Modelling Big, Heterogeneous, Non-Gaussian Spatial and Spatio-Temporal Data using **FRK**

Matthew Sainsbury-Dale **Andrew Zammit-Mangion** **Noel Cressie**

University of Wollongong

Abstract

FRK is a user-friendly R package for spatial/spatio-temporal modeling and prediction with very large data sets that, until recently, has only supported Gaussian data models. In this paper, we describe a major upgrade to **FRK** that allows for non-Gaussian data to be analyzed in a generalized linear mixed model (GLMM) framework. This upgrade allows the analyst to do big–data spatio-temporal non-Gaussian modeling with point-referenced and areal data simultaneously, and predictions or forecasts at any spatial support.

Keywords— areal data, big data, change-of-support, fixed rank kriging, non-Gaussian data, spatial statistics

1 Introduction

Non-Gaussian spatial and spatio-temporal data arise from a vast array of sources. The statistical modeling of these data is pertinent, as accurate predictions, and uncertainty quantification of those predictions, give informed answers to real-world problems (Wikle *et al.*, 2019).

Despite the many modeling approaches available, software for spatial and spatio-temporal model fitting with non-Gaussian data is relatively limited. Some examples include the R packages (R Core Team, 2021) **ngspatial** (Hughes, 2014), **spBayes** (Finley *et al.*, 2015), **mgcv** (Wood, 2017), **spNNGP** (Finley *et al.*, 2020), and **georob** (Papritz, 2020). Each of these packages is limited in some way: Some cater for point-referenced spatial data only; for computational reasons, some are limited to modeling only smooth spatial processes; some are not designed for large data sets; and some cater for only a small number of non-Gaussian distributions. Further, these software packages do not cater for spatial change-of-support (i.e., the handling of data and processes at different levels of spatial aggregation). Some general-purpose packages can, in principle, handle the wide array of modeling challenges posed by non-Gaussian spatial and spatio-temporal data; however, they are not specifically designed for this purpose and can be difficult for an unfamiliar user to implement.

FRK (**?**) is an R package for spatial/spatio-temporal statistical modeling and prediction based on fixed rank kriging, developed by **??** Our article presents a major upgrade to **FRK** that allows one to cater for many distributions within the exponential family using the spatial GLMM framework (Diggle *et al.*, 1998); we henceforth refer to it as **FRK** v2 and the original version as **FRK** v1. **FRK** v2 provides a unifying framework that handles large, spatial and spatio-temporal non-Gaussian (and Gaussian) data, and it can seamlessly ingest point-referenced and area-referenced data to solve spatial change-of-support problems. User-friendliness is a central focus of the package: Challenging statistical analyses may be tackled with only a few lines of intuitive, readable code, requiring a minimal number of user-level decisions. **FRK** v2 also allows for the use of substantially more basis functions than **FRK** v1 when modeling the spatial process. Therefore, in a Gaussian setting, it is often able to achieve more accurate predictions than **FRK** v1. **FRK** v2 has been made available for download from CRAN.

The remainder of the paper is organized as follows. In Section 2 we establish the statistical framework for **FRK** v2. In Section 3 we demonstrate new features in **FRK** v2 using several illustrative examples. Section 4 concludes the paper.

2 Methodology

The statistical model used in **FRK** v2 is a spatial or spatio-temporal hierarchical statistical model consisting of two conditional-probability layers; the *process layer* (Section 2.1) and the *data layer* (Section 2.2). In Section 2.3, we discuss parameter estimation; in Section 2.4, we discuss spatial prediction and uncertainty quantification of the predictions; and in Section 2.5, we present our approach to modeling spatio-temporal data.

2.1 The process layer

For ease of exposition we discuss the spatial case here; the spatio-temporal case is briefly discussed in Section 2.5. We denote the latent spatial process as $Y(\cdot) \equiv \{Y(\boldsymbol{s})\colon \boldsymbol{s} \in D\}$, where $\boldsymbol{s}$ indexes space in the spatial domain of interest D. The model for $Y(\cdot)$ is

$$Y(\boldsymbol{s}) = \boldsymbol{t}(\boldsymbol{s})^\top \boldsymbol{\alpha} + v(\boldsymbol{s}) + \xi(\boldsymbol{s}); \quad \boldsymbol{s} \in D, \tag{1}$$

where spatially referenced covariates $\boldsymbol{t}(\cdot)$ with associated regression parameters $\boldsymbol{\alpha}$ capture spatial variation that is linked to known, usually large-scale, explanatory variables; the spatially correlated random effect $v(\cdot)$ captures medium-to-small-scale spatial variation; and the 'almost' uncorrelated random process $\xi(\cdot)$ captures fine-scale spatial variation.

