Image Forecasting Using Dynamic Functional Time-Series Models

Julian Austin

February 1, 2021

1 Introduction

Recent advances in satellite technology have enabled the collection of high spatial resolution satellite imagery over the same scene with high frequency. Such remote sensing data sets offer a rich source of information which can be utilised for various reasons. Take for example the Sentinel satellite constellation series described succinctly in Aschbacher and Milagro-Pérez (2012) the uses of such remote sensing data cover ocean, land and ice monitoring. We refer the reader to Malenovský et al. (2012) for a detailed description of the various studies available utilising Sentinel data.

Often a primary question of interest when using such data sets is how the imagery varies over time. An understanding of this would allow future variations to be quantified and the resultant forecasts to be fed back to decision makers as an additional source of information. As such producing forecast imagery from a time series of remote sensing images could be of great use. In the following, we propose a framework based on functional data analysis of time series of remote sensing imagery which describes the variation of the images

over time. In addition we utilise a dynamical functional time series model to produce forecast of the images.

Functional data analysis considers the modelling of data where observations are considered as functions over some continuous domain; for examples surfaces over a two dimensional space. The monographs of Ramsay and Silverman (2010) and Wang et al. (2016) both describe the details of analysing such functional data. One popular method in the functional data analysis literature for investigating modes of variation is Functional Principal Components Analysis (FPCA), see Ramsay and Silverman (2010) and Yao et al. (2005). This is closely linked to the multivariate Principal Components Analysis (PCA), Jolliffe (2002), and is often though of as the functional extension to it. Another more recent technique in dealing with functional variation is the Maximal Autocorrelation Factor Rotation (MAFR) introduced by Hooker and Roberts (2016). Such a technique is inspired by the multivariate Maximal Autocorrelation Factor (MAF) rotation of PCA, Switzer and Green (1984). The MAFR techniques aims to enhance the FPCA decomposition by finding a rotation of the principal components which promotes smoothness in the leading components. In the following, we consider utilising the decomposition methodologies of FPCA and MAFR to study the variation of a time series of remotely sensed imagery. We compare and contrast the ability of both techniques under different noise scenarios.

Finally we consider using such decompositions for forecasting imagery through time. A method which is closely linked to FPCA has been proposed by Shang (2013b) for use on univariate functional data and is commonly referred to a Functional Time Series Analysis (FTSA). Such a method proposed forecasting the functional data by first decomposing into its functional decomposition and forecasting the score functions of that resultant decomposition. In doing so we reduce the complexity of forecasting highly dependent observations to forecast-

ing univariate time series which can use a vast array of common time series approaches. We consider the ability of using FTSA approach with our imagery presented as bivariate functional data. In addition we propose the use of the MAFR decomposition under a similar framework as FTSA for forecasting. We compare the ability of the FTSA forecasting using both the standard FPCA and MAFR decomposition for various step ahead forecasts under different noise scenarios.

We illustrate the ability of these methodologies on both synthetically generated data as well as real world interferometric synthetic aperture radar data set.

The remainder of the article is structured as follows. In Section 2 we set out the functional representation of a time series of remotely sensed imagery and the FPCA and MAFR decomposition techniques. In Section 3 we set out the forecasting framework utilising the FTSA model with the FPCA and MAFR decompositions. We consider the effectiveness of the models using a simulated data set in Section 4 and using a real world interferometric synthetic aperture radar data set in Section 5. Finally we draw conclusions of the article and propose future direction in Section 6.

2 Functional Data Analysis Methodology

In a time series of remotely sensed images each observed pixel of the image can be indexed by three dimensions; one temporal and two spatial. The most usual presentation of functional data analysis is to consider discrete observations as samples from univariate functions indexed by a spatial location. For example see Liu et al. (2017) on their treatment of spatially correlated functional data. However in our case we will consider our discrete observations as samples from bivariate functions or surfaces indexed by a temporal location.

