Isotropic and anisotropic kriging approaches for interpolating surface-level wind speeds across large, geographically diverse regions

Carol J. Friedland, T. Andrew Joyner, Carol Massarra, Robert V. Rohlic, Anna M. Treviño, Shubharoop Ghosh, Charles Huyck and Mark Weatherhead

ABSTRACT
Windstorms result in significant damage and economic loss and are a major recurring threat in many countries. Estimating surface-level wind speeds resulting from windstorms is a complicated problem, but geostatistical spatial interpolation methods present a potential solution. Maximum sustained and peak gust weather station data from two historic windstorms in Europe were analyzed to predict surface-level wind speed surfaces across a large and topographically varied landscape. Disjunctively sampled maximum sustained wind speeds were adjusted to represent equivalent continuously sampled 10-minute wind speeds and missing peak gust station data were estimated by applying a gust factor to the recorded maximum sustained wind speeds. Wind surfaces were estimated based on anisotropic and isotropic kriging interpolation methodologies. The study found that anisotropic kriging is well-suited for interpolating wind speeds in meso- and macro-scale areas because it accounts for wind direction and trends in wind speeds across a large, heterogeneous surface, and resulted in interpolation surface improvement in most models evaluated. Statistical testing of interpolation error for stations stratified by geographic classification revealed that stations in coastal and/or mountainous locations had significantly higher prediction errors when compared with stations in non-coastal/non-mountainous locations. These results may assist in mitigating losses to structures due to excessive wind events.

1. Introduction
Twice-daily upper level wind speeds interpolated from coarse networks of radiosonde-based observations are commonly used to estimate surface wind speeds in windstorms (e.g. Della-Marta et al. 2009), but this type of estimation from atmospheric sampling does not adequately reflect the complexity of winds at the surface. Surface-level meteorological station data best depict wind fluctuations at the local scale and are therefore most appropriate for surface wind speed estimation (Smits et al. 2005). However, the use of station data to develop a footprint or interpolated surface of wind speeds also presents unique issues. First, most sustained wind speeds are disjunctively sampled at
weather stations operated under World Meteorological Organization (WMO) standards, typically sampled 10 minutes prior to each hour, resulting in historical records with 50 minutes of unknown sustained wind speeds each hour. Only a handful of studies have addressed the effects of disjunct sampling for extreme wind speeds (e.g. Larsen & Mann 2009; Gatey 2011). Second, many stations either do not collect peak gust wind speeds or these data are not reported for all sample times. Estimation of peak gust values is therefore necessary when peak gust data are missing, to obtain the ‘probable maximum wind speed’ over a shorter averaging period, because buildings and other structures are most affected by wind gusts of approximately 3 seconds in duration (Krayer & Marshall 1992). The Durst (1960) method of applying a gust factor has proven to be an accurate method for estimating wind speeds of different averaging times in tropical cyclones (Krayer & Marshall 1992) — similar to mid-latitude cyclones, or windstorms, in this study. Once properly adjusted, station data provide a time history of wind speed magnitude and direction; however, the final issue is that spatial resolution of stations is coarse in many areas and wind speeds between stations are unknown. This necessitates not only the development of a continuous wind speed surface via interpolation, but also a surface that maintains low error levels in areas of high wind speed variability while maintaining accurate estimates elsewhere. Spatial interpolation methods produce either deterministic, geostatistical, or knowledge-based estimates of unknown variables across a surface and must be adapted to the specific variable(s) being evaluated (Luo et al. 2008).

Deterministic methods create surfaces from measurements at specific locations either globally (i.e. calculating predictions using the entire dataset) or locally (i.e. calculating predictions using a neighbourhood within a dataset). Common deterministic interpolation methods include polynomial regression, triangular irregular network, nearest neighbour, splines, and inverse distance weighting (IDW). Geostatistical, or stochastic, methods use a probabilistic approach for data regularization (Lanza et al. 2001; Cellura et al. 2008; Zlatev et al. 2010). Common geostatistical interpolation methods include ordinary kriging, universal kriging, and cokriging, which calculate the spatial autocorrelation between measurements and also use the spatial structure of measurements around the prediction location. Geostatistical methods create an unbiased surface where a polynomial function has not been forced to fit as is the case for deterministic methods (e.g. IDW), thus reducing spurious edge and bullseye effects common in other interpolation methods (Akkala et al. 2010). Knowledge-based methods of interpolation typically include multilevel statistical applications (e.g. neural networks, Bayesian modelling) and require intricate knowledge of the study area, which is generally very localized (e.g. Öztupal 2006; Cellura et al. 2008; Li and Shi 2010; Carta et al. 2011). For these reasons, it is difficult to apply knowledge-based methods across large study areas.

Wind speed surface interpolation results suggest that deterministic methods should be avoided because they fail to account for spatial autocorrelation (Bentamy et al. 1996; Phillips et al. 1997; Sterk and Stein 1997; Venäläinen & Heikinheimo 2002; Öztupal 2006, Cellura et al. 2008; Luo et al. 2008; Zlatev et al. 2009; Akkala et al. 2010; Zlatev et al. 2010) and various forms of kriging have been shown to outperform other methods for interpolation of surface-level wind speeds (Lanza et al. 2001; Luo et al. 2008; Akkala et al. 2010; Zlatev et al. 2010). Although kriging has been shown to improve interpolation results, most previous studies that examined kriging focused on local- to regional-scale wind surfaces within a single country. Interpolated surfaces for meso- and macro-scales (i.e. large areas), which include multiple countries over geographically diverse regions, have not been studied extensively.

Kriging uses probability and spatial correlation to create a surface that is weighted by observed values through a distance- and direction-based semi-variance function that can account for anisotropic spatial patterns and trends in wind behaviour (Luo et al. 2008). Isotropy (uniform values in all directions) is assumed during the kriging process unless anisotropy is specified. Consequently, comparisons between isotropic and anisotropic semivariogram-derived surfaces are not often made. Thus far, the use of anisotropy within kriging has been shown to be superfluous for local- and regional-scale modelling, although Luo et al. (2008) hypothesized that it may be more useful for meso- and macro-scale modelling.
In addition to issues of scale and anisotropy, the impact of heterogeneous terrain (or geographic diversity) on wind speed interpolations is also poorly understood. Etienne and Beniston (2012) examined extreme station data (i.e. top 10% of wind speeds) for wind storms in Europe using ‘basic’ kriging. The results found that topography greatly influences wind speeds and likely contributed to error effects not normally seen in interpolations of larger areas. Additionally, Zlatev et al. (2010) divided the United Kingdom into five areas of homogenous terrain to analyze wind speed interpolations and reduce the potential impact of over-smoothing and over-fitting. From a standard error surface, the areas incorporating more diverse terrain contained noticeably ‘spottier’ error patterns. Joyner et al. (2015) used cokriging to interpolate wind speeds for multiple European windstorms to account for errors associated with aspect, elevation, and land cover, further alluding to an impact of geographic diversity on wind speed interpolations.

This research investigates the effectiveness of anisotropic kriging for interpolation of surface-level maximum sustained and peak gust wind speeds across an expansive, heterogeneous area through evaluation of two of the most damaging windstorms that occurred on back-to-back dates in Europe — 1999 Lothar and Martin. To accomplish this, a methodology is applied to estimate missing peak gust wind speed records from reported sustained wind speed station data and to correct for the effects of disjunct sampling of wind speeds. Ordinary kriging, with and without anisotropy, is performed to derive surface-level wind speed footprints from the two windstorm events, and the effects of inclusion of anisotropy are quantified through evaluation of cross validation average error of prediction (CVAEP), cross validation standardized standard error of prediction (CVSSEP), cross validation standard error of prediction (CVSEP), minimum/maximum wind speed ranges, and the number of ‘high-error’ stations. Finally, statistical testing is performed to evaluate differences between stations located in coastal and mountainous areas with those located in non-coastal/non-mountainous areas to evaluate the effects of geographic heterogeneity on the absolute cross validation average error of prediction (ACVAEP) for each interpolation model. Improved modelling of windstorm-induced surface winds is critical to the advancement of extreme-event disaster science and wind engineering, while also testing the limits of geospatial analytical techniques. Development of surface wind estimates through spatial analytic techniques will potentially improve modelling efforts at all spatial scales and improve our understanding of windstorm meteorology.

