GI_Forum 2021, Issue 1
Page: 13 - 23
Research Paper
Corresponding Author:
malini.krishnan@research.iiit.ac.in
DOI: 10.1553/giscience2021_01_s13

Generation of Spatial Profiles & Mapping of Volcanic Ash Distribution

Malini Krishnan[1] and Rajan Krishnan Sundara[1]
[1]International Institute of Information Technology, India

Abstract

Defining spatial distribution of airborne volcanic ash in the neighbourhood of an erupting volcano is a synoptic scale problem, severely impacting lives and livelihoods. Robust algorithms are needed to model such complex phenomenon from sparse field data. This study investigated optimal modelling of the spatial dispersion of ash using Empirical Bayesian Kriging (EBK): a geostatistical, probabilistic algorithm. Both distance and ash temperature values of samples from the 2010 Icelandic eruption were spatially correlated using semivariograms to generate prediction and error surfaces. Results showed that block averages were 90% accurate as validated against NCEP NWP model data. The work supports the utility of EBK in datasets where spatial autocorrelation is not significant. Furthermore, the results could help generate risk maps to delineate safety zones for aircrafts.

Keywords: volcanic ash, kriging, geostatistics, spatial analysis

1 Introduction

Waldo Tobler's First Law of Geography, states "Everything is related to everything else, but near things are more related than distant things." This law provides the foundation of the fundamental concepts in spatial dependence and spatial autocorrelation, and is utilized specifically in spatial interpolation techniques. Spatial autocorrelation (Zhu et al., 2019) is a key concept that is used to analyse the degree of dependency among observations (samples) in a given geographic space. Distance between neighbours, lengths of shared borders, and orientation are just some of the measurements used in conjunction, when modelling a given field, to estimate the unknowns.

When given a random spatial field with unbounded variation causing high or low spatial autocorrelation, it is necessary to analyse how the choice of the geostatistical method can accurately model the variable of interest. This paper will investigate the appropriateness of the spatial interpolation technique Kriging, in particularly for clustered, heteroskedastic datasets.

In addition, the generation of highly accurate prediction estimates, even in severe weather scenarios over synoptic scales: embracing a pure spatial analysis approach can be a powerful method to supplement grid-based models. Deterministic techniques, in general do not model uncertainties accurately. Therefore, stochastic geostatistical methods are needed to model even small-scale spatial variances. To demonstrate and evaluate this, we have chosen a variant of Kriging named Empirical Bayesian Kriging (EBK), and applied it in this study.

Kriging is primarily a spatial algorithm. When spatiotemporal data must be analysed, usually the datasets are either grouped or split based on temporal criteria, to apply kriging, or to study the patterns (van Stein et al., 2020 and Krivoruchko et al., 2020). In this investigation, we chose four main data clusters spatially disjointed in both 2D and 3D (Altitude wise), as well as temporally (across four days). While performing kriging, the assumption was to treat the input data (May 16th, 17th and 18th) samples as pure spatial data. However, the temperature prediction and error estimate outputs have been rigorously evaluated against the available fourth day's test data (May 14th), which in reality, was also spatially and temporally disjoint from the input dataset. A process has been defined on how to customize spatiotemporal data sampled in transects, and appearing spatially random to be redefined as a spatially clustered dataset. Meaning, a technique like EBK, which was primarily designed purely for transect samples, can still be applied in other spatiotemporal contexts. Therefore, the site under study can be modelled as accurately as possible.

2 Study Site

The 2010 eruption of an Icelandic volcano, called Eyjafjallajokull, was selected for this study. The ash was dispersed across the European airspace for several days. Facility for Airborne Atmospheric Measurements (FAAM) aircrafts were flown in-sync with satellite overpasses for multiple days, near potentially hazardous ash laden regions to collect a variety of scientific data. British Atmospheric Data Centre (BADC, 2013) released a subset of the weather data for research purposes.