The medium-to-small-scale term $v(\cdot)$ is constructed as

$$v(\boldsymbol{s}) = \sum_{l=1}^{r} \phi_l(\boldsymbol{s})\eta_l = \boldsymbol{\phi}(\boldsymbol{s})^\top \boldsymbol{\eta}; \quad \boldsymbol{s} \in D,$$

where $\boldsymbol{\eta} \equiv (\eta_1, \dots, \eta_r)^\top$ is an r-dimensional vector of random coefficients for the r-dimensional vector $\boldsymbol{\phi}(\cdot) \equiv \big(\phi_1(\cdot), \dots, \phi_r(\cdot)\big)^\top$ of pre-specified spatial basis functions. Cressie and Johannesson (2008) called this the spatial random effects (SRE) model. The fine-scale term, $\xi(\cdot) \equiv \{\xi(\boldsymbol{s}) : \boldsymbol{s} \in D\}$, is modeled as white noise after discretisation, which we discuss next.

To cater for different observation supports and facilitate solutions to spatial change-of-support problems, we assume a discretized domain of interest, $D^G \equiv \{A_i : i = 1, \dots, N\}$, that is made up of N small, non-overlapping basic areal units (BAUs) such that $D = \cup_{i=1}^N A_i$. Let $Y(A_i)$ denote a representative value of $\{Y(\boldsymbol{s}) : \boldsymbol{s} \in A_i\}$, where commonly that value is the spatial integral or the spatial average over A_i. Define the discretized latent spatial process $Y(\cdot)$ evaluated over the N BAUs as $\boldsymbol{Y} \equiv (Y_1, \dots, Y_N)^\top$, where $Y_i \equiv Y(A_i)$, $i = 1, \dots, N$. Then, a vectorised version of (1) is

$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{\alpha} + \boldsymbol{S}\boldsymbol{\eta} + \boldsymbol{\xi}, \tag{2}$$

where $\boldsymbol{T}$ and $\boldsymbol{S}$ are known design matrices constructed from $\boldsymbol{t}(\cdot)$ and $\boldsymbol{\phi}(\cdot)$ respectively, $\boldsymbol{\alpha}$ is a fixed effect, and $\boldsymbol{\xi}$ is a vector associated with the fine-scale process which, like $\boldsymbol{\eta}$, is treated as a random effect.

The elements of $\boldsymbol{\xi}$ are often modeled as i.i.d. Gaussian random variables with mean zero and variance proportional to σ^2_ξ, and $\boldsymbol{\eta}$ is modeled as a mean-zero multivariate-Gaussian random vector. In **FRK** v2, $\mathrm{cov}(\boldsymbol{\eta}, \boldsymbol{\eta})$ is modeled either as positive-definite $\boldsymbol{K}$ or $\boldsymbol{Q}^{-1}$, where $\boldsymbol{Q}$ is a positive-definite precision matrix. Irrespective of the specific parameterization, we assume that $\mathrm{cov}(\boldsymbol{\eta}, \boldsymbol{\eta})$ depends on a parameter vector $\boldsymbol{\vartheta}$, which needs to be estimated.

Following standard generalised-linear-model theory (McCullagh and Nelder, 1989), we use an invertible link function, $g(\cdot)$, to model $Y(\cdot)$ as a transformation of a mean process, $\mu(\cdot)$:

$$g\left(\mu(\boldsymbol{s})\right) = Y(\boldsymbol{s}); \quad \boldsymbol{s} \in D. \tag{3}$$

The mean process evaluated over the BAUs is $\boldsymbol{\mu} \equiv g^{-1}(\boldsymbol{Y})$, where $g^{-1}(\cdot)$ is applied element-wise.

2.2 The data layer

We denote the vector of m observations (the data vector) as $\boldsymbol{Z} \equiv (Z_1, \dots, Z_m)^\top$. We define the set of observation supports in terms of BAUs as $D^O \equiv \{B_j : j = 1, \dots, m\}$, where $B_j \equiv \cup_{i \in c_j} A_i$ and c_j denotes the indices of the BAUs associated with datum Z_j.

