DOI:10.28974/idojaras.2024.2.6

IDŐJÁRÁS

Quarterly Journal of the HungaroMet Hungarian Meteorological Service
Vol. 128, No. 2, April – June, 2024, pp. 237–249

Spatiotemporal imputation of missing rainfall values to establish climate normals

Brian O'Sullivan* and **Gabrielle Kelly**

Department of Mathematics and Statistics,
University College Dublin
UCD Earth Institute

**Corresponding author E-mail: brian.osullivan3@ucdconnect.ie*

(Manuscript received in final form January 23, 2024)

Abstract— Spatial kriging interpolation has been a widely popular geostatistical method for decades, and it is commonly used to predict both gridded and missing climatic variables. Climate data is typically monitored across a variety of timescales, from daily measurements to thirty-year periods, known as long-term averages (LTAs). LTAs can be constructed from daily, monthly, or annual measurements so long as any missing values in the data are infilled first. Although spatial kriging is an available method for the prediction of missing data, it is limited to a single moment in time for each imputation. Not only can missing values only be predicted with observations measured at the same instance in time, but the entire imputation process must be repeated up to the number of timesteps in which missing data is present. This study investigates the imputation performance of spatiotemporal regression kriging, an extension of spatial regression kriging which simultaneously accounts for data across both space and time. Hence, missing data is predicted using observations from other points in time, and only a single imputation process is required for the entire data set.

Spatiotemporal regression kriging has been evaluated against a variety of geostatistical methods, including spatial kriging, for the imputation of monthly rainfall totals for the Republic of Ireland. Across all tests, the spatiotemporal methods presented have outperformed any purely spatial methods considered. Furthermore, three different regression methods were considered when de-trending the data before interpolation. Of those tested, generalized least squares (GLS) was shown to provide the best results, followed by elastic-net regularization when GLS proved computationally unavailable. Finally, the data set has been infilled using the best performing imputation method, and precipitation LTAs are presented for the Republic of Ireland from 1981–2010.

Key-words: spatiotemporal kriging, rainfall, long-term averages, missing data, imputation, kriging, Ireland, elastic-net

1. Introduction

Climate normals or LTAs are the standard measure by which the climate is described, providing the average climatic conditions experienced by a region over a thirty-year period. A thorough understanding of precipitation is crucial to numerous endeavours in Ireland, ranging from agriculture to flood risk management (*Charlton et al.,* 2006; *Naughton et al.,* 2017). During the measuring period, missing data entries commonly arise for rain gauge stations. There are various potential causes for this, such as stations being opened, closed, or moved, malfunctioning equipment, or insufficient observations taken by the station monitor. Missingness present in the data set must first be addressed before continuing with any climatological or hydrological research, including the production of LTAs. This creates the need for robust and sophisticated methods which impute missing entries as accurately as possible in order to minimize bias caused by missingness in future analysis. In this study, the performance of spatiotemporal regression kriging is explored for the purpose of imputing missing monthly precipitation totals. Elastic-net regularization (*Zou* and *Hastie,* 2005) is also investigated as a potential model to de-trend the data for regression kriging.

Kriging is a popular geostatistical method for predicting variables of interest over a spatial field. Developed originally by *Matheron* (1963), it is widely applied in a variety of fields such as environmental science, mining, and remote sensing (*Tavares et al.,* 2008; *Mondal et al.,* 2017). The method interpolates values as weighted averages of observations from nearby stations, where weights are calculated according to the estimated variance between sample stations and the target point. Numerous extensions of kriging have been explored, and popular examples include universal kriging, cokriging, and Bayesian kriging (*Handcock and Stein,* 1993; *Myers,* 1982). Universal kriging is a particularly widespread method that includes a regression of the target variable against auxiliary variables present in the data (elevation, latitude, longitude, etc.). As kriging assumes a second order stationarity across the field, it is necessary to remove any trends initially present in the data. Hence, the popularity of universal kriging when kriging interpolation is employed with climate data. Small variations to universal kriging exist, namely, regression kriging and kriging with external drift. These three methods differ slightly in their implementation, but are all generally the same technique. Regression kriging divides the approach into a two-step process, where trends in the data are first removed by regression and the remaining residuals are interpolated by ordinary kriging. The regression method achieved by universal kriging is known as generalized least squares (GLS), and this approach is theoretically optimal for a linear estimator. However, the modular approach of regression kriging allows one to consider alternative regression methods such as elastic-net regularization or principle component regression (*Zou* and *Hastie,* 2005; *Jolliffe,* 1982).