The data collected by the BOMEM Michelson interferometer over four days (May 14, May16, May 17, May 18) was chosen for this study, and depicted in the Minimum Bounding Region (MBR) created, including the vent location as shown in the Figure 1. While the field sampling durations extended several hours, a small portion of the recorded temperature data considered to be from an ash-significant regions was prepared. The processing involved mapping the attribute data against the flight path information by referring to the discussions made amongst the scientific crew on board the sorties.

Figure 1: Map showing the MBR with Data Locations w.r.t. Volcanic Vent over Europe

For the 4 days of flight data, 16th, 17th and 18th were used as input, while 14th data was considered as test dataset for evaluating the accuracy of estimations. The MBR encompasses around 5 lakh square Kilometers of area. The temperature distribution across those days were compared and plotted in Figure 2.

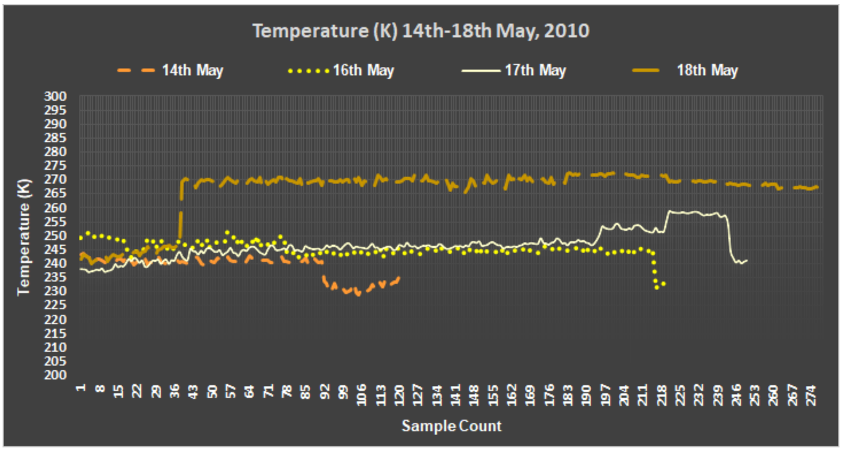

Figure 2: Temperature Distribution Plot of Data Samples

2.1 Validation Dataset

An Ash Dispersal Forecast and Civil Aviation Workshop [9] was conducted post eruption to benchmark dispersion models based on ash & weather data from the Hekla eruption in 2000.Ash concentration contour maps were generated at different flight levels.

While all the operative models were tested and compared based on properties of ash, our paper focuses on temperature variable as a proxy to model the ash dispersion. The NCEP/NCAR (National Centre for Environmental Prediction/National Centre for Atmospheric Research) reanalysis climate/weather dataset from the USA used in the workshop was therefore chosen for validation. Data for each day was downloaded from the repository (NCEP/NCAR 20th Century Reanalysis Weather Data Repository, 2016) according to the pressure altitude of the flight routes, and time duration (set to European Projection configuration).

The initial step was to understand the temperature profiles simulated by Numerical Weather Prediction (NWP) models such as NCEP, theoretically, over continental and oceanic Europe for the same period and region of interest. Daily composites for the period between May 14-May18 were compared annually from 2008-2011, minimum and maximum temperature values predicted at 350/400/700/800 mb Pressure Altitudes it was observed that there were no variations in temperature greater than 8K in total. Contrastingly, May 17th 2010 samples (collected by flight) revealed a variation of up to 22K at very short spatial scales. Furthermore, up to a 27K drop in air temperature was observed on May 17th when compared against the usual Environment Lapse Rate (ELR) (expected at 700 mb).