2.1 Functional Representation

In the functional data analysis approach we consider our observed images as discrete observations from a realisation of an underlying stochastic process, $\mathcal{X}(\cdot)$, that is square integrable. Our observed sample of a single image consists of N individual pixel values with locations denoted by $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N$ where each location lies in some domain $\mathcal{S} \subset \mathbb{R}^2$ representing our whole domain of the image. We assume our pixel values are observed with some error. Our data set then consists of J images each of which is sampled as described giving our observation model as:

$$y_i\left(\boldsymbol{s}_{ij}\right) = \chi_i\left(\boldsymbol{s}_{ij}\right) + \varepsilon_i\left(\boldsymbol{s}_{ij}\right) \tag{1}$$

where $y_i(\boldsymbol{s}_{ij})$ represents our observed pixel value for the i^{th} image at spatial location $\boldsymbol{s}_{ij}$ for $i = 1, 2, \ldots, J$ and $j = 1, 2, \ldots, N$. $\chi_i(\cdot)$ represents the i^{th} realisation of the stochastic process $\mathcal{X}$ which is our noise free image at time i. $\varepsilon_i(\cdot)$ represent the i^{th} realisation of a noise process, which we will assume is independent across time but not necessarily independent in space.

In order to estimate our noise free image $\chi_i(\cdot)$ from our observed data $\{y_i(\boldsymbol{s}_{ij}); j = 1, 2, \ldots, N\}$ we opt to utilise a smoothing methodology which is popular in the functional data literature known as penalised least square smoothing using a basis expansion, Ramsay and Silverman (2010). In order to use such a methodology we assume that our process $\mathcal{X}$ can be represented in a known bivariate basis system. That is:

$$\mathcal{X}\left(\boldsymbol{s}\right) = \sum_{\substack{1 \le k_1 \le K_1 \\ 1 \le k_2 \le K_2}} \theta_{k_1 k_2} \boldsymbol{\phi}^1_{k_1}\left(s_1\right) \boldsymbol{\phi}^2_{k_2}\left(s_2\right) \tag{2}$$

where $\boldsymbol{\phi}^1(s) = \left(\phi^1_1(s), \phi^1_2(s), \ldots, \phi^1_{K_1}(s)\right)$ is a known univariate basis system

over the first spatial dimension. Similarly, $\boldsymbol{\phi}^2(s)$ is a univariate basis system over the second spatial dimension. The θ_{kl} are unknown random coefficients to be determined and $\boldsymbol{s} = (s_1, s_2)^\top$ is the vector of spatial coordinates. In our work we consider the known basis system to be the B-spline basis system, see Piegl and Tiller (1997) for a detailed description of such a basis system. We can write such a basis expansion more succinctly by using the tensor product notation. Let $\bar{\boldsymbol{\phi}}(\boldsymbol{s}) = \boldsymbol{\phi}^2(s_2) \otimes \boldsymbol{\phi}^1(s_1)$ where $\otimes$ represents the Kronecker product. Let $\boldsymbol{\theta} = \text{Vec}(\Theta)$ where Vec is an operator which stacks columns of a matrix and $\Theta \in \mathbb{R}^{K_1 \times K_2}$ be the matrix formed of elements $\theta_{k_1 k_2}$. The Equation 2 can be written as:

$$\mathcal{X}(\boldsymbol{s}) = \bar{\boldsymbol{\phi}}^\top(\boldsymbol{s})\,\boldsymbol{\theta} \tag{3}$$