Originally proposed as a purely spatial method, kriging has been extended to spatiotemporal contexts. Known as spatiotemporal kriging, the distance between points across a temporal field is considered alongside spatial distance, such that interpolation can be achieved using a data set not bound to a single point in time (*Montero et al.,* 2015). This approach lends itself particularly well to imputation, as missing observations can have nearby temporal neighbors either before or after the target point which are observed from the same station. Additionally, the entire data set can be imputed at once, removing the need to undergo a separate kriging procedure for each time step. Spatiotemporal kriging requires the production of more sophisticated spatiotemporal covariance models (*Gräler et al.,* 2016), but otherwise it is formulated similarly to spatial kriging. Spatiotemporal kriging has been applied successfully in many contexts. For example, *Hengl* (2012) predicted daily temperatures through spatiotemporal regression kriging with a sum-metric variogram model. The remainder of this study is reported as follows: First, a short overview of the data is provided. Then, the methodology behind both elastic-net regularization and spatiotemporal regression kriging is outlined. The results comprise of the imputation performance of all considered methods, followed by a brief discussion. All research has been implemented using the *R* programming language (*R Core Team,* 2021), with a particular emphasis on the `glmnet` (*Friedman et al.,* 2010) and `gstat` (*Gräler et al.,* 2016) packages.

2. Methods

The island of Ireland has a temperate oceanic climate with an abundance of rainfall throughout the year (*Lennon,* 2015). Precipitation is monitored by Met Éireann, the Irish meteorological service, using over 1100 rain gauge stations located around the country. The considered data set consists of monthly precipitation totals from 474 stations over a thirty-year monitoring period of 1981–2010. At least 50% of data-entries from these stations are recorded as non-missing. Additionally, the data set has been considered at different levels of completeness, i.e., only considering stations with at least 70% (365 stations), 50% (474 stations), or 30% (679 stations) of their entries recorded non-missing, respectively. The rain gauge distribution of the monitoring network at station completeness cutoffs of 100%, 70%, and 30% are displayed in *Fig. 1*.

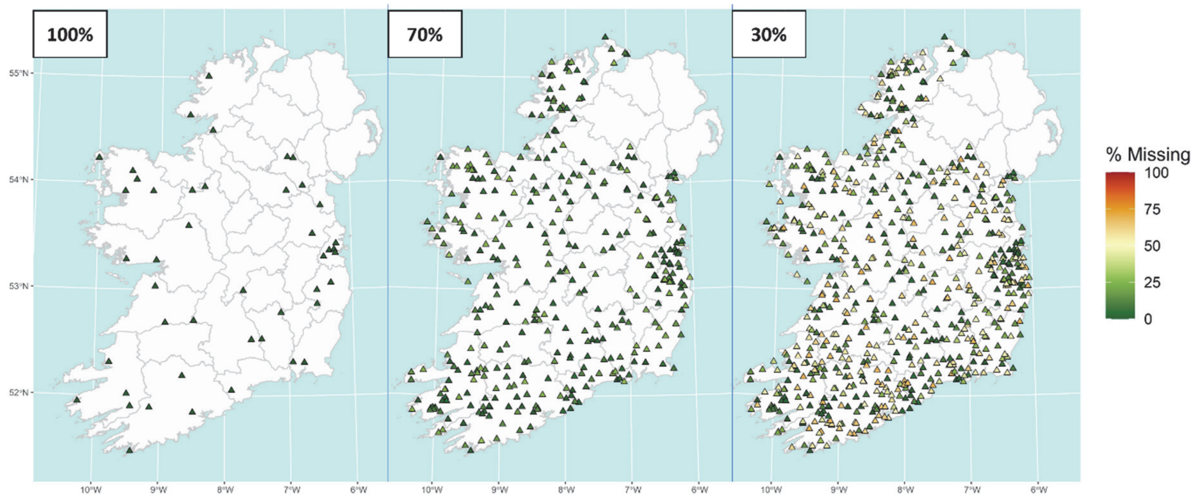

Fig. 1. Met Éireann precipitation monitoring network for the Republic of Ireland from 1981–2010. The network is presented at different completeness cutoffs (100%, 70%, and 30%, respectively).

