Research Article

Spatiotemporal forecast with local temporal drift applied to weather patterns in Patagonia

Eduardo Henrique de Moraes Takafuji[1] · Marcelo Monteiro da Rocha[1] · Rodrigo Lilla Manzione[2]

Received: 8 February 2020 / Accepted: 22 April 2020

Abstract
Geostatistics was developed to generate maps or 3D models interpolating observed values in space. The so-called spatiotemporal geostatistic applies the same principles to estimate observed values that have both spatial and temporal distribution. Moreover, time series analysis can decompose and extrapolate its main trends and seasonality, preparing data for geostatistical assumptions. Using this principle, this study aims to decompose the time series of a spatiotemporal dataset as external drifts and estimate its residuals by spatiotemporal kriging. Since each observation point is a time series, it is possible to decompose its trend and seasonality locally and map its parameters, preferable, by traditional geostatistics. Aftermath, it is possible to extrapolate the trend and seasonality at each pixel. This procedure can achieve great long-term forecasting maps even in regions with poor sampling due to its time series analysis. As well as, the geostatistics guarantee that the spatio-temporal correlation is maintained. This method is especially good for prediction in regions that the time series pattern depends on its location, which is a common problem in large areas and the problem is worsened in poorly sampled regions. This study presents a 10 years map forecast (2008–2017) comparison by spatiotemporal geostatistics, the first with original data, with ARIMA Models Panels, then with global decomposition, finally, with the local decomposition approach. The target variable is temperature captured by the 18 active weather stations in Patagonia between 1973 and 2007. To validate the results, they are compared to Land Surface Temperature (LST), which is an image product MOD11C3 derived from the MODIS sensor onboard on Terra/Aqua satellites. The proposed method can make long-term forecasts with low error, low smoothing effect and similar spatiotemporal statistics (mean and variance) of the stations and the LST product. Finally, its results are comparable with the ARIMA Models Panels with the advantage that it can generate maps with spatiotemporal correlation and better than the often-used methods (st-kriging and global decomposition) to forecast large areas maps.

Keywords Time series analysis · Local decomposition · Space–time geostatistics · Map forecasting · Patagonia

1 Introduction

This study objective is a procedure that uses methods that are well-known, straightforward, statistics-based and data-driven to achieve robust long-term results in low information regions with high spatial variability of its time series (i.e. consider the local time series due to its patterns are different in mountain and beaches). The major contribution of this study is the procedure that can make long-term map prediction of a large and poorly sampled area (e.g. mapping the effects of climate change over an entire county for a long period such as a century) with low error and low smoothness.

✉ Eduardo Henrique de Moraes Takafuji, eduardo.takafuji@alumni.usp.br; Marcelo Monteiro da Rocha, mmrocha@usp.br; Rodrigo Lilla Manzione, lilla.manzione@unesp.br | [1]Institute of Geosciences, University of São Paulo (IGc-USP), Rua Do Lago, 562 Cidade Universitária, São Paulo, SP 05508-080, Brazil. [2]Department of Biosystems Engineering, School of Sciences and Engineering, São Paulo State University (FCE-UNESP), Rua Domingos da Costa Lopes, 780 Jd. Itaipu, Tupã, SP 17602-496, Brazil.

Published online: 02 May 2020

The climatological processes operate over a wide range of timescales and time series analysis supports the detection, description, and modeling of the climatic variability and impacts [1]. Weather forecasts are made by solving the equations of motion for the atmosphere, the nonlinear partial differential equations of dynamics, moisture conservation, thermodynamics, and mass continuity [2]. These equations try to keep track of the myriad of complex nonlinear interactions between winds, temperature, and moisture at thousands of locations and altitudes around the world [2]. Numeric modeling attempts to add a better understanding of the modus operandi of the Earth's climate system [3]. Climatology and meteorology study chaotic and non-linear signals that are unstable and nonperiodic to model all possible interaction between the natural events and the dependency to initial states as shown in Chaos Theory [4]. Slightly differing initial states can evolve into considerably different states [4]. Furthermore, the errors in different stages were analyzed and it is assumed that there would be no prediction error if we could observe an initial state without error [5]. Other studies in chaos theory applied to spatiotemporal variables are cited: reconstructing spatiotemporal dynamics using support vector machines [6]; nonlinear ensemble prediction of chaotic daily rainfall [7]; polynomial Chaos methods are used to quantify uncertainties in ocean forecasting [8]; fuzzy Bayesian network-based data-driven framework for spatiotemporal prediction with an elegant approach of dealing with intrinsic chaos in time series [9].