Then for the i^{th} realisation of $\mathcal{X}$ we can denote the coefficients to be determined by $\boldsymbol{\theta}^i$ from the observations $\boldsymbol{Y}_i = \{y_i(\boldsymbol{s}_{ij}); j = 1, 2, \ldots, N\}$. We employ penalised least squares to estimate such coefficients using Equation 3. That is our estimated coefficients are given by, Ramsay and Silverman (2010):

$$\hat{\boldsymbol{\theta}}^i = \left(\bar{\boldsymbol{\phi}}^\top \boldsymbol{W} \bar{\boldsymbol{\phi}} + \boldsymbol{P}(\boldsymbol{\lambda})\right)^{-1} \bar{\boldsymbol{\phi}}^\top \boldsymbol{W} \boldsymbol{Y}_i \tag{4}$$

where $\boldsymbol{W}$ is a known weighting matrix, $\boldsymbol{P}(\boldsymbol{\lambda})$ is a penalty matrix whose size is controlled by regularisation parameter $\boldsymbol{\lambda}$ to control over fitting. In particular we consider the use of the following form of the penalty matrix for a two dimensional B-spline basis system proposed by Wood (2006) given by:

$$\boldsymbol{P}(\boldsymbol{\lambda}) = \lambda_1 \boldsymbol{P}^1 \otimes \boldsymbol{I}_{K_2} + \lambda_2 \boldsymbol{I}_{K_1} \otimes \boldsymbol{P}^2 \tag{5}$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)^\top$ is our two dimensional regularisation parameter controlling the regularisation across each dimension. $\boldsymbol{P}^1 \in \mathbb{R}^{K_1 \times K_2}, \boldsymbol{P}^2 \in \mathbb{R}^{K_2 \times K_2}$ are marginal second order penalty matrices for each spatial dimension respectively.

$\boldsymbol{I}_{K_1}, \boldsymbol{I}_{K_2}$ are identity matrix of order K_1 and K_2 respectively. The form of the one dimensional penalty matrices are fairly common and are formed through elements of:

$$P_{l,m} = \int \boldsymbol{\phi}_l''(s)\boldsymbol{\phi}_k'' ds \tag{6}$$

where P_{lm} is the $(l,m)^{\text{th}}$ element of matrix $\boldsymbol{P}$ and we replace $\boldsymbol{\phi}$ by the appropriate basis system for each dimension to give $\boldsymbol{P}^1$ and $\boldsymbol{P}^2$. We display the second order penalty we have chosen to use since it will penalise high curvature surfaces along each dimension. Our smoothed estimate for the i^{th} realisation , $\chi_i(\boldsymbol{s})$, of $\mathcal{X}$ is then given by:

$$\hat{\chi}_i(\boldsymbol{s}) = \bar{\boldsymbol{\phi}}^{\top}(\boldsymbol{s})\,\hat{\boldsymbol{\theta}}^i \tag{7}$$

With a functional representation of our discretely observed data from each image over time we can examine the variation between image functions that we observe. We consider the FPCA and MAFR methodology for such.

2.2 Functional Principal Component Analysis

First developed as a theory for the optimal expansion of a continuous stochastic process proposed in Karhunen (1946) and Loève (1946). This theory was then to become known as the Karhunen-Loéve expansion and was applied to functional data in early works such as Tucker (1958). For a summary of recent advances to the FPCA methodology see Shang (2013a).

We state the basis properties of FPCA without proof; for more details see Ramsay and Silverman (2010). Alike in the multivariate technique PCA, FPCA seeks to find components which decreasingly capture variation in our observed functions.

Suppose as described in Section 2.1 we have J surfaces $\chi_i(\boldsymbol{s})$ for $i = 1, 2, \ldots, J$

and alike PCA we wish to find modes of maximal variation. That is initially we wish to find an eigenfunction which depict the dominant mode of variation. Given the Karhunen-Loève expansion of $\mathcal{X}$ by:

$$\mathcal{X}(\boldsymbol{s}) - \mu(\boldsymbol{s}) = \sum_{k=1}^{\infty} \zeta_k \psi_k(\boldsymbol{s}) \tag{8}$$

where $\mu(\boldsymbol{s}) = E\left(\mathcal{X}\left(\boldsymbol{s}\right)\right)$ and $\psi_1, \psi_2, \ldots$ are the orthonormal eigenfunctions of the linear Hilbert-Schmidt operator induced by $G\left(\boldsymbol{s}, \boldsymbol{s}'\right) = \mathrm{Cov}\left(\mathcal{X}\left(\boldsymbol{s}\right), \mathcal{X}\left(\boldsymbol{s}'\right)\right)$. ζ_k is the principal component associated with the k^{th} eigenfunction ψ_k and defined by:

$$\zeta_k = \int_{\mathcal{S}} \left(\mathcal{X}\left(\boldsymbol{s}\right) - \mu\left(\boldsymbol{s}\right)\right) \psi_k\left(\boldsymbol{s}\right) d\boldsymbol{s} \tag{9}$$