Typically, when geospatial climate data is interpolated, any trend present in the data is removed. This is a necessity for regression kriging as the ordinary kriging step assumes a second-order stationarity with no external trend or "drift". Drift is described as the continuous change of the underlying target variable. It is modeled as a function of available predictor variables, e.g., longitude, latitude, elevation, etc., and how changes in these variables correspond with changes in the target. For linear effects, this function is simply expressed as the well-known linear model: $y = X\beta + \epsilon$. Removing drift is a fairly straightforward task, and can be done in a variety of ways. Ordinary least squares (OLS) is an elementary option, where regression parameters are determined according to the estimator: $\hat{\beta} = (X^T X)^{-1} X^T y$. However, OLS relies on the assumption that the underlying model residuals are independent and uncorrelated. Upon inspection of the variogram of residuals after regression is conducted (*Fig. 2*), this is clearly not the case. Alternatively, GLS accounts for auto-correlation between residuals by including the variance-covariance matrix of the residuals, C, in the estimator: $\hat{\beta} = (X^T C^{-1} X)^{-1} X^T C^{-1} y$. In combination with kriging, GLS gives the best linear unbiased estimator (BLUP). Besides these two unbiased regression techniques, a third method known as elastic-net regularization has also been considered.

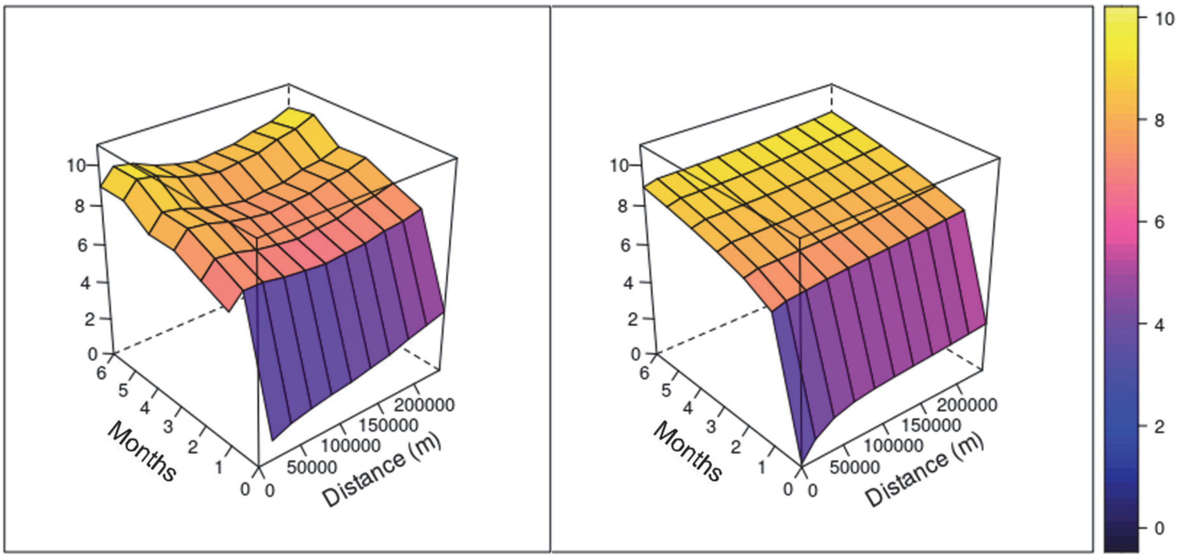

Fig. 2. Left: Empirical spatiotemporal variogram of residuals. Right: Fitted sum-metric variogram model consisting of spatial, temporal, and spatiotemporal Matérn structures.

Developed by *Zou* and *Hastie* (2005), elastic-net regularization is a technique designed to address high-dimensionality and/or highly correlated variables in a regression context. The objective function of elastic-net differs from that of OLS by the addition of a regularization penalty, L, i.e.,

$$\min_{\beta_0,\beta}\left\{\frac{1}{N}\|y-\beta_0-X\beta\|^2+L\right\}. \tag{1}$$