2. Data and methods

2.1. Wind station data

Wind data were obtained from WMO observation stations sourced through Speedwell Weather (http://www.speedwellweather.com/), a weather data compiler and provider, with support from Guy Carpenter & Company, LLC and ImageCat. The dates for data collection were 25—27 December 1999 for windstorm Lothar and 26—28 December 1999 for windstorm Martin. The WMO Standard (World Meterological Organization 2008) for measuring sustained wind is the disjunctively sampled average of values obtained for a period of 10 minutes prior to the recorded observation time. The WMO Standard for measuring peak gust is a continuously sampled average of values over a 3-second period. WMO-compliant wind instruments are to be located at a height of 10 m in open terrain, with wind data adjusted for local topographic effects through the use of a correction factor. For example, data from a station located on a hilltop would be adjusted downward to account for the local increase in wind speed caused by the station’s position. According to WMO wind standards (World Meterological Organization 2008), recorded wind speeds are to be representative of the region rather than the individual station. In this study, all data and instruments were assumed to be in accordance with these WMO standards. A preliminary quality control analysis of the data was conducted by plotting the sustained wind and peak gust values against the mean of each variable. Outliers were examined to determine whether the values were reasonable for the given climatological conditions.
Station wind data from windstorms Lothar and Martin were analyzed (Table 1). MATLAB® was used to extract sustained and peak gust wind speed values from the original data files, which contained thousands of rows of observation data for each event. The maximum sustained and peak gust wind speeds recorded at each station were extracted for each observation day and the station maxima were identified from the daily maxima. The hourly maximum sustained wind speed data were reported consistently, with only 1.4% of the data values missing for each storm. Peak gust measurements were missing for 80.0% of the hourly observations for windstorm Lothar and 90.5% of the hourly observations for windstorm Martin. Three classification schemes (CSs) were defined (Table 2) to describe the prevalent geographic classification (GC) of each station. All stations were fundamentally classified in CS 1 as: both coastal (within ~1.6 km of the coast) and mountainous (according to the Geographic Information System at the European Commission (GISCO)); coastal/non-mountainous; non-coastal/mountainous; or non-coastal/non-mountainous (Figure 1). All coastal and all mountainous stations were also defined

Table 1. Number and percentage of stations reporting wind data for each windstorm. Total numbers reflect at least one report of a wind speed type from that station, while percentage missing refers to hourly readings. Percentage GC-A, B, or C refers to the percentage of total stations located in these delineated areas.

| Storm | Countries                  | Maximum sustained | Peak gust | % missing (sustained) | % missing (peak gust) | % GC-A, B, or C |
|-------|----------------------------|-------------------|-----------|-----------------------|-----------------------|-----------------|
| Lothar| France, Switzerland, Germany | 321               | 202       | 1.4                   | 80.0                  | 50.0            |
| Martin| France, Switzerland, Germany | 324               | 195       | 1.4                   | 90.5                  | 49.8            |

Table 2. Geographic classification schemes to test for differences between coastal, mountainous, and non-coastal/non-mountainous stations.

| Classification scheme (CS) | Geographic classifications (GCs)                                      |
|----------------------------|-------------------------------------------------------------------|
| 1                          | Coastal/mountainous (A)                                           |
|                            | Coastal/non-mountainous (B)                                        |
|                            | Non-coastal/mountainous (C)                                        |
|                            | Non-coastal/non-mountainous (D)                                    |
| 2                          | All coastal (mountainous and non-mountainous) (E = A + B)         |
|                            | Non-coastal/mountainous (C)                                        |
|                            | Non-coastal/non-mountainous (D)                                    |
| 3                          | All mountainous (coastal and non-coastal) (F = A + C)             |
|                            | Coastal/non-mountainous (B)                                        |
|                            | Non-coastal/non-mountainous (D)                                    |

Figure 1. Geographic classifications within classification scheme 1 for each station that reported data during Lothar (a) and Martin (b).
in CSs 2 and 3. For both storms, approximately half of all stations (161 of 322 for Lothar and 163 of 325 for Martin) were located in non-coastal/non-mountainous areas.

To develop maximum sustained and peak gust wind speed interpolations, missing sustained or peak gust wind speed values were calculated from corresponding recorded values. Observed wind speeds were converted and adjusted in accordance with the Durst (1960) method. This study developed three data cases and procedures for adjustment (Figure 2): Case I — only gust wind speed is reported, which is used to estimate a continuously sampled 10-minute sustained wind speed using the Durst curve provided in ASCE 7-10 (ASCE 2010); Case II — only 10-minute, disjunctively sampled maximum sustained wind is reported, which is used to calculate equivalent continuously sampled 1-hour wind speed using the method developed by Larsén and Mann (2006), then converted to continuously sampled 10-minute and 3-second wind speeds using the Durst (1960) method; and Case III — both disjunctively sampled maximum sustained wind and continuously sampled peak gust wind speeds are reported. Case III combines the methodologies of Cases I and II, and results in the maximum derived sustained wind speed as

$$U_{t}^{\text{max}} = \frac{U_{d,t}^{\text{max}}}{U_{d}^{\text{max}}}$$

where $U_{t}^{\text{max}}$ = maximum wind speed averaged over time $t$, $U_{d,t}^{\text{max}}$ = disjunctively sampled maximum wind speed averaged over time $t$, and $U_{d}^{\text{max}}/U_{\text{max}} = b - a(\log_{10}(x))^{2}$ where $x$ = wind speed recurrence frequency, $a = 0.0209(T_{d} - 1/6)^{0.4627}$, $b = 1 - 0.0342(T_{d} - 1/6)^{0.5436}$, and $T_{d}$ = sampling interval in hours (Larsén & Mann 2006). For this study, a recurrence frequency of 10 years was selected. To apply the Durst curve, the appropriate $V_{t}/V_{3600}$ factors were selected from the ASCE 7 standard (ASCE 2010, p. 509).

### 2.2. Wind surface interpolation

During surface construction, ordinary kriging was chosen to interpolate wind station data based on its superiority over other techniques for representing wind speeds (e.g. Luo et al. 2008; Akkala et al. 2010; Zlatev et al. 2010; Luo et al. 2011). Ordinary kriging is represented as

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \gamma_i Z(s_i)$$
where $Z(s_i)$ is the measured value at the $i$th location, $\gamma_i$ is an unknown weight for the measured value at the $i$th location, $s_0$ is the prediction location, and $N$ is the number of measured values. Specifically, the spherical semivariogram model of ordinary kriging was chosen to examine both peak gust and maximum sustained wind speeds and is represented by

$$\gamma(h) = (s - n) \left( \frac{3h}{2r} - \frac{h^3}{2r^3} \right) 1_{(0,r)}(h) + 1_{(r,\infty)}(h) + n 1_{(0,\infty)}(h)$$

where $h$ is an approximate, or lag, distance; $s$ represents the sill parameter that defines the limit at a certain lag distance; $n$ represents the nugget, or the height above the origin where the semivariogram begins; and $r$ represents the range, defined as the distance at which the variogram and sill differences are insignificant (often when semivariance reaches 95% of sill). While multiple variogram models can be applied (e.g. spherical, circular, exponential, Gaussian, linear), the spherical version is the most commonly used across multiple data types, and is ideal when nugget variance plays a role and when autocorrelation has a clear limit (i.e. sill and range) (Robeson 1997; Ver Hoef et al. 2004; Luo et al. 2008). When considering the sill and range of modelled wind speeds across a surface, the spherical model is preferred over the exponential model because it accounts for a progressive decrease in spatial autocorrelation, as characteristic in windstorms, in contrast to an exponential decrease in spatial autocorrelation.