Define the conditional mean of the data as $\boldsymbol{\mu}_Z \equiv \big(\mathbb{E}(Z_1 \mid \boldsymbol{\mu}), \dots, \mathbb{E}(Z_m \mid \boldsymbol{\mu})\big)^\top$. Since each $B_j \in D^O$ is either a BAU or a union of BAUs, one can construct an $m \times N$ matrix

$$\boldsymbol{C}_Z \equiv \Big(w_{ij}\mathbb{I}(i \in c_j) : i = 1, \dots, N; j = 1, \dots, m\Big), \tag{4}$$

where $\mathbb{I}(\cdot)$ is the indicator function, such that

$$\boldsymbol{\mu}_Z = \boldsymbol{C}_Z \boldsymbol{\mu}. \tag{5}$$

The matrix $\boldsymbol{C}_Z$ aggregates the BAU-level process $\boldsymbol{\mu}$ over the observation supports and, depending on the weights in (4), it can correspond to a weighted average or a weighted sum over the BAUs.

We assume that all observations are conditionally independent given the latent spatial process, and that they are all from the same exponential-family member: Specifically,

$$[\boldsymbol{Z} \mid \boldsymbol{\mu}_Z, \psi] = \prod_{j=1}^{m} \mathrm{EF}(\mu_{Zj}, \psi),$$

where EF corresponds to a probability distribution in the exponential family with dispersion parameter ψ and, for generic random quantities A and B, $[A \mid B]$ denotes the probability distribution of A given B.

The model employed by **FRK** v2 can be summarized as follows.

$$Z_j \mid \boldsymbol{\mu}_Z, \psi \overset{\text{ind}}{\sim} \text{EF}(\mu_{Z_j}, \psi); \quad j = 1, \dots, m, \tag{6}$$
$$\boldsymbol{\mu}_Z = \boldsymbol{C}_Z \boldsymbol{\mu}, \tag{7}$$
$$g(\boldsymbol{\mu}) = \boldsymbol{Y}, \tag{8}$$
$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{\alpha} + \boldsymbol{S}\boldsymbol{\eta} + \boldsymbol{\xi}, \tag{9}$$
$$\boldsymbol{\eta} \mid \boldsymbol{\vartheta} \sim \text{Gau}(\boldsymbol{0}, \boldsymbol{Q}^{-1}), \tag{10}$$
$$\boldsymbol{\xi} \mid \sigma^2_\xi \sim \text{Gau}(\boldsymbol{0}, \sigma^2_\xi \boldsymbol{V}), \tag{11}$$

where $\boldsymbol{V}$ is a known, positive-definite, diagonal matrix.

2.3 Estimation

The complete-data likelihood function for our model is

$$L(\boldsymbol{\theta}; \boldsymbol{Z}, \boldsymbol{\eta}, \boldsymbol{\xi}) \equiv [\boldsymbol{Z}, \boldsymbol{\eta}, \boldsymbol{\xi} \mid \boldsymbol{\theta}] = [\boldsymbol{Z} \mid \boldsymbol{\mu}_Z, \psi][\boldsymbol{\eta} \mid \boldsymbol{\vartheta}][\boldsymbol{\xi} \mid \sigma^2_\xi],$$

where $\boldsymbol{\theta} \equiv (\boldsymbol{\alpha}^\top, \boldsymbol{\vartheta}^\top, \sigma^2_\xi, \psi)^\top$. The observed-data likelihood is found by integrating out the unobserved random effects $\boldsymbol{u} \equiv (\boldsymbol{\eta}^\top, \boldsymbol{\xi}^\top)^\top$ from $L(\boldsymbol{\theta}; \boldsymbol{Z}, \boldsymbol{\eta}, \boldsymbol{\xi})$. When the data are non-Gaussian, this integral is typically intractable and must be approximated. In **FRK** v2, we use a Laplace approximation (see, e.g., Tierney and Kadane, 1986), implemented using the `R` package **TMB** (Kristensen *et al.*, 2016).

2.4 Prediction and uncertainty quantification

The Laplace approximation of the conditional distribution of $\boldsymbol{u}$ is an (unnormalised) Gaussian distribution; since $\boldsymbol{Y}$ is a linear function of $\boldsymbol{u}$, Laplace-approximate inference on $\boldsymbol{Y}$ can be carried out using well-known formulas. However, the posterior distribution of a non-linear function of $\boldsymbol{Y}$, for example the mean $\boldsymbol{\mu}$ in (8), is typically not available in closed form. In **FRK** v2, we therefore use a Monte Carlo (MC) approach to make inference on non-linear functions of $\boldsymbol{Y}$ and to make predictions of data over the BAUs, which we assume to be from the same exponential family model as that of the original data.