The elastic-net penalty is in fact a convex sum of penalties from two other methods, lasso regression and ridge regression, $L_1=\lambda\|\beta\|_1$ and $L_2=\lambda\|\beta\|_2^2$, respectively. Both penalties are designed to shrink the regression coefficients, β, present in the standard linear model by constraining either the L_1-norm or L_2-norm of all β below a constant, i.e., $\|\beta\|_1=\sum_{i=1}^{p}|\beta_i|<c$ for lasso regression and $\|\beta\|_2^2=\sum_{i=1}^{p}\beta_i^2<c$ for ridge regression. The tuning parameter, λ, determines the degree of shrinkage that is applied to the regression parameters. It is normally preselected before regression or can be fit by 10-fold cross validation using the `cv.glmnet` function available in *R* package, `glmnet` (*Friedman et al.,* 2010). By considering a linear combination of both L_1 and L_2 penalties, the elastic-net penalty is expressed as follows:

$$L_{ENet}=\lambda\sum_{i=1}^{p}(\alpha\beta_i^2+(1-\alpha)|\beta_i|). \tag{2}$$

An additional parameter, $\alpha \in [1, 0]$, is introduced which describes how closely the elastic-net penalty is to lasso ($\alpha = 1$) or ridge ($\alpha = 0$) regression. Elastic-net boasts the advantages of feature selection (from lasso) and robustness in the presence of multicollinearity (from ridge). *Table 1* describes the covariates that were considered in every regression model implemented during this study. As normally distributed data is desirable when using regression prediction methods, a square-root transformation has also been applied to the initial data before imputation. This transformation is reversed at the last step in imputation and has shown to greatly improve the imputation results.

Kriging predicts values by a weighted average method, $\bar{Y} = \sum_{i=1}^{N} w_1 Y(x_i)$, where weights, w, are assigned to nearby observations according to their spatiotemporal relationship with the target. A Gaussian random field, Y, over the spatial and temporal domains, S and T, is assumed, and each value is observed at a distinct point in space and time. That is, $Y(s, t)$ is the total monthly precipitation measured by the station located at s for the month, t. The target variable is considered as a sum of deterministic components, $m(s, t)$, and random components, $\epsilon(s, t)$:

$$Y(s, t) = m(s, t) + \epsilon(s, t). \tag{3}$$

The central idea of kriging is the assumption that Y is second-order stationary so long as the deterministic component, $m(s, t)$, is constant. Here is why any drift present in the data must initially be removed before kriging. Once second-order stationarity is achieved, the covariance between any pair of observations does not depend on their positions, but only on the distance between them. This allows the introduction of the variogram which models the semivariance of point pairs, i.e., the dependence between them with respect to their separation:

$$\gamma(s_1, t_1, s_2, t_2) = \frac{1}{2} Var\{Y(s_1, t_1) - Y(s_2, t_2)\} = \gamma(h, u). \tag{4}$$

All data entries, $x = x(s, t)$, are separated both spatially, h, and temporally, u. The construction of γ is done over multiple steps. An empirical variogram is first created from the observed data, where all available point pairs are grouped into bins according to their separation. The average γ of each bin is calculated and plotted, upon which a parametric representation of γ is fitted to the empirical variogram using the limited-memory BFGS algorithm in `gstat` (L-BFGS). Generally, the parametric form of a purely spatial variogram is as follows:

$$\gamma(h) = \tau^2 + \sigma^2(1 - \rho(h)) \tag{5}$$

The correlation, $\rho(h)$, is a monotonic decreasing function where $\rho(0) = 1$ and $\rho(h) = 0$ as $h \to \infty$. Three parameters are needed to represent the variogram model – the nugget τ^2, the sill $\tau^2 + \sigma^2$, and the range ϕ (*Diggle* and *Giorgi,* 2019). For Matérn covariance models, an additional parameter must be considered, κ, which is known as the "shape".

With these parameters, various covariance models can be fit to the spatial empirical variogram. Popular models include the Spherical, Gaussian, Exponential, and Matérn models. A Matérn structure was used throughout this study and was found to provide the best imputation performance. Its structure contains a modified Bessel function of order, $\kappa(K_\kappa(h/\phi))$, and a Gamma function $\Gamma(\kappa)$. It is presented below:

$$\rho(h) = \{2^{\kappa-1}\Gamma(\kappa)\}^{-1}(h\phi^{-1})^{\kappa}K_\kappa(h\phi^{-1}). \quad (6)$$