Wind speeds often exhibit directionality in which they are increasing or decreasing across a surface; however, microclimatological effects sometimes produce pockets or patches of high/low wind speed over a surface that can create confusion during kriging surface construction. The anisotropic function calculates a dominant surface trend, but this trend may not align with the actual correlation of wind speeds relative to a storm track, making it necessary to verify that the anisotropic azimuth direction reflects a reasonable wind speed correlation pattern. Anisotropic semivariograms were created during the interpolation procedure to account for directional dependence of wind speeds at varying distances, creating a spatial relationship for each direction that cannot be described by a single formula. Dominant directional trends were detected automatically for each storm and wind type. This most often resulted in directional trends that corresponded logically to storm tracks, but occasionally dominant trends were difficult to determine and directionality was adjusted within the semivariogram model accordingly. Additionally, a variable search radius was determined based on an optimized number of points through cross validation using the Geostatistical Analysis tool in ArcGIS Version 10.2. In this study a radius of 15 points was determined to reflect spatial covariance adequately, meaning that appropriate range and sill values could be determined by incorporating this number of points to estimate local surface trends, similar to a moving window. For semivariogram surface creation, an eight-sector elliptical search with three neighbours per sector was specified to optimize surface variability and smoothness.

The interpolation parameters were selected to obtain the best accuracy based on the station data. CVAEP, CVSSEP, CVSEP, minimum/maximum wind speed range, and the number of ‘high-error’ stations were used to evaluate the interpolated surface and compare the accuracy for each interpolated surface. These statistical measures evaluate accuracy related to observed and predicted variability of wind speed on the surface within leave-one-out cross validation (LOOCV). CVAEP values near zero indicate the surface interpolation as a whole did not discernibly under- or over-predict measured values; however, CVAEP only represents the error trend of the surface as a whole. Negative (positive) CVAEP values infer that measured values were underestimated (overestimated) overall across the surface. CVSSEP provides an assessment of the normalized sum of station prediction error, allowing evaluation of the variability of error across the surface, with values very close to one indicating low prediction error variability. CVSEP evaluates the variability of the prediction, with lower values representing a surface with lower prediction error. A wider (narrower) minimum/maximum range estimate infers more (less) variability in wind speeds across the interpolated surface.
'High-error' stations were identified as those that had a standardized prediction error (StddPE) of ±1.96. CVAEP, CVSSEP, CVSEP, and StddPE test statistics, based on LOOCV, were calculated as

\[
\text{CVAEP} = \frac{1}{N} \sum_{i=1}^{N} \left[ \hat{Z}(s_{-i}) - Z(s_i) \right]
\]

\[
\text{CVSSEP} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left[ \hat{Z}(s_{-i}) - Z(s_i) \right]^2}
\]

\[
\text{CVSEP} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left[ \hat{Z}(s_{-i}) - Z(s_i) \right]^2}
\]

\[
\text{StddPE} = \sum_{i=1}^{N} \left[ \frac{\hat{Z}(s_{-i}) - Z(s_i)}{\text{PSE}(s_{-i})} \right]
\]

where \(\hat{Z}(s_{-i})\) is the predicted wind speed for the \(i\)th observation when the \(i\)th observation is omitted from the model (i.e. LOOCV), \(Z(s_i)\) is the measured (or observed) wind speed, \(N\) is the number of weather stations used for the prediction, and \(\text{PSE}(s_{-i})\) is the prediction standard error of the \(i\)th observation when the \(i\)th observation is omitted from the model.

To verify the interpolation output with known storm tracks and magnitudes, additional reports of windstorms were obtained. Storm tracks were compared with general trends, locations of highest/lowest winds and maximum wind speed values seen in the interpolation (Zimmerli 2003). Maximum wind speed and storm damage data were also evaluated for storms where data were available (EQECAT 2000).

### 2.3. Geographic heterogeneity statistical testing

The three CSs were used to facilitate statistical testing of model performance within and between each GC. To evaluate the influence of geographic heterogeneity on interpolation results, accuracy metrics for each storm, wind type, and semivariogram were calculated for each storm by GC. The total number of stations within the GC was recorded, as was the mean LOOCV predicted wind speed within the GC, \(\frac{\sum_{i=1}^{n_j} \hat{Z}(s_{-i})}{n_j}\). To test the performance of the models in GC \(j\), the absolute value of the mean prediction error was used to avoid the effects of summing negative and positive values of prediction error. The ACVAEP for GC \(j\) is defined as

\[
\text{ACVAEP}_j = \frac{\sum_{i=1}^{n_j} |\hat{Z}(s_{-i}) - Z(s_i)|}{n_j} = \frac{\sum_{i=1}^{n_j} \text{PE}_{ij}}{n_j}
\]

where \(|\text{PE}_{ij}|\) is the absolute value of the prediction error for the \(i\)th observation within GC \(j\), and \(n_j\) is the number of stations in GC \(j\). CVSEP was calculated for each GC as

\[
\text{CVSEP}_j = \sqrt{\frac{\sum_{i=1}^{n_j} (\hat{Z}(s_{-i}) - Z(s_i))^2}{n_j}}
\]

and the number and percentage of high-error stations within each GC were recorded.
One-way analysis of variance (ANOVA) testing was performed to determine if significant differences in ACVAEP exist based on station GC. The ANOVA $F$-test is computed as the ratio between the mean square error between the GCs and the mean square error within the GC, or

$$F = \frac{MS_{\text{interGC}}}{MS_{\text{intraGC}}}$$

where $MS_{\text{interGC}}$ is the mean square error between the GCs and $MS_{\text{intraGC}}$ is the mean square error within the GC. The mean square error between GCs is calculated by dividing the sum of squares error between GCs by the corresponding degrees of freedom:

$$MS_{\text{interGC}} = \frac{SS_{\text{interGC}}}{df_{\text{interGC}}}$$

$SS_{\text{interGC}}$, the sum of squares error between GCs, examines the differences among the GC prediction error means by calculating the variation of each GC prediction error mean $ACVAEP_j = \frac{\sum_{i=1}^{nj} |PE_{ij}|}{nj}$ around the overall prediction error mean $ACVAEP = \frac{\sum_{j=1}^{a} \sum_{i=1}^{nj} |PE_{ij}|}{N}$, where $nj$ is the number of stations in GC $j$, $a$ is the number of GCs in each CS, and $N$ is the total number of stations in each CS. $N = \sum_{j=1}^{a} nj$

$$SS_{\text{interGC}} = \sum_{j=1}^{a} nj (ACVAEP_j - ACVAEP)^2$$

$$df_{\text{interGC}} = a - 1$$

The mean square error within the GC is calculated by dividing the sum of squares error within the GC by the corresponding degrees of freedom or

$$MS_{\text{intraGC}} = \frac{SS_{\text{intraGC}}}{df_{\text{intraGC}}}$$

$SS_{\text{intraGC}}$, the sum of squares error within each GC, examines the variation of an individual station’s prediction error around the GC prediction error mean, where $|PE_{ij}|$ represents the absolute value of prediction error at station $i$ within GC $j$ and $N$ is the total number of stations in the CS:

$$SS_{\text{intraGC}} = \sum_{j=1}^{a} \sum_{i=1}^{nj} (|PE_{ij}| - ACVAEP_j)^2$$

$$df_{\text{intraGC}} = N - a$$

If the $F$-test results are significant, pair-wise comparison of means based on Fisher’s least significant difference (LSD) was applied to determine which GCs had statistically significant differences in mean prediction error. The value of the $t$ statistic to evaluate the difference in ACVAEP between two GCs, 1 and 2, is given by

$$t = \frac{ACVAEP_1 - ACVAEP_2}{\sqrt{MS_{\text{intraGC}} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$
where ACVAEP\textsubscript{1} is the absolute cross validation average error of prediction of GC 1, ACVAEP\textsubscript{2} is the absolute cross validation average error of prediction of GC 2, MS\textsubscript{intraGC} is the mean square error within the GC, \(n_1\) is the number of stations in GC 1, and \(n_2\) is the number of stations in GC 2.