If we assume the eigenfunctions are ordered such that the corresponding eigenvalues are ordered as $\omega_1 \geq \omega_2, \ldots$. Then it can be shown that the first eigenfunction ψ_1 depicts the dominant mode of variation, that is:

$$\psi_1 = \underset{\|\psi\|=1}{\arg\max} \left(\mathrm{Var} \left(\int_{\mathcal{S}} \left(\mathcal{X}\left(\boldsymbol{s}\right) - \mu\left(\boldsymbol{s}\right)\right) \psi\left(\boldsymbol{s}\right) \right) d\boldsymbol{s} \right) \tag{10}$$

where $\|\cdot\|$ is the L^2 norm. The k^{th} eigenfunction then similarly corresponds to the k^{th} dominant mode of variation subject to being orthogonal to the previous $k-1$ eigenfunctions.

We use the Principal Component Analysis Though Conditional Expectation (PACE) methodology introduced in Yao et al. (2005) to estimate the above model components including error variance (σ_ε^2), mean function (μ), eigenfunctions ($\psi_k,\ k = 1, 2, \ldots, K$), and scores ($\zeta_i, i = 1, 2, \ldots, J$). The details of the estimation methodology can be found in Yao et al. (2005). In addition we can use our basis expansion representation as discussed in Section 2.1 to simplify the calculation of these estimates using the methodology discussed in Ramsay and Silverman (2010) to simplify the estimate for $G\left(\boldsymbol{s}, \boldsymbol{s}'\right)$.

2.3 Maximal Autocorrelation Factor Rotation

The FPCA methodology utilises linear combinations of observed functions to find transformations that maximise the variance of the projected scores. However maximising variance may not be the optimal criterion to priorities components in our decomposition. For example we may prefer components which are more interpretable. One technique in the multivariate literature is to consider re-expressing components of the PCA which emphasis smoothness through a factor rotation. Hooker and Roberts (2016) introduces such a factor rotation for the FPCA methodology which re expresses the functional subspace formed by the components of the FPCA decomposition in terms of directions of decreasing smoothness as represented by some smoothing metric. In the following we briefly discuss the components of such a rotation and refer the reader to Hooker and Roberts (2016) for more details.

The methodology start by assuming we have already performed the FPCA decomposition and we retain the leading K components writing:

$$\boldsymbol{\psi}(\boldsymbol{s}) = (\psi_1(\boldsymbol{s}), \psi_2(\boldsymbol{s}, \ldots, \psi_K(\boldsymbol{s}))^\top$$
$$\boldsymbol{\zeta}_i = (\zeta_{i1}, \zeta_{i2}, \ldots, \zeta_{iK})^\top \text{ for } i = 1, 2, \ldots, J$$

for the retained eigenfunctions and the corresponding score vectors for the J observed surfaces.