A variety of covariance models are available for the spatiotemporal context (*Gräler et al.,* 2016). The sum-metric structure was applied by *Hengl et al.* (2012) for daily temperature data and has been found to be similarly suitable for this study. The sum-metric model consists of three components, a spatial variogram $\gamma_s(h)$, a temporal variogram $\gamma_t(u)$, and a joint variogram, $\gamma_{joint}(\sqrt{h^2 + (\chi \cdot u)^2})$:

$$\gamma(h,u) = \gamma_s(h) + \gamma_t(u) + \gamma_{joint}(\sqrt{h^2 + (\chi \cdot u)^2}). \quad (7)$$

Each component has been modeled using a Matérn correlation structure (Eqs. 5 and 6) with their own distinct fitted parameters (σ^2, τ^2, ϕ, and κ). The third term, γ_{joint}, also contains an anisotropy term, χ, allowing temporal separation to be scaled relative to an equivalent spatial distance. The anisotropy adds an additional parameter to the sum-metric model, bringing the number of parameters needed to be fit by L-BFGS to thirteen (four for each Matérn structure and χ). *Fig. 2* displays the fitted spatiotemporal variogram for the de-trended precipitation data. Notably, the dependency is observed to be much stronger spatially than temporally. As the data is expressed in monthly time steps, the weaker temporal dependency may be attributed to the long period of time between data entries.

Once a variogram of the residuals has been produced and fitted, prediction weights, w, can be calculated through the kriging system of equations. For ordinary kriging, the mean, $m(s,t)$, is assumed unknown, however, this can be relaxed by introducing the constraint that all weights add to one: $\sum_{i=1}^{M} w_i = 1$. The missing value at target point, $x_0 = x(s_0, t_0)$, is predicted by a weighted average of the $M = 700$ nearest observations over all space and time, x_i, $by\ \sum_{i=1}^{M} w_i Y(x_i)$. The ordinary kriging system of equations is given by Eq. (8), where $\gamma(x_i, x_j)_{M\times M}$ is an $M \times M$ matrix of the semivariances between all observed point pairs considered in the kriging system, and $\gamma(x_0, x_i)_{M\times 1}$ is a column vector of the semivariances between the target and the observed points:

$$\gamma(x_i, x_j)_{M\times M} w_{M\times 1} = \gamma(x_0, x_i)_{M\times 1} \qquad \sum_{i=1}^{N} w_i = 1 \tag{8}$$

Inverse distance weighting (IDW) is a simple benchmark method which also predicts target points by weighted average. Instead of modeling a dependency structure like kriging however, an inverse spatial power law is assumed, and the corresponding weights are described as $w = h^{-2}$. The imputation performance of three methods have been evaluated: inverse distance weighting, spatial regression kriging, and spatiotemporal regression kriging. Any trend present in the data (*Table 1*) has been removed prior to interpolation for all methods. Furthermore, three trend removal procedures were considered for spatiotemporal kriging: OLS, elastic-net, and GLS. When implementing GLS, the variance-covariance matrix, C, is obtainable through the modeled theoretical variogram. This is ultimately achieved by an iterative process, where residuals are first obtained to produce the variogram, and regression is repeated using GLS and the now modeled dependency structure. Unfortunately for larger data sets, inversion of C proved computationally unattainable. Tests have been conducted on two data sets from 1981–2010: 474 stations across the Republic of Ireland and a smaller subset of 27 stations in Greater Dublin. Inclusion of the smaller data set allows GLS spatiotemporal kriging to be considered. The imputation of each method was evaluated under 10-fold cross-validation using three performance metrics: root mean squared error (RMSE), relative RMSE ($RMSE_R$), the RMSE normalized by the deviation of the observed data, and R^2, the percentage of variance explained between observed (y) and predicted values ($\hat{y}$):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}}; \quad RMSE_R = \frac{RMSE}{\sigma}; \quad R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}. \tag{9}$$

Table 1. All covariates used when removing trends in precipitation data by regression.

Covariate	Description
east & east^2	Easting, Irish Grid TM75 (m/m^2)
north & north^2	Northing, Irish Grid TM75 (m/m^2)
east × north	Easting/Northing interaction (m^2)
points5 & points5^2	Mean elevation in 5 km radius around station (m/m^2)
exp25k	Ocean cover within 25 km radius of station (%)
t	Time of observation in months from January 1981 to December 2010 (i.e., $t \in (1, 360)$)

3. Results

Imputation performance is reported in *Table 2*. Across all three metrics (RMSE, $RMSE_R$, and R^2), it is clear that the inclusion of a temporal component in the imputation method greatly benefits the prediction accuracy for both large and small data sets. Moreover, this improvement is shown to be consistent throughout the year in *Fig. 3*, where the spatiotemporal methods tested provide a steady improvement for every month. With regards to the regression method applied, elastic-net has shown to slightly outperform OLS in both cases, although GLS expectedly produces the best results when available. Notably, the choice of regression method has a considerably lower impact on imputation accuracy when compared to the introduction of spatiotemporal approaches.