For each quantity, **FRK** v2 uses the posterior expectation as the predictor; for uncertainty quantification, it provides the root-mean-squared prediction error as well as the widths of the prediction intervals.

2.4.1 Arbitrary prediction regions

Often, one does not wish to predict over single BAUs but over regions spanning multiple BAUs. These regions may overlap and may not coincide with entire BAUs. We define the set of prediction regions in terms of BAUs as $D^P \equiv \{\tilde{B}_l : l = 1, \dots, N_P\}$, where $\tilde{B}_l \equiv \cup_{i \in \tilde{c}_l} A_i$ and $\tilde{c}_l$ denotes the indices of the BAUs associated with the lth prediction region, for $l = 1, \dots, N_P$.

To predict $\boldsymbol{\mu}_P \equiv (\mu_{P,1}, \dots, \mu_{P,N_P})^\top$ over D^P, we aggregate the mean process $\boldsymbol{\mu}$ over the associated BAUs. Since each $\tilde{B}_l$ is a BAU or a union of BAUs, one can construct an $N_P \times N$ matrix

$$\boldsymbol{C}_P \equiv \left(\tilde{w}_{il} \mathbb{I}(i \in \tilde{c}_l) : i = 1, \dots, N; l = 1, \dots, N_P\right), \tag{12}$$

such that

$$\boldsymbol{\mu}_P = \boldsymbol{C}_P \boldsymbol{\mu}. \tag{13}$$

Inference on (13) and predictions of data over aggregations of BAUs is done using MC sampling.

2.5 Spatio-temporal framework

FRK accommodates spatio-temporal data by using spatio-temporal basis functions constructed via a tensor product of spatial and temporal basis functions. For computational reasons, $\text{cov}(\boldsymbol{\eta}, \boldsymbol{\eta})$ is modeled as being separable in space and time, although $\text{cov}(\boldsymbol{Y}, \boldsymbol{Y}) = \boldsymbol{S}\text{cov}(\boldsymbol{\eta}, \boldsymbol{\eta})\boldsymbol{S}^\top + \sigma^2_\xi \boldsymbol{V}$ is non-separable due to the presence of spatio-temporal basis functions in $\boldsymbol{S}$. For the random coefficients associated with the temporal basis functions, **FRK** uses a first-order autoregressive model.

2.6 High-resolution spatial and spatio-temporal modeling

The efficiency of **TMB** and our use of a sparse precision matrix when modeling $\text{cov}(\boldsymbol{\eta}, \boldsymbol{\eta})$ means that **FRK** v2 is now better equipped than **FRK** v1 to use a large number of basis functions. Specifically, **FRK** v2 is equipped to use an order of magnitude more basis functions ($> 10,000$) than **FRK** v1. This is important, as the predictive performance of fixed rank kriging is often determined by the number of basis functions.

3 Illustrative examples

We now demonstrate key new features in **FRK** v2. We present several illustrative examples chosen to show the utility of **FRK** v2 in doing complicated data analyses with little effort. In Section 3.1, we use data on poverty figures in Sydney, Australia, to demonstrate the spatial change-of-support functionality of **FRK** v2 in a non-Gaussian setting. In Section 3.2, we provide a non-Gaussian spatio-temporal example through modeling crime counts in the city of Chicago during the first two decades of the 21st century. In Section 3.3, we show the potential improvement in predictive performance of **FRK** v2 over **FRK** v1 when the data are Gaussian, owing to an increase in the maximum number of basis functions allowed in **FRK** v2. All results presented in the remainder of this paper can be generated using the reproducible code at https://github.com/msainsburydale/FRKv2_src.

3.1 Non-Gaussian spatial change-of-support (statistical downscaling)

The Australian Statistical Geography Standard (ASGS) defines a series of nested geographical areas in Australia known as Statistical Area Levels. Statistical Area Level 3 (SA3) regions are aggregations of Statistical Area Level 2 (SA2) regions, and SA2 regions are aggregations of Statistical Area Level 1 (SA1) regions. In this example, we consider a region around Sydney, and we aim to infer 'poverty' levels at the SA1 and SA3 regions from a data set containing mostly SA2 data and a small amount of SA1 data.