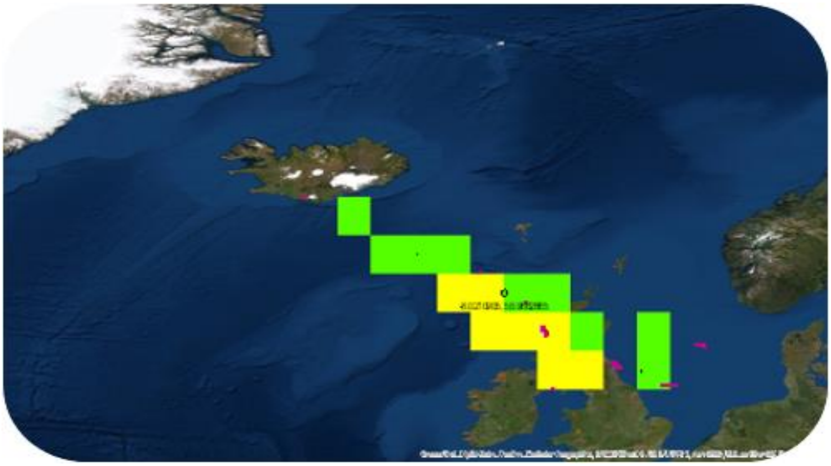

Figure 3: Map showing Overlay of Grids of NCEP Rasters from May 14th to May 18th 2010

Figure 3 clearly shows that coarse grid sizes used in NWP models do not accurately represent the state of the atmosphere even during large volcanic eruptions in any given region. The average temperature of the overlay created from using rasters of each day was ~253K. This paper (Threnbert et al., 1988) describes the interpolation approach used in NCEP models, and, discusses the limitations arising in accuracy of model outputs in the context of large geographic regions.

3 Methodology – Kriging

Linear regression techniques can produce good estimates of global mean, but are not very effective in modelling the observed small-scale variations accurately. Consequently, a robust spatial interpolation technique, based on stochastic geostatistical theory, called Kriging, originally drafted for mining industry, is cross-applied on air temperature data sampled from the affected region, at various altitudes to interpolate values at unknown locations. Kriging or Gaussian Process is a weighted average technique that assigns higher weights on nearby observations, based on the distance and direction characteristics.

The process involves, the generation of a semivariogram, which expresses the rate of change of regionalized variable w.r.t. different distance bands. By interpreting the sampled data as the result of a random process, kriging builds a methodological basis to provide a scope for estimating the spatial inference of quantities in unobserved locations. Kriging is also useful in quantifying uncertainty associated with the estimator since the sample values are expected to be correlated between themselves owing to their locational proximity. Using Linear Mixed Model framework in a Bayesian context, clusters are modelled using EBK. This method calculates, structured drift, spatial variations and errors separately. EBK produces surface outputs for prediction by fitting different transitive functions.

4 Empirical Bayesian Kriging

EBK implemented in ArcGIS software (Gribov et al., 2020 and Krivoruchko et al., 2019) effectively represents the stochastic spatial process locally as non-stationary random field, where the parameters vary across space. Local models are built by simulating multiple theoretical semivariograms, created by sub setting the input data to apply the REML (Restricted Maximum Likelihood Estimation) method.

In EBK, the Bayesian framework estimates only prior distributions using observed marginal distributions. The estimates were predicted by considering temperature concentrations as a response variable; while location variables, derived from flight data, were used as predictors. EBK model is calculated by:

$$\gamma(h) = Nugget + b|h|^{\alpha} \tag{1}$$

ϒ is the semivariance, b is the positive slope; α is the power between 0.25-1.75, Nugget which has a positive value.

4.1 Block Grade – Prediction and Error Estimates

To compare the NCEP temperature averages (measured in Kelvin) with the prediction estimates of kriging, 1x1 degree grids were created. EBK block averages, shown in Figure 4, reveal a narrow range of global temperature estimates: ranging between 241K to 251K. The global mean is ~243K, around 10K less than NCEP average.