Table 2. 10-fold cross validation results for Republic of Ireland and Greater Dublin. For the 474 stations, ST-kriging with GLS was unavailable due to computational intractability

	Republic of Ireland (474 stations)			**Greater Dublin (27 stations)**		
Method	RMSE	$RMSE_R$	R^2	RMSE	$RMSE_R$	R^2
IDW	21.21mm	30.28%	0.909	15.69 mm	29.81%	0.911
Spatial Kriging	20.61mm	29.41%	0.914	18.13 mm	34.44%	0.884
ST-Kriging	17.47mm	24.94%	0.938	14.05 mm	26.70%	0.929
ST-Kriging ENET	17.42mm	24.86%	0.938	13.89 mm	26.39%	0.930
ST-Kriging GLS	-	-	-	13.80 mm	26.23%	0.931

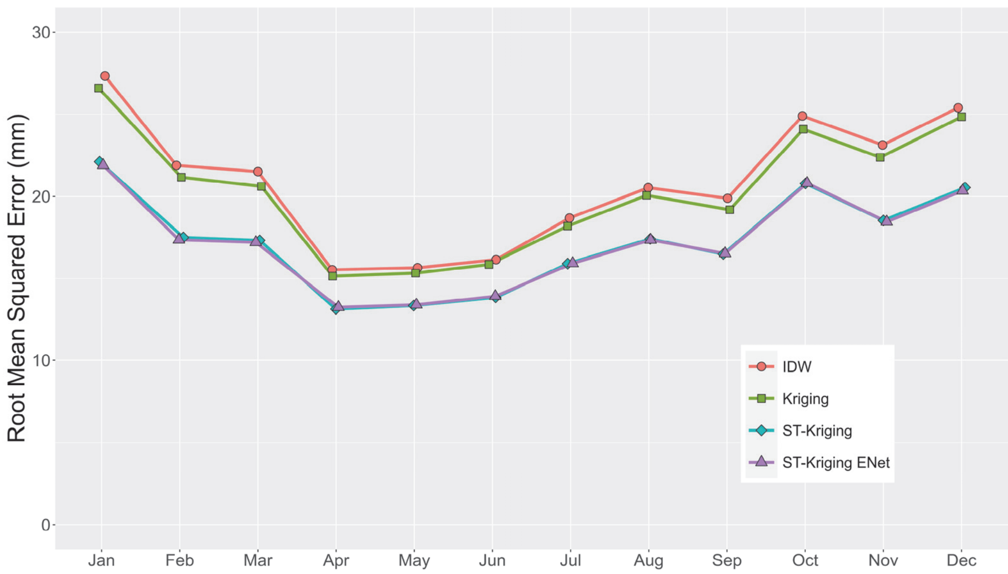

Fig. 3. Monthly RMSE calculated by 10-fold cross validation of the network with a completeness cutoff of 50%. Spatiotemporal methods are shown to consistently outperform throughout the year.

In addition to *Table 2*, further tests have been conducted where the data set comprises of increasingly incomplete stations. Three data sets have been tested where the included stations have a completeness of at least 70%, 50%, or 30%, respectively. The best performing method for large data sets, spatiotemporal kriging with elastic-net, was evaluated alongside these data sets to explore how the method performs against progressively sparser data. Results from *Table 3* demonstrate that the inclusion of less complete stations improves the overall RMSE. However, no improvement in $RMSE_R$ was observed between a cutoff completeness of 50% and 30%. $RMSE_R$ allows imputation performance to be compared across different data sets, a benefit which is unavailable to RMSE (*Hengl,* 2007). Thus the results indicate that the appropriate completeness cutoff to consider during imputation lies somewhere close to the range of 50% to 30%.

Table 3. 10-fold cross validation results according to different completeness cutoffs. Imputation is achieved by spatiotemporal kriging with elastic-net.

