillon, William R., (1984)." Multivariate analysis : methods and applications". Wiley. New York. NY, U.S.A

Erfani, R. (2007), "De-aggregated Hazard of Freezing Rain Events". Ph.D. Thesis, in progress, Dept. of Civil Engineering, McGill University.

Field, E. H. "Probabilistic Seismic Hazard Analysis (PSHA) tutorial", URL: www.relm.org/tutorial_materials/PSHA_Primer/index.html. field@usgs.gov.

Gyakum J. R., Roebber P. (2001). "The 1998 ice storm-analysis of a planetary-scale event", Monthly Weather Review. Washington. Vol. 129, Iss. 12; p. 2983 (15 pages)

Hosking; J. R. M., Wallis, J. R. (1987) "Parameter and Quantile Estimation for the Generalized Pareto Distribution", Technometrics, Vol. 29, No. 3., pp. 339-349.

Hosking, J. R. M., and Wallis, J. R. (1997). "Regional frequency analysis: an approach based on L-moments". Cambridge University Press, Cambridge, U.K.

Jones, K. Thorkildson, B., Lott, N., (2002), "The Development of an US Climatology of Extreme Ice Loads", NOAA/ NESDIS National Climate Centre.

Jones, K. (1998), "A simple model for freezing rain ice loads". Atmospheric Research 46, 87–97.

Kalnay et al.,(1996) "The NCEP/NCAR 40-year reanalysis project", Bull. Amer. Meteor. Soc., 77, 437-470.

Laflamme, J. (1995) "Spatial variation of extreme values for freezing rain", Atmospheric Research. Elsevier Science, 36: 195-206.

Makkonen, L., (1998) "Modeling power line icing in freezing precipitation", Atmospheric Research, Volume 46, Issues 1-2, Pages 131-142.

Preisendorfer, R. W. (1988). "Principal component analysis in meteorology and oceanography". Elsevier, New York, NY, U.S.A.

Rauber R. M., Olthoff L. S, Ramamurthy M. K., Miller D., Kunkel K. E.. (2001) "A synoptic weather pattern and sounding-based climatology of freezing precipitation in the United States east of the Rocky Mountains", Journal of Applied Meteorology. Boston. Vol. 40, Iss. 10; p. 1724 (24 pages)

Wang, Q. J. (1991), "The POT model described by the generalized Pareto distribution with Poisson arrival rate", Journal of Hydrology, Volume 129, Issues 1-4, Pages 263-280