The smoothness we wish to promote in our component eigenfunctions is specified similarly to the roughness penalty as in Section 2.1. Let L be a linear differential operator which captures such a smoothness constraint and define $\boldsymbol{P}^{\text{mafr}}$ to be the matrix formed through:

$$\boldsymbol{P}^{\text{mafr}}_{kl} = \int_{\mathcal{S}} L\boldsymbol{\psi}_k(\boldsymbol{s})\, L\boldsymbol{\psi}_l(\boldsymbol{s})\, d\boldsymbol{s} \tag{11}$$

Then the MAFR rotation can be found by the Eigen decomposition of $\boldsymbol{P}^{\text{mafr}} \in \mathbb{R}^{K\times K}$ (see Hooker and Roberts (2016) for details). Write $\boldsymbol{P}^{\text{mafr}} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\top$ for the Eigen decomposition of $\boldsymbol{P}^{\text{mafr}}$ then the MAFR components correspond to:

$$\boldsymbol{\psi}_{\text{mafr}} = \boldsymbol{U}^\top \boldsymbol{\psi} \tag{12}$$

Similarly the MAFR scores corresponding to such rotated components can be found through $\boldsymbol{\zeta}_i^{\text{mafr}} = \boldsymbol{U}\boldsymbol{\zeta}_i$ for $i = 1, 2, \ldots, J$.

Such a methodology promotes smooth eigenfunctions in the sense of the constraint L whilst retaining the total variance explained from the decomposition under the FPCA methodology. The methodology is also inexpensive to compute as the additional work required from the FPCA methodology is the Eigen decomposition of a $K \times K$ matrix where K is typically relatively small.

3 Forecasting Methodology

Forecasting of remotely sensed imagery is typically considered by forecasting each time series of individual pixels. Such a methodology is often complex due to the spatial dependency observed between neighbouring pixels which induces spatial dependency between the observed time series. By considering the data set as a collection of surfaces over space as we do in Section 2.1 we aim to simplify the forecasting methodology since the spatial dependency is already taken into account in our representation of the surface. We are thus left with a time series of functional variables.

Recent work by Shang (2013b) has considered this case for univariate functional data. They propose a frame work known as Functional Time Series Analysis (FTSA) using the FPCA decomposition. The reason for forecasting using the FPCA decomposition is that it allows the complex case of forecasting

functional variables to be reduced to forecasting univariate time series through the scores of the decomposition. We summarise the FTSA methodology in the following and refer the reader to Shang (2013b) and the references within for further details.

As before we assume we have an K component FPCA decomposition of our observed data $\boldsymbol{Y}$ as described in Section 2.2. That is we can recover an estimate of our smooth signal surface $\chi_i(\boldsymbol{s})$ by:

$$\hat{\chi}_i(\boldsymbol{s}) = \hat{\mu}(\boldsymbol{s}) + \sum_{k=1}^{K} \hat{\zeta}_{ik}\hat{\psi}_k(\boldsymbol{s}) + \hat{\epsilon}_i(\boldsymbol{s}) \tag{13}$$

where $\hat{\mu}$, $\hat{\zeta}_{ik}$ and $\hat{\psi}_k$ are the sample mean surface estimate, estimated scores, and estimated eigenfunctions from the FPCA decomposition respectively. The error term, $\hat{\epsilon}_i$ is error due to using the truncated series expansion. Shang (2013b) propose to utilise a univariate forecast of each score series, $\{\hat{\zeta}_{ik}\}_{i=1}^{J}$, which is then used to forecast the full series $\{\chi_i(\boldsymbol{s})\}_{i=1}^{J}$.

By conditioning on the set of smoothed surfaces

$$\hat{\boldsymbol{\chi}}(\boldsymbol{s}) = (\hat{\chi}_1(\boldsymbol{s}), \hat{\chi}_2(\boldsymbol{s}), \ldots, \hat{\chi}_J(\boldsymbol{s}))^{\top}$$

and the fixed principal components $\hat{\boldsymbol{\psi}} = \left(\hat{\psi}_1(\boldsymbol{s}), \hat{\psi}_2(\boldsymbol{s}), \ldots, \hat{\psi}_K(\boldsymbol{s})\right)^{\top}$, the h-step ahead forecasts $y_J(\boldsymbol{s})$ are given by:

$$\hat{y}_{J+h|J}(\boldsymbol{s}) = \hat{\mu}(\boldsymbol{s}) + \sum_{k=1}^{K} \hat{\zeta}_{J+h|J,k}\psi_k(\boldsymbol{s}) \tag{14}$$

where $\hat{\zeta}_{J+h|J,k}$ is the h-step ahead forecast of the univariate score series for the k^{th} component. The exact method of forecasting the univariate score series is not prescribed but there exists many univariate time series forecasting methodologies that can be chose, see Hyndman and Athanasopoulos (2018) for

a variety of examples. Since under FPCA methodology we have independent scores across components we can perform K univariate forecasts to obtain all the forecasts we wish.