This \(t\) value follows a student’s \(t\) distribution with \(N - a\) degrees of freedom. The ratio \(t\) would be declared significant at a given \(\alpha\) significance level (\(\alpha = 0.05\) for this study) if the value of \(t\) is larger than the critical value of \(t_{\text{df intraintraGC}}\) where \(\text{df intraintraGC} = N - a\). Rewriting this ratio, the difference between the ACVAEP of GCs 1 and 2 is significant if

\[
|\text{ACVAEP}_1 - \text{ACVAEP}_2| > \text{LSD} = \sqrt{\text{MS}_{\text{intraGC}} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}
\]

3. Results

3.1. Wind station data

To test the performance of the wind speed adjustment methodology, the Lothar dataset was evaluated as an initial test case using the methodology detailed in Figure 2. A comparison of peak gust wind speeds calculated under Case II (i.e. assuming all peak gust wind speeds were missing) with known peak gust wind speeds showed that the methodology did not significantly over-predict gust wind speed data (an average increase of 2%, standard deviation of 5%). Further, the methodology accounted for the under-reporting of sustained wind speeds as a result of the disjunct sampling period. These initial results indicate that the proposed calculation methodology provides an adequate representation of continuously sampled maximum sustained and peak gust wind speeds that accounts for missing values and disjunct sampling for use in the surface interpolations.

3.2. Interpolation results

Following the initial evaluation of Lothar, the wind speed data for both storms were processed according to Figure 2. Anisotropic and isotropic spherical kriging were performed using the derived equivalent continuously sampled maximum sustained and peak gust wind speeds at each station for both windstorms. This resulted in four interpolations for each storm, for a total of eight interpolations. Accuracy metrics for each storm and wind type are provided in Table 3.

The four interpolated surfaces created for each storm represent the greatest maximum sustained wind speeds and the peak gust wind speeds for the duration of each storm. The maximum sustained wind speed interpolations illustrate 10-minute continuously sampled sustained wind in open terrain, and the peak gust wind speed interpolations illustrate continuously sampled 3-second gust in open terrain. Differences between isotropic and anisotropic surfaces were mapped for each storm to evaluate variability between the methods.

| Storm name | Wind type | Range | Mean wind speed | CVAEP | CVSSEP | CVSEP | High-error stations (#) |
|------------|-----------|-------|-----------------|-------|--------|-------|------------------------|
| Lothar     | MS        | 25.86 | 17.11          | 0.08  | 1.01   | 6.48  | 15                     |
|            | MS w/anisotropy | 21.97 | 17.02          | -0.01 | 1.03   | 6.33  | 17                     |
|            | PG        | 40.69 | 25.64          | 0.13  | 1.01   | 10.49 | 19                     |
|            | PG w/anisotropy | 28.33 | 25.55          | 0.04  | 1.00   | 10.37 | 17                     |
| Martin     | MS        | 19.80 | 16.81          | 0.12  | 0.94   | 6.66  | 15                     |
|            | MS w/anisotropy | 19.70 | 16.75          | 0.06  | 0.96   | 6.66  | 16                     |
|            | PG        | 31.00 | 24.91          | 0.12  | 0.95   | 10.84 | 17                     |
|            | PG w/anisotropy | 30.99 | 24.94          | 0.14  | 0.97   | 10.83 | 18                     |
3.2.1. Lothar windstorm

A minimum/maximum range estimate of 21.97 m s$^{-1}$ and a mean wind speed of 17.02 m s$^{-1}$ were produced by the anisotropic maximum sustained wind speed interpolation, while a range estimate of 25.86 m s$^{-1}$ and a mean wind speed of 17.11 m s$^{-1}$ were produced by the isotropic maximum sustained wind speed interpolation (Table 3). Anisotropic interpolation resulted in a CVAEP closer to zero (−0.01) and a lower CVSEP (6.33), indicating improvement over the isotropic surface; however, the higher CVSSEP (1.03) indicates slight underestimation of variability and a small uptick in high-error stations (17 compared with 15) was observed. The anisotropic maximum sustained wind speed interpolation is illustrated in Figure 3(a) and the difference between the anisotropic and isotropic surfaces is shown in Figure 3(b). The anisotropic surface predicted higher wind speeds (red) in isolated areas of western and eastern Switzerland, south-western Germany, and south-eastern coastal France, while lower wind speeds (blue) were predicted in isolated areas of south-central Switzerland, south-eastern Switzerland, south-eastern France, and central Germany. A minimum/maximum range estimate of 28.33 m s$^{-1}$ and a mean wind speed of 25.55 m s$^{-1}$ were produced by the anisotropic peak gust wind speed interpolation, while a range estimate of 40.69 m s$^{-1}$ and a mean wind speed of 25.64 m s$^{-1}$ were generated by the isotropic peak gust wind speed interpolation.

![Figure 3](image-url)

Figure 3. Anisotropic kriging maximum sustained wind speed interpolation for Lothar (a) and the difference between anisotropic and isotropic interpolations (b). Anisotropic kriging peak gust wind speed interpolation for Lothar (c) and the difference between anisotropic and isotropic interpolations (d). (To view this figure in colour, see the online version of the journal.)
Anisotropic interpolation resulted in a CVAEP closer to zero (0.04), a CVSSEP closer to one (1.00), a lower CVSEP (10.37), and fewer high-error stations (17 compared with 19), all indicating improvement over the isotropic surface. The anisotropic peak gust wind speed interpolation is illustrated in Figure 3(c) and the difference between the anisotropic and isotropic surfaces is depicted in Figure 3(d). The anisotropic surface predicted higher wind speeds (red) in many areas of south-western and south-eastern France, north-western coastal France, eastern Switzerland, south-western Germany, and eastern Germany, while lower wind speeds (blue) were predicted in northern coastal France, multiple areas of southern France, central and south-eastern Switzerland, and central and south-eastern Germany. The Lothar windstorm followed a west—east track, making landfall in northern France and moving into central Germany with the strongest winds (33—42 m s\(^{-1}\) peak gust) occurring south of this track in France and southern Germany. Some reports of wind speeds 8—9 m s\(^{-1}\) higher than those shown in the interpolation occurred in parts of France (EQE 2000). These extremes were most likely localized and thus smoothed by the interpolation. Similar to other reports, stronger wind speeds also occurred near the foothills of mountainous regions in south-eastern Germany bordering Switzerland and Austria.