The global climate models (GCM) can make forecasts over a grid that covers the globe by numerically solving sets of finite-differences equations [2]. Because models are imperfect, and observations have inherent errors and inadequate coverage, neither model nor observations alone will provide a full, comprehensive description of the Earth's climate system [3]. Likewise, some physical processes are not known well enough to provide exact physical laws or, for other situations, the exact physical processes are so complicated or computationally unwieldy that, in these circumstances, a simpler parameterization can provide an answer that could be considered good enough [2]. A major barrier to improvement in climate model simulation is the insufficiency of detailed investigation for climate processes with high temporal and spatial resolutions [3]. Further, all models are simplifications of the real world and the climate models can be considered as a series of equations expressing physical, chemical, biological and social laws [10]. Moreover, the more complex the representation, the costlier the model is to use, and the result can only be approximations [10].

With all this information and the right parameters, GCMs can generate great predictions with a lot of details. On the other hand, setting all climate parameters can be truly demanding and its processing will request a supercomputer and long periods. The climate system is extremely complex and statistical analysis is essential in providing useful simplifications of the system variability. In order to generate long-term trend maps, the accuracy in the specific location and date is not as important as its local trends and statistics, so the use of geostatistics and time series is adequate. With the advantage of being statistics-based and data-driven models that do not demand excessive processing time in supercomputers.

Geostatistics is defined as the study of distribution in space of useful values for mining [11]. Its application was transferred to other areas such as hydrogeology, environmental studies, agriculture, climate, epidemiology and everything else presenting a correlated spatial distribution. After that, this theory started being applied in a spatiotemporal framework to estimate the random variables at unobserved points at any specific time. However, its basics concepts explained in [11] and deepen in [12] should be followed, like second-order stationarity in space and time. Temporal data collected over time usually present some seasonality. In other words, it is non-stationary. Qualitatively, a stationary time series presents statistical equilibrium, containing no trends and, most methods of handling non-stationary time series remove the non-stationary parameters [13].

Considering any spatially distributed dataset with values of a specific variable being recorded periodically, e.g. monitoring stations (weather stations, water level, pollution stations, and others), the dataset is spatiotemporal due to its temporal and spatial components. The spatiotemporal variables (e.g. temperature or pollution) $Z(s, t)$ of both space (s) and time (t) domains are usually registered at constant temporal intervals at established points [14]. Space represents a state of coexistence, where occur multiple dimensions (or directions) and it is frequently interpolated [15]. On the other hand, time is the state of successive existence with a defined order (nonreversible) in only one dimension and the extrapolation is commonly the main interest [15]. Considering one natural spatiotemporal variable Z(s, t) where s is the location in space and the t is an instant in time, its spatiotemporal framework tends to be cyclical in the temporal component. In other words, natural events usually present a seasonality, it can be short as day and night or long as a visible comet pass near Earth. However, it is harder to create a deterministic model of space and that is why one commonly estimated its values with geostatistics.

Based on the spatiotemporal observations, it is possible to reproduce the behavior of the spatiotemporal process, or simply predict its value at given space–time points. Geostatistics takes advantage of the spatio-temporal correlations present in the spatio-temporal data to make

predictions at unobserved space–time locations [16]. Spatiotemporal geostatistics can be applied to generate maps that represent the spatial distribution of a feature at any desired time, that is, it is possible to forecast entire regions with a few time-series sampling points. The combined space–time domain should be regarded only as a coordinate system, where observations are tagged by a spatial coordinate vector s and a temporal coordinate t [17]. Essentially, the definition of isotropy cannot be stated for spatio-temporal random fields, since space and time have to be considered as distinct and non-comparable entities [18]. Furthermore, units and scales are distinct between space and time, as well as, they cannot be compared in a physical sense. Similarly, it is shown that every spatiotemporal location is a point on $R^d \times R$ with R^d being the d-dimensional Euclidean space and R the time dimension [16].