Similarly we propose using the FTSA methodology with the MAFR decomposition as described in Section 2.3. The exact same methodology can be employed but using the MAFR scores and eigenfunctions. It is worthwhile to note that we no longer have independent score series as under the FPCA decomposition since they are correlated with the MAFR rotation $\boldsymbol{U}$ however practically we can still employ univariate time series method for each MAFR score series independently.

4 Simulated Experiment

This simulated study comprises of a data set of surface displacement measurements from a earthquake model simulation. The data comprises of 128 image over time which are equally spaced apart from prior to the onset of earthquake to its end. Each image comprises of 512×512 pixels with a resolution of $90m$ in both directions. We take such a simulated data set as our ground truth observations and add known measurement error processes $\{\varepsilon_i(\boldsymbol{s})\}_{i=1}^{128}$ to see how well the forecasting methodologies discussed in Section 3 perform under the various errors.

4.1 Experimental Design

Our simulated experiment is setup to compare and contrast the two forecasting methodologies using the FPCA and MAFR decompositions. To do so effectively we must keep our representation of the surfaces fixed in bot cases, as such we specify a tensor product B-spline basis with 32 basis functions in each dimension as our basis expansion for the representation discussed in Section 2.1. Such a

basis is chosen as it is flexible to accommodate variation among images and across space whilst maintaining computational feasibility. In order to reduce computation time for such methodology we tile our image in 128×128 sections and apply the decomposition and forecasting techniques on each tile of the full image independently. We can then recover the full image by stitching together our forecasted tiles respectively. The regularisation parameter, $\boldsymbol{\lambda}$, we choose through Generalised Cross Validation (see Lukas (2006) for details).

For both decomposition techniques we choose five components, that is $K = 5$ in Equation 13. The final design choice is the univariate forecasting methodology to use for the score processes. We choose to model such univariate series as a Gaussian Process Regression, Williams and Rasmussen (2006). That is:

$$\zeta_{ik} = \zeta_k(t_i) \sim \mathcal{GP}(\mathbf{0}, A_k(t, t')) \tag{15}$$

where $A_k(t, t')$ is covariance kernel for component k which we choose to be the Matèrn covariance, Abramowitz and Stegun (2013). We choose to fix the shape parameter to be 1.5 for the Matèrn covariance but leave the length scales and variance parameters to be chosen through empirical maximum likelihood estimation (See Williams and Rasmussen (2006) for implementation details). We allow the Matèrn covariance to be different for each component.

We employ three metrics for evaluation. We use the standard root mean square error (RMSE), mean absolute error (MAE), and the structured similarity index measure (SSIM), Zhou Wang et al. (2004). We employ these metrics on the our ground truth images with our h-step ahead forecast for $h = 1, 3, 10, 25$. To do so we utilise 90 time steps as training images, the remaining 35 images are left as test imagery under the h-step ahead time series standard procedure, Hyndman and Athanasopoulos (2018).

Finally, we specify three types of measurement error processes to compare

against. The first process we consider is a low variance white noise process with $\sigma_\varepsilon^2 = 10.0$, the second is a high variance white noise process with $\sigma_\varepsilon^2 = 20.0$, and lastly we consider a spatially structured noise process with high variance. The spatial structure is again specified through a Gaussian Process with Matèrn covariance. The covariance has shape parameter 1.5 with variance 20 and isotropic length scale of 10.