The data were collected in the Australian Census of 2011, and they consist of the number of families of various types within a range of weekly income brackets; we base our definitions of poverty lines on a Melbourne Institute of Applied Economic and Social Research (MIAESR) report that was published in March 2011 (Melbourne Institute of Applied Economic and Social Research, 2011). Data at the SA1 regions are available, and we use these to validate our predictions.

The first step when using **FRK** v2 is to construct and fit an '`SRE`' (short for spatial random effects) object using `FRK()`. In this example, we use a binomial data model with a logit link function. We use the SA1 regions stored as a '`SpatialPolygonsDataFrame`' object from the package **sp** (Pebesma and Bivand, 2005) as the BAUs, which also store the BAU-level 'number-of-trials' parameters (here the total number of families in each SA1 region). Since we have not explicitly provided basis functions, they are automatically constructed by `FRK()`. As BAUs, we use the SA1 regions. We note that `FRK()` automatically constructs BAUs when these are not pre-specified. We use the SA2-region (and some SA1-region) data for fitting the model and store them in a '`SpatialPolygonsDataFrame`' called `SA2s_and_some_SA1s`; see Figure 1.

```
R> S <- FRK(f = Z ~ 1,
+    data = SA2s_and_some_SA1s, BAUs = SA1s,
+    response = "binomial", link = "logit")
```

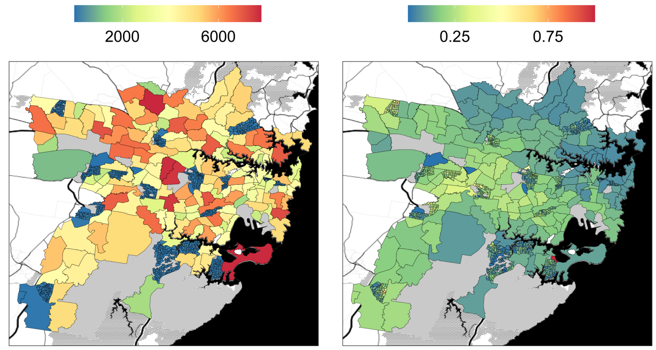

Figure 1: Data on SA1/SA2 regions used for modeling the proportion of families 'in poverty'. (Left) The total number of families in each region. (Right) The observed proportion of families in poverty in each region. Solid-grey regions correspond to SA1/SA2 regions in which the total number of families is zero. The data are overlayed on a Stamen base map, where the textured grey areas correspond to bushland, large black areas correspond to ocean or large bodies of water, and solid-black lines correspond to major arterial roads. Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.

Spatial prediction over all of the SA1 regions is obtained as follows.

```
R> SA1_prediction <- predict(S)
```

Predicting over different spatial supports is straightforward with **FRK** v2. We predict over the SA3 regions by passing them as a '`SpatialPolygonsDataFrame`' object to the argument `newdata`.

```
R> SA3_prediction <- predict(S, newdata = SA3s)
```

The function `predict()` returns predictions for several quantities of interest; here, we focus on the proportion of families in poverty in each region. The predictions and associated uncertainty over the SA1 and SA3 regions are shown in Figure 2; this figure was generated using a method of `plot()` exported by **FRK** v2.

We assessed the model's ability to quantify uncertainty over the SA1 regions by computing the empirical coverage from nominal 90% prediction intervals obtained via MC-simulated data at the SA1 level. We found the empirical coverage to be 90.8%, which is very close to the nominal value.

3.2 Non-Gaussian spatio-temporal data

The city of Chicago is divided into 77 so-called community areas (CAs). In this study, we model the number