% Missing	**No. Stations**	**RMSE (mm)**	**$RMSE_R$ (%)**	**R^2**
70%	365	17.89	25.18	0.937
50%	474	17.42	24.86	0.938
30%	679	17.15	24.86	0.938

Once the data set is fully imputed, LTAs can finally be produced. *Fig. 4* demonstrates the monthly precipitation LTAs for the island of Ireland, created using the fully imputed data set with a 50% completeness cutoff. All missing data entries were first imputed by elastic-net spatiotemporal kriging, then monthly averages were calculated and interpolated on to a 1 km × 1 km grid. Notably from *Fig. 1*, no rain gauge stations were available for Northern Ireland in this study, and as such, the interpolated values are not expected to sufficiently represent the precipitation over this region. Overall, a mean monthly rainfall of mm is reported from 1981–2010, with an increase in rainfall observed in the west and southwest of the island, particularly during the winter months.

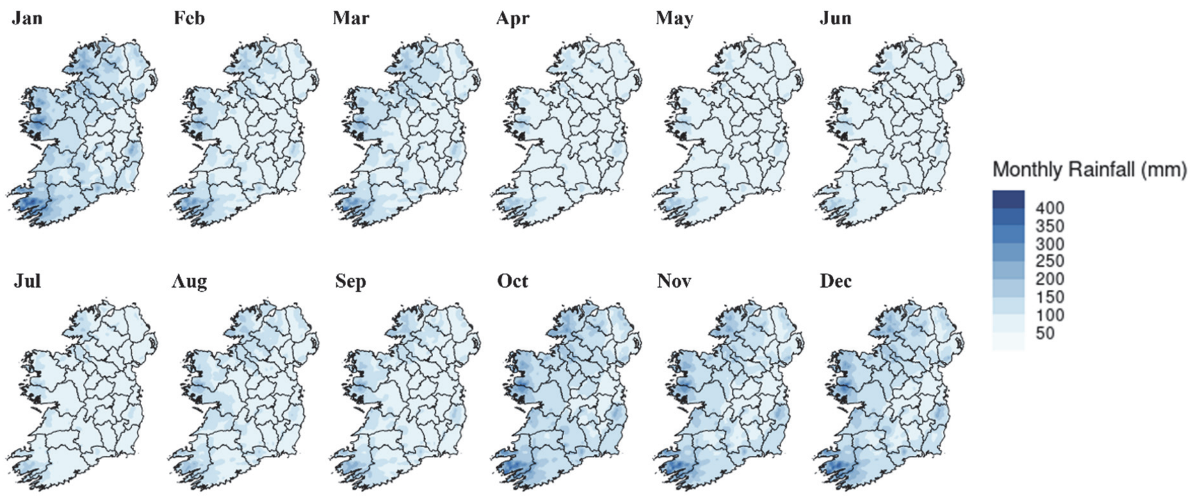

Fig. 4. Gridded long-term averages of monthly precipitation from 1981–2010 (1 km × 1 km).

4. *Conclusion*

A spatiotemporal regression kriging method has been demonstrated to show improved imputation capabilities when compared to its purely spatial counterpart. Furthermore, elastic-net regularization has shown to be a suitable regression method to remove any trend present in a climatic data set. Although GLS provides slightly better results, it is not always available when working with large data sets, which is typically the case in a spatiotemporal context. Given the achieved results, spatiotemporal kriging is presented as a viable option for the imputation of incomplete precipitation data sets. The small adjustment from OLS to elastic-net when removing trend may also be worthwhile, as improved imputation can be achieved for very little additional computational cost. For spatiotemporal kriging, however, it is noted that the increased computation is significant, even when GLS is not considered. The recorded computational time needed to impute for the 50% completeness cutoff 474 stations using spatial kriging was 29.29 seconds, far smaller than the 6.48 hours required for spatiotemporal kriging.

For future research, improving the computational viability of GLS is of utmost concern, potentially via a likelihood-based or Bayesian approach to model fitting. Additionally, the entire island of Ireland may be considered by including data from the Northern Ireland rainfall monitoring network managed by the United Kingdom Meteorological Office. This is generally standard procedure in other climatological research from Met Éireann (*Walsh,* 2016), and would allow for a more comprehensive overview of the complete island. Spatiotemporal kriging may also lend itself well in the imputation for other climatic variables such as temperature or wind speed. Evidently, the temporal dependency is much stronger amongst such variables, a property that which would allow spatiotemporal kriging to yield substantial improvements. However, for Ireland, the observation coverage of rainfall is considerably higher than many other variables, and the lack of sufficient data may limit the capabilities of these more sophisticated imputation methods.