### 3.2.2. Martin windstorm

A minimum/maximum range estimate of 19.70 m s\(^{-1}\) and a mean wind speed of 16.75 m s\(^{-1}\) were produced by the anisotropic maximum sustained wind speed interpolation, while a similar range estimate of 19.80 m s\(^{-1}\) and a mean wind speed of 16.81 m s\(^{-1}\) were generated by the isotropic maximum sustained wind speed interpolation (Table 3). Anisotropic interpolation resulted in a CVAEP closer to zero (0.06), a CVSSEP closer to one (0.96), and the same CVSEP (6.66 compared to 6.66) for the anisotropic interpolation, indicating improvement over the isotropic surface; however, a marginal increase in high-error stations (16 compared with 15) was calculated. The anisotropic maximum sustained wind speed interpolation is illustrated in Figure 4(a) and the difference between the anisotropic and isotropic surfaces is shown in Figure 4(b). The anisotropic surface predicted higher wind speeds (red) in isolated areas of south-western and north-eastern France, coastal south-eastern France, eastern Switzerland, and southern Germany, while lower wind speeds (blue) were predicted in south-eastern France, southern and south-eastern Switzerland, and south-eastern Germany. A minimum/maximum range estimate of 30.99 m s\(^{-1}\) and a mean wind speed of 24.94 m s\(^{-1}\) were produced by the anisotropic peak gust wind speed interpolation, while a range estimate of 31.00 m s\(^{-1}\) and a mean wind speed of 24.91 m s\(^{-1}\) were generated by the isotropic maximum sustained wind speed interpolation (Table 3). A CVSSEP closer to one (0.97) and a lower CVSEP (10.83) indicate improvement over the isotropic surface; however, a higher CVAEP (0.14) and a marginal increase in high-error stations (18 compared with 17) were calculated. The anisotropic peak gust wind speed interpolation is illustrated in Figure 4(c) and the difference between the anisotropic and isotropic surfaces is shown in Figure 4(d). The anisotropic surface predicted higher wind speeds (red) in multiple areas of southern Germany and northern coastal Germany, parts of eastern Switzerland, and scattered areas across France especially in south-western France and coastal south-eastern France, while lower wind speeds (blue) were predicted in south-eastern France, south-central Switzerland, and south-eastern Germany. The Martin windstorm followed closely behind Lothar by only one day, making landfall in western coastal France. The track of this storm was slightly more southern than that of Lothar and the interpolation results demonstrate this trend. Consistent with other reports, the location of the highest wind speeds (33—42 m s\(^{-1}\)) was in western France and the mountainous areas of southern Germany.

### 3.3. Geographic classification results

The analysis produced fairly consistent results between the two storm events and semivariogram types (Table 4), allowing discussion of average findings by wind speed type and GC. Averaging wind speeds for both storms and semivariogram types, GC-A (coastal/mountainous) had the highest
wind speeds (MS: 19.86 m s\(^{-1}\), PG: 29.63 m s\(^{-1}\)), followed by GC-F (all mountainous; MS: 19.10 m s\(^{-1}\), PG: 23.47 m s\(^{-1}\)) and GC-E (all coastal; MS: 18.50 m s\(^{-1}\), PG: 27.37 m s\(^{-1}\)). The lowest wind speeds were found in GC-C (non-coastal/mountainous; MS: 16.62 m s\(^{-1}\), PG: 25.21 m s\(^{-1}\)) and GC-D (non-coastal/non-mountainous; MS: 16.61 m s\(^{-1}\), PG: 24.56 m s\(^{-1}\)). Averaging CVAEP for both storms and semivariogram types, GC-A (coastal/mountainous) had the largest negative average CVAEP (MS: \(-5.98\), PG: \(-10.63\)), followed by GC-E (all coastal; MS: \(-3.21\), PG: \(-6.54\)), GC-B (coastal/non-mountainous; MS: \(-2.08\), PG: \(-4.88\)), and GC-F (all mountainous; MS: \(-0.27\), PG: \(-0.79\)). These negative values of CVAEP indicate that the interpolation models under-predict wind speeds in GC-A, B, E, and F for all wind speed types and semivariograms for both storms. Positive average CVAEP was found for GC-D (non-coastal/non-mountainous; MS: 0.44, PG: 1.34) and lowest for GC-C (non-coastal/mountainous; MS: 0.39, PG: 0.34). Averaging ACVAEP and CVSEP for both storms and semivariogram types, GC-A (coastal/mountainous) had the largest average ACVAEP (MS: 9.03, PG: 16.29) and CVSEP (MS: 11.53, PG: 20.19). GC-E (all coastal) had the next highest ACVAEP (MS: 5.75, PG: 10.52) and an MS CVSEP slightly lower than GC-F (MS: 7.59, PG: 13.67), followed by GC-F (all mountainous; ACVAEP - MS: 5.67, PG: 9.46; CVSEP - MS: 7.75, PG: 12.81), GC-C (non-coastal/mountainous; ACVAEP - MS: 5.29, PG: 8.69;

![Figure 4](https://example.com/figure4.png)
CVSEP - MS: 7.19, PG: 11.68), and GC-B (coastal/non-mountainous; ACVAEP - MS: 4.42, PG: 7.89; CVSEP - MS: 5.20, PG: 9.85). GC-D (non-coastal/non-mountainous) had the lowest average ACVAEP (MS: 3.77, PG: 5.73) and CVSEP (MS: 5.03, PG: 7.73). GC-A had the highest percentage of high-error stations (MS: 21.2%, PG: 28.8%), but no classification consistently showed the lowest percentage of high-error stations. GC-A also had the lowest number of total stations, which may have affected this statistic.

ANOVA test results (significance level $\alpha = 0.05$) show that there are significant differences among the station absolute mean prediction error (i.e. ACVAEP) for each storm for both maximum sustained and peak gust interpolations, with and without anisotropy for each CS (Table 5). Therefore, post hoc ANOVA testing and pair-wise comparison of station ACVAEP based on Fisher’s LSD