The spatiotemporal geostatistical modeling requires second-order stationary dataset in space and time domains. Thereby, one should consider a procedure to access the stationary components of the time series. The presence of the drift does not allow it to be estimated directly from the variogram [12]. The first important step in all geostatistical modeling is to establish the correct variable to model and make sure that this property can be modeled as stationary over the domain of the study [19]. Moreover, if data show a systematic trend, this trend must be modeled and removed before the geostatistics and added back to the estimated values at the end [19]. Because trend makes the variable non-stationary and it is unreasonable to expect the mean value to be independent of location [19].

Thus, in order to obtain the second-order stationarity, one way is to remove its global trend and seasonal components of the variable. Some studies formulate a general space–time model with a purely spatial mean function, purely temporal mean function and the intrinsically stationary space–time error process and considering that this model is sufficiently flexible to model a large range of spatiotemporal data sets [20]. However, if there is seasonality, it can be detected and removed by inspecting each time series graphs and kriged the residuals to get the values in the future and at the monitoring location [21, 22]. More fully, it is possible to decompose the space–time considering the case where the dynamical structure is unknown and it can be written in terms of the deterministic means and four mutually statistically independent mean-zero random effects [23]: (i) location-specific variability common to all times; (ii) time-specific variability common to all locations; (iii) capturing the spatiotemporal interaction; (iv) representing the microscale spatiotemporal variability. As well as, decompositions approaches in the spatiotemporal framework [24], such as (i) the trend component is constant over the space–time domain; (ii) the trend component is described by a deterministic model in space and time; (iii) the trend component is described by a stochastic model in space and time. All interpolation and results were compared at the sample locations.

Other authors applied the global trend/seasonal decomposition to generate a map thought spatiotemporal geostatistics—where the global trend and/or seasonality components and the residual variogram model are used to predict future maps. Considering the sub-regions of the domain, it is possible to remove the spatiotemporal drift component from each observed sub-region in order to perform a residual geostatistical analysis [25]. Another approach is to model the trend applying the generalized linear model with a random component and a linear function of regressors with the long-term trend and/or periodic components as the large-scale variation and the small-scale residuals are the stationary residuals [26, 27]. Furthermore, to assess its residuals (remove the trend and the seasonality), some authors [28–30] used the differencing time series. Time series decomposition for climate dataset was also done by multifractal characterization [31], spectrum analysis for investigating chaos [32], pattern scaling decomposition [33]; modeling big data with MapReduce framework [34]; spatiotemporal prediction using dimension reduced local states [35]; ensemble empirical mode decomposition [36]; and, finally, a comparison of methods for extracting annual cycle with changing amplitude shows the importance of the non-linear mode decomposition [37].

The geostatistical methods found in the literature works well if the study is focused on a specific location or considering sub-regions. However, if the objective is to generate a map of a long-term forecast in a large area (where the time series depends on its location), one should consider the use of the local temporal drift. The local temporal drift decomposes (removing the trend and seasonality) each sampling point (time series) and consider its parameters as regionalized variables. These new regionalized variables can be mapped with geostatistical or non-geostatistical interpolators. Then, it is possible to generate an equation of time trend and oscillation at all studied locations (grid points) with every parameter of the time series at each pixel in the grid. Then one can estimate every time series parameter at each domain point and forecast the future values with extrapolation of the time series regression. The residuals of those deterministic equations are the stationary component that can be forecasted by spatiotemporal geostatistics, they are composed of a mixture of an intrinsic non-separable spatiotemporal framework, natural oscillations of non-decomposed scales and randomness. Finally, it is possible to combine the forecasted results with the estimation computed by spatiotemporal geostatistical

of the residuals. This procedure, spatiotemporal forecast with local temporal drift, can be applied to any spatiotemporal variable that one may want to generate future maps (spatial distribution). Variables with spatiotemporal characteristics are increasing with big data technology, once smartphones users are always providing geotagged information about traffic jams, health, crimes, among others. Traditional variables are exemplified by meteorological/air quality stations that contain environmental variables; satellite images with spectrum analysis that can represent temperature, air pollution, vegetation quality, soil exposure, and others; security/health stations that report the city criminality and well-being.