Acknowledgements: We wish to extend our sincere gratitude to Met Éireann for their invaluable provision of the monthly precipitation data for Ireland used throughout this study.
This work was conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number SFI 18/CRT/6049.

References

Charlton, R., Fealy, R., Moore, S., Sweeney, J., and *Murphy, C.,* 2006: Assessing the Impact of Climate Change on Water Supply and Flood Hazard in Ireland Using Statistical Downscaling and Hydrological Modelling Techniques. *Climatic Change 74,* 475–491. https://doi.org/10.1007/s10584-006-0472-x

Diggle, P. and *Giorgi, E.,* 2019: *Model-based Geostatistics for Global Public Health: Methods and Applications.* Chapman and Hall/CRC, New York. 35–37. https://doi.org/10.1201/9781315188492

Friedman, J. H., Hastie, T., ***and*** *Tibshirani, R.,* 2010: Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Software 33,* 1–22. https://doi.org/10.18637/jss.v033.i01

Gräler, B., Pebesma, E., and *Heuvelink, G.,* 2016: Spatio-Temporal Interpolation using gstat. *The R Journal 8, 1,* 204–218. url: https://journal.r-project.org/archive/2016/RJ-2016-014/index.html

Handcock, M. S. and *Stein, M. L.,* 1993: A Bayesian Analysis of Kriging. *Technometrics 35,* 403–410. https://doi.org/10.2307/1270273

Hengl, T., 2007: A Practical Guide to Geostatistical Mapping of Environmental Variables. Office for Official Publications of the European Communities, Luxembourg, 118–119. https://publications.jrc.ec.europa.eu/repository/handle/JRC38153

Hengl, T., Heuvelink, G. B. M., Tadíc, M. P., and *Pebesma, E. J.,* 2012: Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theor. Appl. Climatol. 107*, 265–277. https://doi.org/10.1007/s00704-011-0464-2

Jolliffe, I. T., 1982: A Note on the Use of Principal Components in Regression. *J. Roy. Stat. Soc. Series C: Appl. Stat. 31,* 300–303. https://doi.org/10.2307/2348005

Lennon, J. J., 2015: Potential impacts of climate change on agriculture and food safety within the island of Ireland. *Trends Food Sci. Technol. 44,* 1–10. https://doi.org/10.1016/j.tifs.2014.07.003

Matheron, G., 1963: Principles of geostatistics. *Econ. Geol. 58,* 1246–1266. https://doi.org/10.2113/gsecongeo.58.8.1246

Mondal, A., Khare, D., Kundu, S., Mondal, S., Mukherjee, S., and *Mukhopadhyay, A.,* 2017: Spatial soil organic carbon (SOC) prediction by regression kriging using remote sensing data. *The Egyptian J. Remote Sens. Space Sci. 20,* 61–70. https://doi.org/10.1016/j.ejrs.2016.06.004

Montero, J., Fernándes-Avilés, G., and *Mateu, J.,* 2015: Spatial and Spatio-Temporal Geostatistical Modeling and Kriging. Wiley, New York. 266–279. https://doi.org/10.1002/9781118762387

Myers, D.E., 1982: Matrix formulation of co-kriging. *Math. Geol. 14,* 249–257. https://doi.org/10.1007/BF01032887

Naughton, O., Johnston, P. M., and *Gill, L. W.,* 2017: Groundwater flood risk mapping and management: examples from a lowland karst catchment in Ireland. *J. Flood Risk Manage. 10,* 53–64. https://doi.org/10.1111/jfr3.12145

R Core Team, 2021: *R:* A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Tavares, M.T., Sousa, A.J., and *Abreu, M.M.,* 2008: Ordinary kriging and indicator kriging in the cartography of trace elements contamination in São Domingos mining site (Alentejo, Portugal). *J. Geochem. Explor. 98,* 43–56. https://doi.org/10.1016/j.gexplo.2007.10.002
Walsh, S., 2016: Long-term rainfall averages for Ireland, 1981–2010*.* 15, Climatological Note, Met Éireann. http://hdl.handle.net/2262/76135
Zou, H. and *Hastie, T.,* 2005: Regularization and Variable Selection via the Elastic Net. *J. Roy.Stat. Soc. Series B (Stat. Methodol.) 67,* 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x