| Storm name | Wind type | GC | Number of stations | Range | Mean wind speed | CVAEP | ACVAEP | CVSEP | High-error stations |
|------------|-----------|----|--------------------|-------|-----------------|-------|--------|-------|---------------------|
| Lothar     | MS        | A  | 13                 | 7.87  | 18.32           | -5.19 | 8.86   | 10.87 | 2                   |
|            |           | B  | 32                 | 14.50 | 18.75           | -2.00 | 4.35   | 5.03  | 0                   |
|            |           | C  | 116                | 19.76 | 16.40           | 0.30  | 5.15   | 7.02  | 11                  |
|            |           | D  | 161                | 22.62 | 17.05           | 0.41  | 3.71   | 4.91  | 2                   |
|            |           | E  | 45                 | 14.50 | 18.63           | -2.92 | 5.65   | 7.22  | 2                   |
|            |           | F  | 129                | 19.76 | 16.58           | -0.28 | 5.54   | 7.51  | 13                  |
|            | MS w/anisotropy | A  | 13                 | 8.54  | 18.92           | -4.58 | 8.88   | 10.78 | 2                   |
|            |           | B  | 32                 | 14.02 | 18.75           | -1.99 | 4.41   | 5.14  | 0                   |
|            |           | C  | 116                | 21.97 | 16.28           | 0.18  | 5.09   | 7.12  | 13                  |
|            |           | D  | 161                | 17.62 | 17.01           | 0.36  | 3.71   | 4.91  | 2                   |
|            |           | E  | 45                 | 14.02 | 18.81           | -2.74 | 5.70   | 7.24  | 2                   |
|            |           | F  | 129                | 21.97 | 16.52           | -0.34 | 5.48   | 7.58  | 15                  |
| PG         |           | A  | 13                 | 13.33 | 27.12           | -8.79 | 16.57  | 19.53 | 4                   |
|            |           | B  | 32                 | 21.12 | 27.88           | -4.82 | 7.28   | 8.84  | 1                   |
|            |           | C  | 116                | 30.48 | 24.98           | 0.26  | 8.37   | 11.38 | 11                  |
|            |           | D  | 161                | 36.35 | 25.27           | 1.18  | 5.58   | 7.45  | 3                   |
|            |           | E  | 45                 | 21.12 | 27.66           | -5.96 | 9.97   | 12.87 | 5                   |
|            |           | F  | 129                | 28.33 | 24.93           | -0.95 | 9.50   | 12.80 | 14                  |
|            | PG w/anisotropy | A  | 13                 | 12.59 | 27.34           | -8.57 | 15.93  | 19.19 | 3                   |
|            |           | B  | 32                 | 18.23 | 27.56           | -5.14 | 7.14   | 9.31  | 2                   |
|            |           | C  | 116                | 28.33 | 24.72           | -0.01 | 8.78   | 11.84 | 11                  |
|            |           | D  | 161                | 22.70 | 25.52           | 1.41  | 5.63   | 7.45  | 1                   |
|            |           | E  | 45                 | 18.23 | 27.50           | -6.13 | 9.68   | 12.96 | 5                   |
|            |           | F  | 129                | 28.33 | 24.93           | -0.95 | 9.50   | 12.80 | 14                  |
| Martin     | MS        | A  | 13                 | 8.50  | 20.72           | -7.46 | 9.52   | 12.45 | 4                   |
|            |           | B  | 32                 | 16.83 | 17.24           | -2.07 | 4.52   | 5.21  | 0                   |
|            |           | C  | 117                | 13.30 | 17.06           | 0.68  | 5.44   | 7.25  | 7                   |
|            |           | D  | 163                | 18.74 | 16.15           | 0.45  | 3.80   | 5.14  | 4                   |
|            |           | E  | 45                 | 17.99 | 18.24           | -3.63 | 5.96   | 8.00  | 4                   |
|            |           | F  | 130                | 13.98 | 17.42           | -0.13 | 5.85   | 7.93  | 11                  |
|            | MS w/anisotropy | A  | 13                 | 8.98  | 21.50           | -6.68 | 8.86   | 12.02 | 3                   |
|            |           | B  | 32                 | 15.91 | 17.05           | -2.26 | 4.40   | 5.40  | 0                   |
|            |           | C  | 117                | 14.55 | 16.76           | 0.39  | 5.48   | 7.37  | 9                   |
|            |           | D  | 163                | 19.70 | 16.24           | 0.54  | 3.86   | 5.15  | 4                   |
|            |           | E  | 45                 | 17.32 | 18.33           | -3.54 | 5.69   | 7.91  | 3                   |
|            |           | F  | 130                | 14.55 | 17.24           | -0.32 | 5.81   | 7.96  | 12                  |
| PG         |           | A  | 13                 | 14.69 | 31.56           | -13.06| 16.73  | 21.27 | 4                   |
|            |           | B  | 32                 | 24.60 | 25.20           | -4.75 | 8.55   | 10.48 | 1                   |
|            |           | C  | 117                | 21.86 | 25.63           | 0.63  | 8.78   | 11.67 | 9                   |
|            |           | D  | 163                | 29.43 | 23.67           | 1.32  | 5.83   | 8.03  | 3                   |
|            |           | E  | 45                 | 28.61 | 27.04           | -7.15 | 10.91  | 14.45 | 5                   |
|            |           | F  | 130                | 22.50 | 26.23           | -0.74 | 9.58   | 12.96 | 13                  |
|            | PG w/anisotropy | A  | 13                 | 15.61 | 32.49           | -12.13| 15.92  | 20.75 | 4                   |
|            |           | B  | 32                 | 22.17 | 25.14           | -4.80 | 8.60   | 10.77 | 2                   |
|            |           | C  | 117                | 23.41 | 25.49           | 0.49  | 8.84   | 11.82 | 8                   |
|            |           | D  | 163                | 30.99 | 23.79           | 1.44  | 5.87   | 7.97  | 4                   |
|            |           | E  | 45                 | 28.54 | 27.27           | -6.92 | 10.71  | 14.39 | 6                   |
|            |           | F  | 130                | 23.83 | 26.19           | -0.78 | 9.55   | 12.99 | 12                  |
were performed to determine which GCs had statistically significant differences in ACVAEP. Table 6 shows the post hoc ANOVA pair-wise test results for both storms and for all CSs, GC comparison pairs, and wind speed and semivariogram types. Asterisks appearing after p-values denote statistically significant differences in ACVAEP between GC-1 and GC-2. Positive t-values in Table 6 indicate that the ACVAEP of GC-1 is greater than the ACVAEP of GC-2.

For CS 1, post hoc ANOVA found that stations in GC-A (coastal/mountainous) had significantly higher absolute mean prediction error (ACVAEP) than stations in GC-B (coastal/non-mountainous) for Lothar for all wind speed types and semivariogram types. For Lothar, stations in GC-A (coastal/mountainous) also had significantly higher ACVAEP than stations in GC-C (non-coastal/mountainous) for peak gust wind speeds, with and without anisotropy. For both Lothar and Martin, stations in GC-A (coastal/mountainous) had significantly higher ACVAEP than stations in GC-D

| Storm/CS | GC-1 | GC-2 | t-value | p-value | t-value | p-value | t-value | p-value | t-value | p-value |
|----------|------|------|---------|---------|---------|---------|---------|---------|---------|---------|
| Lothar   |      |      |         |         |         |         |         |         |         |         |
| CS 1     | A B  | 2.41 | 0.0308* | 2.45 | 0.0285* | 2.98 | 0.0097* | 2.69 | 0.0168* |
|          | A C  | 1.98 | 0.0680 | 2.08 | 0.0573 | 2.67 | 0.0187* | 2.25 | 0.0417* |
|          | A D  | 2.81 | 0.0152* | 2.90 | 0.0128* | 3.66 | 0.0031* | 3.31 | 0.0060* |
|          | B C  | −1.26 | 0.2092 | −1.03 | 0.3061 | −0.95 | 0.3462 | −1.26 | 0.2130 |
|          | B D  | 1.24 | 0.2222 | 1.31 | 0.1971 | 1.74 | 0.0893 | 1.33 | 0.1924 |
|          | C D  | 2.82 | 0.0053* | 2.62 | 0.0096* | 3.42 | 0.0008* | 3.77 | 0.0002* |
| CS 2†    | C E  | −0.62 | 0.5377 | −0.75 | 0.4573 | −1.12 | 0.2666 | −0.60 | 0.5499 |
|          | D E  | −2.69 | 0.0093* | −2.78 | 0.0074* | −3.40 | 0.0013* | −2.99 | 0.0043* |
| CS 3†    | B F  | −1.85 | 0.0678 | −1.61 | 0.111 | −1.65 | 0.1038 | −1.80 | 0.0766 |
|          | D F  | −3.53 | 0.0005* | −3.35 | 0.001* | −4.32 | <0.0001* | −4.54 | <0.0001* |

| Storm/CS | GC-1 | GC-2 | t-value | p-value | t-value | p-value | t-value | p-value | t-value | p-value |
|----------|------|------|---------|---------|---------|---------|---------|---------|---------|---------|
| Martin   |      |      |         |         |         |         |         |         |         |         |
| CS 1     | A B  | 2.12 | 0.0543 | 1.85 | 0.0859 | 2.07 | 0.0569 | 1.82 | 0.0895 |
|          | A C  | 1.73 | 0.1081 | 1.42 | 0.1794 | 2.06 | 0.0601 | 1.81 | 0.0941 |
|          | A D  | 2.45 | 0.0299* | 2.12 | 0.0548 | 2.86 | 0.0142 | 2.60 | 0.0229* |
|          | B C  | −1.44 | 0.1535 | −1.48 | 0.1418 | −0.18 | 0.8594 | −0.18 | 0.8578 |
|          | B D  | 1.34 | 0.1858 | 0.86 | 0.3941 | 2.31 | 0.0257* | 2.20 | 0.0339* |
|          | C D  | 3.16 | 0.0018* | 3.05 | 0.0026* | 3.53 | 0.0005* | 3.53 | 0.0005* |
| CS 2†    | C E  | −0.56 | 0.5752 | −0.22 | 0.8232 | −1.34 | 0.1861 | −1.15 | 0.2534 |
|          | D E  | −2.55 | 0.0138* | −2.09 | 0.0413* | −3.4 | 0.0013* | −3.21 | 0.0023* |
| CS 3†    | B F  | −2.02 | 0.0465* | −1.92 | 0.0588 | −0.77 | 0.4437 | −0.68 | 0.4981 |
|          | D F  | −3.78 | 0.0002* | −3.57 | 0.0005* | −4.24 | <0.0001* | −4.17 | <0.0001* |

*Statistically significant at 0.05 level.
†GC C and D are compared in CS 1.
‡GC B and D are compared in CS 1.