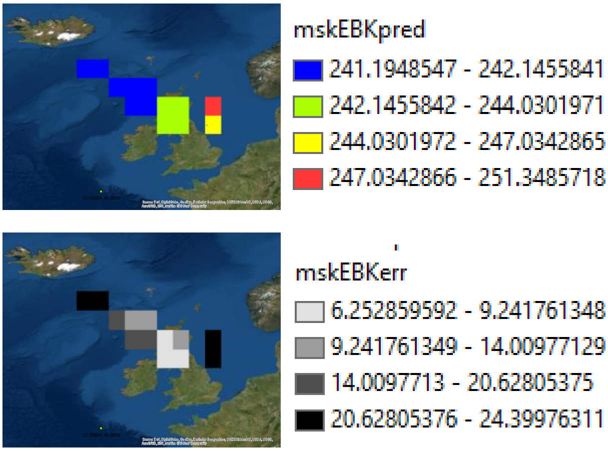

Figure 4: Maps showing Block Grade EBK – Prediction Estimates (above) & Error Estimates (below)

5 Verification & Validation

The interpolated values were verified and validated using the methods below.

5.1 Verification

Error Analysis

While the Root Mean Square (RMS) value is desired to be as low as possible for any interpolation algorithm: a special metric to assess Kriging efficiency is RMS-Standardized, which is expected to be close to 1. EBK had an RMS of 2.596989 and RMSS of 0.938776. RMS values close to zero indicates that the estimates are unbiased. EBK met the criteria with high accuracy (0.018348).

EBK Profile Analysis

Although the correlation between the distances and temperature is low ($R^2 = 0.294$), due to the clustered distribution of the samples, EBK profile (figure 5) reveals a steady decrease in temperature as the distance from the vent gradually increases, as observed in the sampled inputs for the MBR.

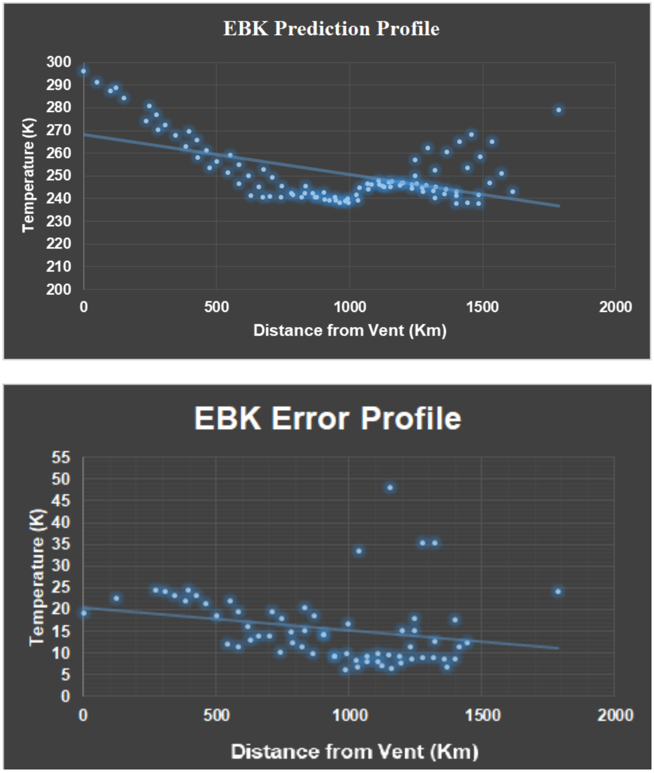

Figure 5: Plot of Temperature Prediction vs Distance Profile (above) and Error Profile (below)

The errors are also not highly correlated with distance, however, are higher in magnitude as the distance from the vent increases.

5.2 Validation

EBK vs NCEP - Profile Analysis

As shown in Figure 6, when EBK averages were validated against the NCEP NWP model values for the same duration in the area of interest, a consistent deviation of 10K was observed. However, the small-scale spatial variations were also accurately estimated using the EBK method with a maximum deviation of ~12K.