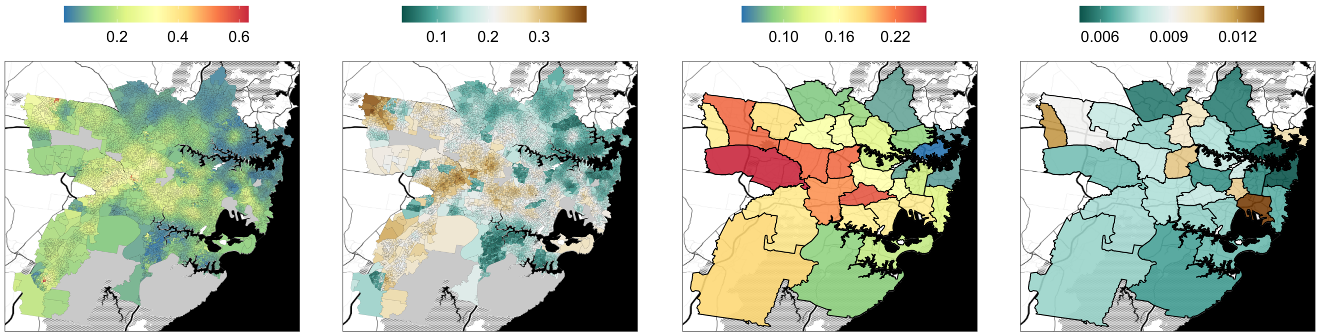

Figure 2: Prediction (left) and prediction-interval width (centre-left) for the proportion of families in poverty in each SA1 region. Prediction (centre-right) and prediction-interval width (right) for the proportion of families in poverty in each SA3 region.

of crimes in each CA between the years 2001 and 2019 inclusive. We considered only crimes labeled as 'assault' or 'battery'; there were roughly 1.75 million such crimes in total between 2001 and 2019. We excluded the last year, 2019, from the data set when fitting and used it to evaluate crime forecasts.

We initialize and fit the '`SRE`' object using `FRK()`. We use a Poisson data model with a log link function and, as spatial BAUs, we use the CAs stored as a '`SpatialPolygonsDataFrame`' object. Internally, spatio-temporal BAUs are created via the tensor product of the spatial BAUs and temporal yearly intervals. The data is stored as a '`STIDF`' object from the package **spacetime** (Pebesma, 2012).

```
R> S <- FRK(
+       f = number_of_crimes ~ 1,
+       data = chicago_crimes,
+       response = "poisson", link = "log",
+       spatial_BAUs  = community_areas,
+       sum_variables = "number_of_crimes"
+       )
```

We predict over the spatio-temporal BAUs using `predict()`.

```
R> pred <- predict(S)
```

The observed (withheld) number of crimes, forecasted number of crimes, and forecast uncertainty are shown in Figure 3. Figure 3 shows agreement between the forecasted and observed number of crimes. Further, the forecast uncertainty is roughly proportional to the forecasted value, as expected when counts are modeled.

3.3 Increased number of basis functions

As discussed in Section 2.6, **FRK** v2 is now better equipped than **FRK** v1 to use a large number of basis

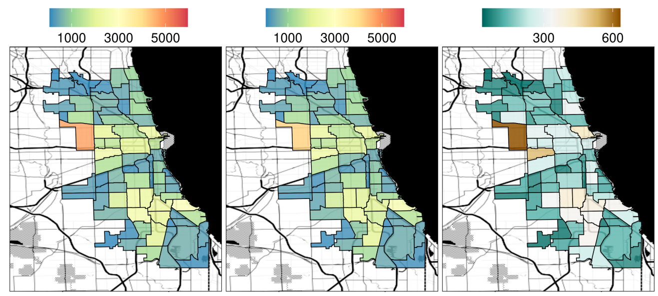

Figure 3: Observed (withheld) number of crimes (left), forecasts (centre), and 90% forecast-interval widths (right) over Chicago in the forecast year, 2019. The large black region in this figure corresponds to Lake Michigan.

functions, and this has implications on predictive performances. To demonstrate this, we re-ran the analysis for the comparative study of Gaussian spatial-prediction methods published in Heaton *et al.* (2019), which included spatial predictions from **FRK** v1. A summary of the diagnostic scores are shown in Table 1, and predictions from **FRK** v1 and **FRK** v2 are shown in Figure 4. These results show that the increased number of basis functions used by **FRK** v2 significantly improves the diagnostic scores compared with **FRK** v1, so that the results for **FRK** v2 in the Gaussian case are now comparable with other state-of-the-art prediction software. To achieve these improvements over **FRK** v1, we only had to specify `nres = 4` rather than `nres = 3`; the rest of the **FRK** code that was used in the comparative study was left unchanged.