Table 5. ANOVA test results for differences in station absolute mean prediction error (ACVAEP) based on geographic classification.

| Storm name | Wind type | F-test | p-value | F-test | p-value | F-test | p-value |
|------------|-----------|--------|---------|--------|---------|--------|---------|
| Lothar     | MS        | 5.01   | 0.0063*| 6.48   | 0.0023*| 6.32   | 0.0025*|
|            | MS w/anisotropy | 4.88   | 0.0072*| 6.23   | 0.0029*| 5.77   | 0.0041*|
|            | PG        | 8.17   | 0.0005*| 10.16  | 0.0001*| 9.80   | 0.0001*|
|            | PG w/anisotropy | 8.09   | 0.0006*| 10.28  | 0.0001*| 10.63  | <0.0001*|
| Martin     | MS        | 5.09   | 0.0059*| 7.09   | 0.0014*| 7.21   | 0.0011*|
|            | MS w/anisotropy | 4.37   | 0.0121*| 6.00   | 0.0037*| 6.36   | 0.0023*|
|            | PG        | 7.32   | 0.0010*| 10.43  | <0.0001*| 10.21  | <0.0001*|
|            | PG w/anisotropy | 6.88   | 0.0015*| 9.98   | 0.0001*| 9.86   | 0.0001*|

*Statistically significant at 0.05 level.
(non-coastal/non-mountainous), except for the Martin MS with anisotropy model. For Martin, stations in GC-B (coastal/non-mountainous) had significantly higher ACVAEP than stations in GC-D (non-coastal/non-mountainous) for peak gust wind speeds, with and without anisotropy. For both Lothar and Martin, stations in GC-C (non-coastal/mountainous) had significantly higher ACVAEP than stations in GC-D (non-coastal/non-mountainous) for all wind speed types and semivariogram types.

For CS-2, stations in GC-D (non-coastal/non-mountainous) had significantly lower ACVAEP than stations in GC-E (all coastal) for both storms and wind speed types, with and without anisotropy. For CS-3, stations in GC-D (non-coastal/non-mountainous) had significantly lower ACVAEP than stations located in GC-F (all mountainous) for both storms and wind speed types, with and without anisotropy. Additionally, for Martin, stations in GC-B (coastal/non-mountainous) had significantly lower ACVAEP than stations located in GC-F (all mountainous) for MS without anisotropy.

3.3.1. Interpolation accuracy and meteorological considerations

Maximum sustained and peak gust wind speed interpolations consistently underestimated the highest of the observed wind speeds and overestimated the lowest of the observed wind speeds. These values were usually in the highest or lowest 3% of observed wind speeds and thus would have been considered outliers by the interpolation when creating the distribution of wind speeds across the surface. Extreme local high/low wind speeds are accounted for by kriging because these observations impact the trend and directional covariance of wind speeds, but more precise prediction would initiate the possibility for spurious ‘eyes’ such as those seen in IDW-interpolated surfaces. This, in turn, would have impacted the accuracy of wind speeds surrounding the anomalous stations and decreased overall accuracy (Luo et al. 2008).

The variability of the estimated wind surface as it relates to the observed surface is the first statistical method of determining the interpolation’s accuracy. Intra-model accuracy metrics (CVAEP and CVSSEP) indicate a reasonable goodness-of-fit for both storms and semivariogram types. The reduction in CVAEP in three of four model comparisons indicates overall improvement from the application of anisotropy. Lower CVSEP values (used for inter-model comparison) for all anisotropic interpolations also indicate that the application of anisotropy to semivariograms at this interpolation scale provides improved models. When comparing wind speed types for both storms, maximum sustained wind speeds produced lower CVSEP values. This alludes to the increased variability and complexity of higher wind speeds (Etienne & Beniston 2012; Roberts et al. 2014). Increased wind variability between adjacent stations may have led to over- or underestimation of variability in some areas. The overall number of high-error stations (stations where estimated wind speeds differed from observations by more than ±1.96 standardized prediction errors; Figures 5(a) and (b)) metric does not appear to reflect the improvement in interpolated surfaces with anisotropy that the CVAEP, CVSSEP, and CVSEP demonstrate. Only one anisotropic model (Lothar peak gust) showed a reduction in high-error stations, while the other three anisotropic models showed slight increases in number of stations (Lothar maximum sustained, Martin peak gust, and Martin maximum sustained).

However, based on the statistical testing of ACVAEP performed on all station data (not only of stations that did not exceed a high-error threshold) that were separated into CS and GC categories, a clearer finding emerges: stations not located in coastal and/or mountainous regions had statistically significant lower ACVAEP. Collectively and separately, stations located in coastal and mountainous areas were often found to have higher ACVAEP, indicating that diverse land surfaces play a major role in impacting surface-level wind speeds to such an extent that interpolation of wind speed station data results in significant variability in surface estimates between coastal/mountainous areas and non-coastal/non-mountainous areas.

Meteorologically, wind should be viewed as a fluid mechanism — acting and reacting based on the varied geographic areas it traverses. In many mountainous locations, the result of this interaction is relatively stronger and more spatially varied wind that is more difficult to estimate through interpolation processes, thus often resulting in higher prediction errors. One possible explanation for
stronger winds in mountainous areas is not only the presence of an exposed ridgeline, but also the manifestation of pressure gradient differences associated with convergent air masses. Steenburgh and Mass (1996) alluded to the tendency for strong winds to flow through mountain gaps/valleys perpendicular to windstorm tracks in North America and it appears that this feature is also present in mountainous regions of Europe. Strong downgradient flow in gaps and channels occurs because of large north–south pressure gradients, so if a storm moves from west to east, then north–south oriented gaps often experience an increase in wind speed (Steenburgh & Mass 1996). Stronger wind speeds in coastal locations are most often the result of winds moving across a frictionless surface (i.e. the ocean) before moving over various forms of friction on land that gradually result in wind speed reductions. When considering higher error values at coastal stations, while a complex surface-level boundary transition interaction does exist, a unique interpolation phenomenon also exists — the exclusion of station data over a large area in one direction from each coastal station. This lack of station data leads to semivariogram and neighbourhood smoothing procedures that may be skewed by the incorporation of more inland station data.

Stations with high positive errors were often adjacent to stations with high negative errors, indicating that wind speed can differ drastically across short distances, depending on many geographic and atmospheric factors. It is not expected that kriging, or other interpolation methods, would fully capture these types of variability because the method predominantly predicts widespread (macro-meteorological) wind speed patterns and not highly local (micro-meteorological) anomalies. To better predict micro-meteorological anomalies associated with windstorms, local-scale (i.e. country or smaller) kriging should be examined in future studies.

4. Conclusions

This study extended previous findings (e.g. Luo et al. 2008; Zlatev et al. 2009) about the appropriateness of kriging for the interpolation of wind data by analyzing isotropic and anisotropic semivariogram-derived kriging surfaces and evaluating a large surface across geographically diverse terrain. Interpolated surfaces for two major European windstorms were created from historical wind records, which were adjusted to account for disjunctive sampling practices and averaging time. Peak gust and maximum sustained wind speed calculations and adjustments improved station-level data by addressing disjunct sampling and generating missing peak gust measurements to provide a more
robust and complete wind speed dataset, thus improving final interpolations. For the study areas evaluated, the application of anisotropy for semivariogram creation across multiple countries resulted in improved interpolation results in most, but not all, cases by capturing the widespread directional distribution of wind speeds in addition to the surface wind speed trends for windstorms Lothar and Martin, confirming the untested conclusions of Luo et al. (2008) that anisotropy could improve surface interpolations in a larger study area. Lower CVSEP scores in all anisotropic models and lower variability in CVAEP and CVSSEP scores in three of four anisotropic models provide evidence of improvement. While the number of high-error stations overall did not appear to be an effective metric in the evaluation of interpolation surfaces, there were generally more of these stations located in coastal or mountainous locations versus non-coastal/non-mountainous locations.