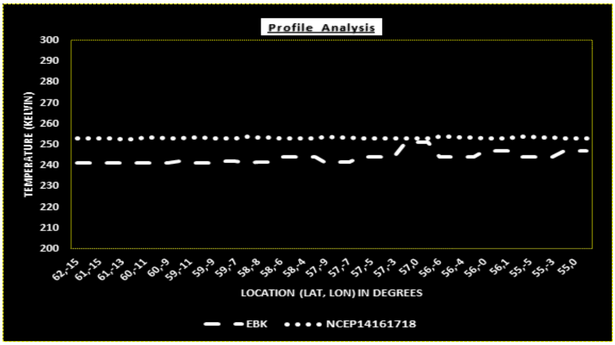

Figure 6: Plot Validating EBK Prediction Profile against NCEP Profile

Figure 7 shows the non-parametric probability density estimation for NCEP and EBK block averages. While EBK estimates had a Standard Deviation of ~3K, NCEP measured at ~0.57K.

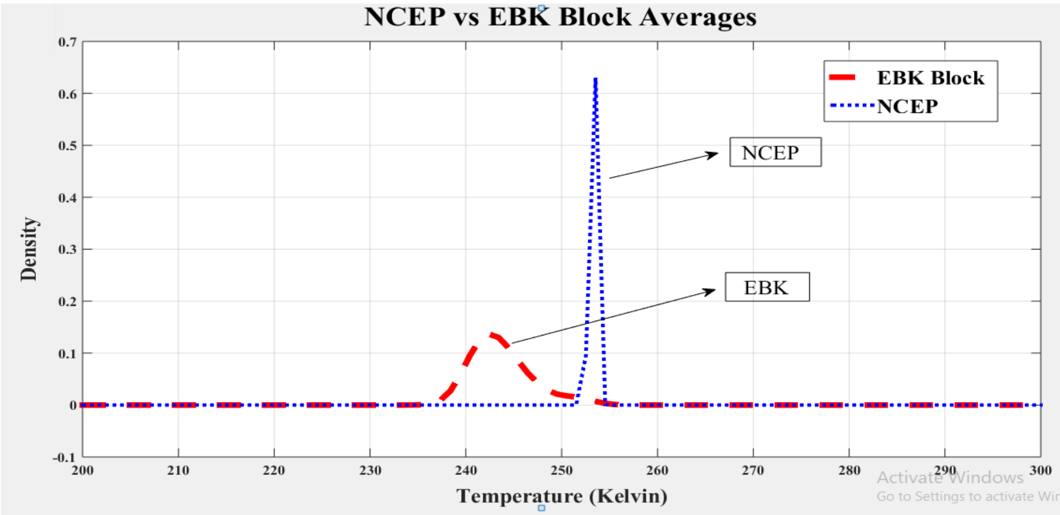

Figure 7: Plot of Probability Density Estimates - EBK Prediction vs NCEP

Against Test Data – 14th May 2010

Out of the four days of samples, three (16th, 17th, 18th May) were used to interpolate data, while one (14th May) was used as test data to validate the predicted results. Figure 8 compares the flight data on May 14th against the kriged output, using 16th/17th/18th data for the same location. Spatially, these test samples were located almost at the centre of the Minimum Bounding Region, and were equidistant from each day's cluster, and the vent. Although altitude information was not used for kriging, The test dataset was from the highest altitude (8000 meter) and hence all values were below 250K. The test dataset had just 122 samples in comparison to the 200+ each from the other 3 datasets, making ideal to be used for verification. Altitude into validation scope solely aided in comparing kriged estimates against NCEP data at specific pressure bands

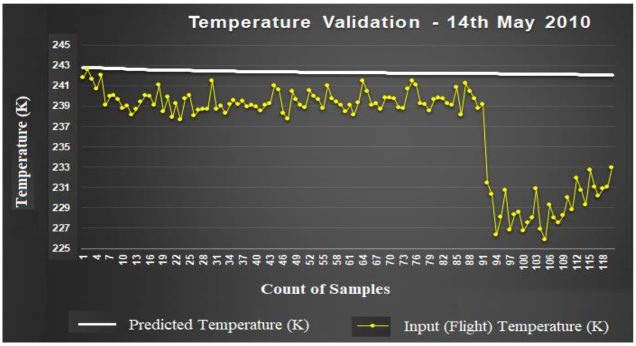

Figure 8: Plot showing Validation of Kriged Temperature Estimates Against May 14th Flight Temperature Samples

The global prediction estimates of EBK using point kriging method had a range spanning approximately 70K. On average, an overestimation error of less than 8K was observed when tested against 14th May 2010 (test data). Thus, the error is within 10% threshold for EBK prediction estimates.