4.2 Simulation Results

5 Real world Experiment

6 Conclusion

References

Abramowitz, M. and Stegun, I. A. (eds.) (2013) *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. No. 1 in Dover books on mathematics. New York, NY: Dover Publ.

Aschbacher, J. and Milagro-Pérez, M. P. (2012) The European Earth monitoring (GMES) programme: Status and perspectives. *Remote Sensing of Environment*, **120**, 3–8. URL: http://www.sciencedirect.com/science/article/pii/S0034425712000612.

Hooker, G. and Roberts, S. (2016) Maximal autocorrelation functions in functional data analysis. *Statistics and Computing*, **26**, 945–950. URL: http://link.springer.com/10.1007/s11222-015-9582-5.

Hyndman, R. J. and Athanasopoulos, G. (2018) *Forecasting: principles and practice*. OTexts. Google-Books-ID: _bBhDwAAQBAJ.

Jolliffe, I. T. (2002) *Principal component analysis*. Springer series in statistics. New York: Springer, 2nd ed edn.

Karhunen, K. (1946) Zur Spektraltheorie stochastischer Prozesse. *Ann. Acad. Sci. Finnicae, Ser. A*, **1**, 34.

Liu, C., Ray, S. and Hooker, G. (2017) Functional principal component analysis of spatially correlated data. *Statistics and Computing*, **27**, 1639–1654. URL: https://doi.org/10.1007/s11222-016-9708-4.

Loève, M. (1946) Fonctions aléatoires à décomposition orthogonale exponentielle. *La Revue Scientifique*, **84**, 159–162.

Lukas, M. A. (2006) Robust generalized cross-validation for choosing the regularization parameter. *Inverse Problems*, **22**, 1883–1902. URL: https://doi.org/10.1088/0266-5611/22/5/021. Publisher: IOP Publishing.

Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M. E., García-Santos, G., Fernandes, R. and Berger, M. (2012) Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment*, **120**, 91–101. URL: http://www.sciencedirect.com/science/article/pii/S0034425712000648.

Piegl, L. A. and Tiller, W. (1997) *The NURBS book*. No. 1 in Monographs in visual communications. Berlin ; New York: Springer, 2nd ed edn.

Ramsay, J. O. and Silverman, B. W. (2010) *Functional data analysis*. New York (N.Y.): Springer Science+Business Media.

Shang, H. (2013a) A survey of functional principal component analysis. *AStA Advances in Statistical Analysis*, **98**, 121–142. URL: https://browzine.com/articles/39384090.

Shang, Lin, H. (2013b) ftsa: An R Package for Analyzing Functional Time Series. *The R Journal*, **5**, 64. URL: https://journal.r-project.org/archive/2013/RJ-2013-006/index.html.

Switzer, P. and Green, A. (1984) Min/Max autocorrelation factors for multivariate spatial imaging. *Technical Report No.6*, 14. Publisher: Department of Statistics, Stanford University, Stanford, CA.

Tucker, L. R. (1958) Determination of parameters of a functional relation by factor analysis. *Psychometrika*, **23**, 19–23. Publisher: Springer.

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016) Functional Data Analysis. *Annual Review of Statistics and Its Application*, **3**, 257–295. URL: https://doi.org/10.1146/annurev-statistics-041715-033624. _eprint: https://doi.org/10.1146/annurev-statistics-041715-033624.

Williams, C. K. and Rasmussen, C. E. (2006) *Gaussian processes for machine learning*, vol. 2. MIT press Cambridge, MA.

Wood, S. N. (2006) Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics*, **62**, 1025–1036.

Yao, F., Müller, H.-G. and Wang, J.-L. (2005) Functional Data Analysis for Sparse Longitudinal Data. *Journal of the American Statistical Association*, **100**, 577–590. URL: https://www.jstor.org/stable/27590579. Publisher: [American Statistical Association, Taylor & Francis, Ltd.].

Zhou Wang, Bovik, A., Sheikh, H. and Simoncelli, E. (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, **13**, 600–612.