Three CSs were tested for both wind speed and semivariogram types to evaluate statistically significant differences in ACVAEP (at the $\alpha = 0.05$ level) based on station GC and consistently found higher ACVAEP in GC-A (coastal/mountainous) than in GC-D (non-coastal/non-mountainous), except Martin PG with anisotropy; higher ACVAEP in GC-C (non-coastal/mountainous) than in GC-D (non-coastal/non-mountainous); lower ACVAEP in GC-D (non-coastal/non-mountainous) than in GC-E (all coastal); and lower ACVAEP in GC-D (non-coastal/non-mountainous) than in GC-F (all mountainous). Additionally, for Lothar, all model types found statistically significant higher ACVAEP in GC-A (coastal/mountainous) than in GC-B (coastal/non-mountainous), indicating the influence of mountains in increasing coastal station ACVAEP. For Lothar PG models, statistically significant higher ACVAEP was found in GC-A (coastal/mountainous) than in GC-C (non-coastal/mountainous), indicating the influence of the coast in increasing mountain station ACVAEP for peak gust wind speeds in that event. Similarly, for Martin PG models, statistically significant higher ACVAEP was found in GC-B (coastal/non-mountainous) than in GC-D (non-coastal/non-mountainous), also indicating the influence of the coast in increasing non-mountain station ACVAEP for peak gust wind speeds in that event.

These findings indicate that, compared to isotropic kriging, anisotropic kriging can improve wind speed interpolations over geographically large areas and that wind speed predictions vary significantly between weather stations in coastal and/or mountains regions when compared to weather stations in non-coastal/non-mountainous regions, with some of the highest error measurements occurring in coastal and/or mountainous regions. Nevertheless, further research is needed to improve understanding of local wind variability in coastal and mountainous environments during high wind events and to address causative factors behind high-error stations, especially for study areas with significant geographic diversity. Development of more local-scale, geography-dependent interpolation models that are subsequently integrated with meso- and macro-scale models may address some of these issues.

**Acknowledgments**

We would like to thank the Austrian Marshall Plan Foundation and ImageCat, Inc. for data resources and collaborative support. We would also like to thank Barry Keim and Michael Martin for their helpful comments and proofreading of the original document.

**Disclosure statement**

The authors do not have any financial interest or benefit arising from the direct applications of this research.

**ORCID**

Carol J. Friedland http://orcid.org/0000-0003-0443-5266
T. Andrew Joyner http://orcid.org/0000-0003-2790-6817
Carol Massarra http://orcid.org/0000-0003-2036-7216
References

Akkala A, Devabhaktuni V, Kumar A. 2010. Interpolation techniques and associated software for environmental data. Environ Prog Sustain Energy. 29:134–141.

[ASCE] American Society of Civil Engineers. 2010. Minimum design loads for buildings and other structures. ASCE Standard 7-10. Reston, VA: American Society of Civil Engineers.

Bentamy A, Quilfen Y, Gohin F, Grima N, Lenour M, Servain J. 1996. Determination and validation of average wind fields from ERS-1 scatterometer measurements. Newark, NJ: International Publishers Distributor.

Carta JA, Vélazquez S, Matías JM. 2011. Use of Bayesian networks classifiers for long-term mean wind turbine energy output estimation at a potential wind energy conversion site. Energy Convers Manag. 52:1137–1149.

Cellura M, Cirrincione G, Marvuglia A, Miraoul A. 2008. Wind speed spatial estimation for energy planning in Sicily: a neural kriging application. Renew Energy. 33:1251–1266.

Della-Marta PM, Mathis H, Frei C, Liniger MA, Kleinn J, Appenzeller C. 2009. The return period of wind storms over Europe. Int J Climatol. 29:437–459.

Durst C. 1960. Wind speeds over short periods of time. Meteorol Mag. 89:181–187.

EQECAT. 2000. EQE summary report: the European storms Lothar and Martin, December 26–28, 1999. Oakland (CA): ABS Consulting.

Etienne C, Beniston M. 2012. Wind storm loss estimations in the Canton of Vaud (Western Switzerland). Nat Hazards Earth Syst Sci. 12:3789–3798.

Gatey D. 2011. The analysis of extreme synoptic winds [Monography]. London: University of Western Ontario.

Joyner TA, Friedland CJ, Rohlí RV, Treviño AM, Massara C, Paulus G. 2015. Cross-correlation modeling of European windstorms: a cokriging approach for optimizing surface wind estimates. Spat Stat. 13:62–75.

Krayer WR, Marshall RD. 1992. Gust factors applied to hurricane winds. Bull Am Meteorol Soc. 73:613–617.

Lanza L, Ramirez J, Todini E. 2001. Stochastic rainfall interpolation and downscaling. Hydrol Earth Syst Sci. 5:139–143.

Larsén XG, Mann J. 2009. Extreme winds from the NCEP/NCAR reanalysis data. Wind Energy. 12:556–573.

Larsén XLG, Mann J. 2006. The effects of disjunct sampling and averaging time on maximum mean wind speeds. J Wind Eng Ind Aerodyn. 94:581–602.

Li G, Shi J. 2010. On comparing three artificial neural networks for wind speed forecasting. Appl Energy. 87:2313–2320.

Luo W, Taylor MC, Parker SR. 2008. A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. Int J Climatol. 28:947–959.

Luo X, Xu Y, Shi Y. 2011. Comparison of interpolation methods for spatial precipitation under diverse orographic effects. Proceedings of the 19th International Conference on Geoinformatics; Jun 24–26. Shanghai, China.

Oztopal A. 2006. Artificial neural network approach to spatial estimation of wind velocity data. Energy Convers Manag. 47:395–406.

Phillips DL, Lee EH, Herstrom AA, Hogsett WE, Tingey DT. 1997. Use of auxiliary data for spatial interpolation of ozone exposure in southeastern forests. Environmetrics. 8:43–61.

Roberts JF, Champion AJ, Dawkins LC, Hodges KC, Shaffrey LC, Stephenson DB, Stringer MA, Thornton HE, Youngman BD. 2014. The XWS open access catalogue of extreme European windstorms from 1979 to 2012. Nat Hazards Earth Syst Sci. 14:2487–2501.

Robeson SM. 1997. Spherical methods for spatial interpolation: review and evaluation. Cartogr Geogr Inf Syst. 24:3–20.

Smits A, Klein Tank AMG, Können GP. 2005. Trends in storminess over the Netherlands, 1962–2002. Int J Climatol. 25:1331–1344.

Steenburgh W, Mass C. 1996. Interaction of an intense extratropical cyclone with complex coastal orography. Mon Weather Rev. 124:1329–1352.

Sterk G, Stein A. 1997. Mapping wind-blown mass transport by modeling variability in space and time. Soil Sci Soc Am J. 61:232–239.

Venäläinen A, Helkinheimo M. 2002. Meteorological data for agricultural applications. Phys Chem Earth Parts A/B/C. 27:1045–1050.

Ver Hoef JM, Cressie N, Barry RP. 2004. Flexible spatial models for kriging and cokriging using moving averages and the fast Fourier transform (FFT). J Comput Graph Stat. 13:265–282.

World Meteorological Organization. 2008. WMO guide to meteorological instruments and methods of observation. Geneva: Chairperson, Publications Board: World Meteorological Organization (WMO).

Zimmerli P. 2003. Natural catastrophes and reinsurance. Switzerland: Swiss Reinsurance Company/Swiss Re Publications.

Zlatev Z, Middleton SE, Veres G. 2009. Ordinary kriging for on-demand average wind interpolation of in-situ wind sensor data. Proceedings of the EWEC 2009. Shanghai, China.

Zlatev Z, Middleton SE, Veres G. 2010. Benchmarking knowledge-assisted kriging for automated spatial interpolation of wind measurements. Proceedings of the 2010 13th Conference on Information Fusion (FUSION). Edinburgh, UK.