4 Conclusion

We have described the spatio-temporal-inference package **FRK** v2, which is a major upgrade to **FRK** v1. Sub-

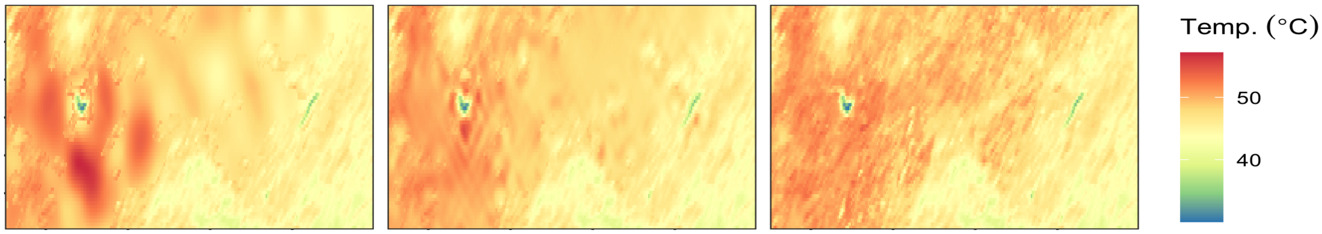

Figure 4: Predictions for **FRK** v1 (left) and **FRK** v2 (centre), and the truth (right) for a subset of the data used in the comparison study of Heaton *et al.* (2019). The increased number of basis functions used by **FRK** v2 leads to significantly less smoothing of the predicted spatial process when compared to the predictions of **FRK** v1.

Method	MAE	RMSPE	CRPS
FRK v1	1.96	2.44	1.44
FRK v2	1.37	1.81	0.98
Average	1.48	1.95	1.10

Table 1: A summary of scores in the comparative study presented in Heaton *et al.* (2019). The average is calculated with respect to all 12 models in the study, other than **FRK**. The scores are the mean absolute error (MAE), the root-mean-squared prediction error (RMSPE), and the continuous ranked probability score (CRPS).

stantial enhancements allow for the spatial and spatio-temporal modeling of, and large-scale prediction from, big, non-Gaussian data sets. The current version now provides a highly accessible and user-friendly approach to spatial and spatio-temporal modeling of big data in both a Gaussian and non-Gaussian setting, and it may prove to be a valuable tool in a variety of disciplines, from official statistics, to ecology, to geophysics.

References

Diggle PJ, Tawn JA, Moyeed RA (1998). "Model-Based Geostatistics." *Journal of the Royal Statistical Society C*, **47**, 299–350.

Finley AO, Banerjee S, Gelfand AE (2015). "**spBayes** for Large Univariate and Multivariate Point-Referenced Spatio-Temporal Data Models." *Journal of Statistical Software*, **63**(13), 1–28.

Finley AO, Datta A, Banerjee S (2020). "**spNNGP** R Package for Nearest Neighbour Gaussian Process Models." *arXiv:2001.09111*.

Heaton MJ, Datta A, Finley AO, Furrer R, Guinness J, Guhaniyogi R, Gerber F, Gramacy RB, Hammerling D, Katzfuss M, Lindgren F, Nychka DW, Sun F, Zammit-Mangion A (2019). "A Case Study Competition Among Methods for Analyzing Large Spatial Data." *Journal of Agricultural, Biological and Environmental Statistics*, **24**, 398–425.

Hughes J (2014). "**ngspatial**: A Package for Fitting the Centered Autologistic and Sparse Spatial Generalized Linear Mixed Models for Areal Data." *The R Journal*, **6**, 81–95.

Kristensen K, Nielsen A, Berg CW, Skaug H, Bell BM (2016). "**TMB**: Automatic differentiation and Laplace approximation." *Journal of Statistical Software*, **70**(5), 1–21.

McCullagh P, Nelder JA (1989). *Generalized Linear Models*. Chapman & Hall, London, UK.

Melbourne Institute of Applied Economic and Social Research (2011). "Poverty Lines: Australia, March Quarter 2011." *Technical report*.

Papritz A (2020). ***georob**: Robust Geostatistical Analysis of Spatial Data*. R package version 0.3-13, URL https://CRAN.R-project.org/package=georob.

Pebesma E (2012). "spacetime: Spatio-Temporal Data in R." *Journal of Statistical Software*, **51**(7), 1–30.

Pebesma EJ, Bivand RS (2005). "Classes and Methods for Spatial Data in R." *R News*, **5**, 9–13.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Tierney L, Kadane JB (1986). "Accurate approximations for posterior moments and marginal densities." *Journal of the American Statistical Association*, **81**, 82–86.

Wikle CK, Zammit-Mangion A, Cressie N (2019). *Spatio-Temporal Statistics with R*. Chapman & Hall/CRC Press, Boca Raton, FL.

Wood S (2017). *Generalized Additive Models: An Introduction with R*. Second edition. Chapman and Hall/CRC, Boca Raton, FL.