Local Estimates

Prediction and error estimates were grouped into intervals of 5K to compare the input data against the kriged outputs for each day. The comparative visualization in Figure 9 reveals the degree of unbiasedness (<1K global error in locations where each day's temperature data is available). The map below (figure 9) compares the variations observed for input data against the predicted data, where samples from the 14th May were located. This clearly shows EBK is an acceptable exact interpolator for variance.

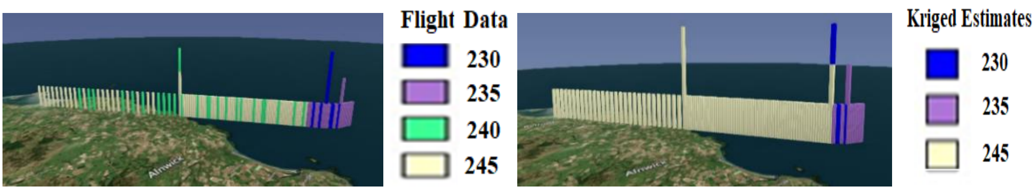

Figure 9: Maps Showing Temperature Variations - Flight Samples against Kriged Estimates on May 14th

6 Results – Discussion

For a three hour forecast of normal weather, the measure of success for prediction of temperatures is defined by UK Met Office (2021) to be within ±2° C 92% of the time it is reported. The smallest size of the grid cell achieved for this study site with kriging was 4x2/2x4 units. The error range for this zone was found to be between 0K-2K. With EBK, the defined success rate was achieved for a spatial resolution as low as 2km x 4km.

In the aerospace industry, this roughly translates the detection of potential ash laden field as early 20 seconds ahead of time by jet aircrafts in cruising altitude with high airspeeds and wind speed conditions. This methodology is highly suited to augment onboard severe weather alert systems, despite its probabilistic origins and simulation scope. The study can also help to define guidelines for sample data collection during future eruptions to assess the safety of an airspace.

7 Mapping Risk Zones

Given a potential use case in the aviation industry, we try to generate Go/No-Go Zones using the point prediction map produced using EBK by comparing against NCEP values. The NCEP has a narrow temperature range of 251.4K-253.9. Figure 10 shows regions with same range of observations highlighted in green (~247K to ~254K). Areas with gradual

Areas with gradual variations in orange reveal EBK underestimations/overestimations against NCEP (±25K), while regions with red depict significant overestimations in comparison against NCEP (~+40K).

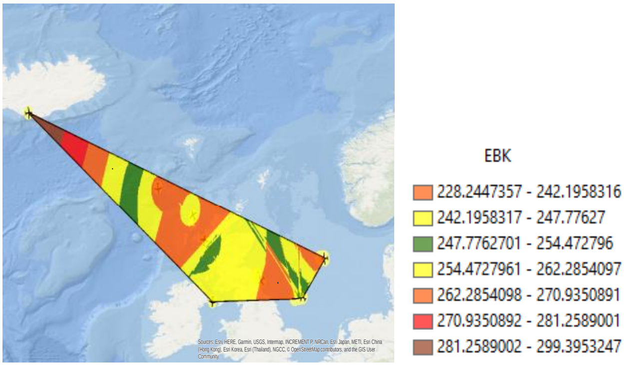

Figure 10: Map Showing Risk Zones Categorized As Go/No-Go Regions

Irrespective of the significant global variations in the input temperature across days, the EBK risk map reflects integration of unbiased global averages and small-scale variations, wherever adequate data is available.

8 Conclusion

In summary, it is observed that the EBK not only produces estimates of block mean with up to 90% accuracy closer to NWP averages, but also models small-scale spatial variances better than NWP models, even at coarser spatial resolutions. In addition, it is also evident that when EBK is applied as a punctual kriging method, it can produce unbiased averages even for spatially clustered, heteroskedastic datasets. Hence, even in nonstationary datasets with absence of significant spatial autocorrelation, EBK can be used to assess the likelihood of volcanic ash concentration exceeding a defined threshold at a given place, so that risk to aviation operations can be determined.

The method involved partitioning the whole dataset into small subsets to model each partition, and then by combining all outputs to predict at unknown locations using a distance metric in a Bayesian framework. The Kriging technique, though originally conceived, designed, and implemented for Gaussian world with higher emphasis on Spatial Autocorrelation, is well suited for ash dispersion modelling. In addition, for smaller datasets, we established that EBK is an appropriate method to model the simultaneous existence of spatial autocorrelation and spatial heterogeneity at different degrees. These are typically observed in events that obey Pareto conditions, and can therefore be used to generate accurate maps for airborne volcanic ash dispersion.

References

Thakur, M., Samanta, B., & Chakravarty, D. (2018). A non-stationary geostatistical approach to multigaussian kriging for local reserve estimation, *Stochastic Environmental Research and Risk Assessment*, 32(8), 2381–2404. https://doi.org/10.1007/s00477-018-1533-1.

Krivoruchko K. & Gribov A. (2014). Pragmatic Bayesian kriging for non-stationary and moderately non- Gaussian data, *Mathematics of Planet Earth: Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences: Lecture Notes in Earth System Sciences, Springer, pp.61-61.* https://doi.org/10.1007/978-3-642-32408-6_15.

Zhu R, Janowicz K, & Mai G. (2019). Making direction a first-class citizen of Tobler's first law of geography. *Transaction in GIS*, 23:398–416. https ://doi.org/10.1002/tgis.12550

Airborne volcanic ash temperature data and flight logs collected of May 2010, using the BAe-146-301 Atmospheric Research Aircraft [ARA], flown by Directflight Ltd & FAAM (NERC & Met Office) obtained via, British Atmospheric Data Center (BADC) repositories. [Accessed on 30 July 2013].

UK Met Office, Global accuracy at a local level, https://www.metoffice.gov.uk/about-us/what/accuracy-and-trust/how-accurate-are-our-public-forecasts. [Accessed on 30 January 2021].

Trenberth & Olson, (1988). The Twentieth Century Reanalysis Project 3. https:/rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.776

Gribov A & Krivoruchko, (2020). Empirical Bayesian kriging implementation and usage. *Science of the Total Environment, 2*0;722:137290. https:// doi: 10.1016/j.scitotenv.2020.137290.

Konstantin Krivoruchko & Alexander Gribov, (2019), Evaluation of empirical Bayesian kriging, *Spatial Statistics, 100368,ISSN:2211-6753.* https://doi.org/10.1016/j.spasta.2019.100368

Bonadonna, C., Folch, A., Loughlin, S. et al (2012). Future developments in modelling and monitoring of volcanic ash clouds: outcomes from the first IAVCEI-WMO workshop on Ash Dispersal Forecast and Civil Aviation. Bull Volcanol 74, 1–10. https://doi.org/10.1007/s00445-011-0508-6

Bas van Stein, Hao Wang et al (2020). Cluster-based Kriging approximation algorithms for complexity reduction, *Applied Intelligence,* 50:778–791 https://doi.org/10.1007/s10489-019-01549-7

Konstantin Krivoruchko & Jorge Mateu, (2020), COVID-19 Epidemic Data Modeling in Space-Time Using Innovation, Diffusion Kriging.

NCEP/NCAR 20th Century Reanalysis Weather Data Repository- https://www.esrl.noaa.gov/psd/cgi-bin/data/composites/plot20thc.day.v2.pl [Accessed on 21 April 2016]