The use of auxiliary information or complex models (e.g. machine learning or global climate models) can produce great results. However, these complex models take longs processing periods and most of the times supercomputers. Although it is not the focus of this study, some procedures with auxiliary variables and machine learning are commented on below. The use auxiliary information to model the trend, it is done with geometric temperature trends which are modeled as a function of the day of year and latitude [38] or considering the values measured in satellites in the regression function [39]. Likewise, dynamic secondary information (net precipitation) can be incorporated to model the soil water trend [15]. Another way is to estimate the slope of the linear trend at each pixel using a digital elevation model as secondary data [40]. Another study [41] used the elevation as an external drift to help the estimation of the temperature map considering a linear relationship between temperature and elevation.

Alternatively, some authors use parametric models, machine learning, neural network or advanced data analysis to models the trend, such as genetic algorithm to extract dynamical rules from the data and forecast in confined systems displaying spatiotemporal chaos [42]; seasonal-trend decomposition based on the non-parametric local regression [43]; characterized spatio-temporal processes through non-linear models utilizing physical arguments of wave mode interactions in which medium scales influence the evolution of large-scale modes [44]; neural network models to generate short-term temporal forecasts and local conditional distributions of the observed values to perform spatial stochastic simulations for the entire geographic area of interest [45]; random forest algorithms to fit the trend model for the regression kriging [46]; introduced the use of hidden physics models (data-efficient learning machines expressed by nonlinear partial differential equations) to extract patterns from high-dimensional data generated from experiments [47]; exponentially weighted moving average to model the temporal trend component [48]; nonlinear dynamic spatiotemporal models [49]; proposed the echo state network machine learning approach that can be used to generate long-lead forecasts of non-linear spatio-temporal processes with reasonable uncertainty quantification [50]; fit trend surface model from the spatiotemporal position of the sample study with median polish [51, 52]; embed a nonparametric trend modeling approach in data-driven real-time predictions [53]; stochastic local interaction model for space–time interpolation [54]; deep learning (integro-difference equation) models for spatio-temporal forecasting [55].

2 Methodology

In general, the procedures mentioned in the literature are enough to remove the trend/seasonality and transform a non-stationary into a stationary dataset, in order to apply spatiotemporal geostatistics. Yet, if the spatiotemporal dataset presents a high spatial variability of its time-series, mainly in the trend and cyclicity, the mentioned trend models may not represent well the spatial variability of the time series at every point of the map. In these cases, it is proposed the geostatistical estimation method with local time drift whose steps can be observed in the flowchart of Fig. 1. This method uses the time series analysis at each spatial point to assess its trend and seasonality parameters (local decomposition) and considered them as regional variables to generate each parameter map. The map can be interpolated by a geostatistical or a non-geostatistical method.

The proposed method intends to transform every time series into a combination of the simplest time series model (iid random variable with zero means), some intrinsic spatiotemporal framework that cannot be captured by temporal decomposition and the oscillations of other scales (i.e. climate oscillations such as El Niño, La Niña, Madden–Julian, Milankovitch cycles, Pacific Decadal, Quasi-biennial, and others). For this, the trend and seasonality decomposition are applied at each time series of the studied region, and if it is necessary spatial trends can be removed, too. The simple decomposition equation of a spatiotemporal variable Z(s, t), is:

$$Z(s,t) = U(s,t) + tr(s,t) + sea(s,t) \quad (1)$$

where $U(s,t)$ is the space–time stationary residual, $tr(s,t)$ is the temporal trend and $sea(s,t)$ seasonality at each sampled point s (i.e. weather station location). One can consider that the variable is monitored at N sites s_i and p time points t_j so that the data are on a space–time mesh $\{Z(s_i, t_j\}$, i = 1, …, N; j = 1, …, p} [14]. In practice, assuming temporal stationary is never trivial, and it should guarantee that all data originate from the same underlying physical process

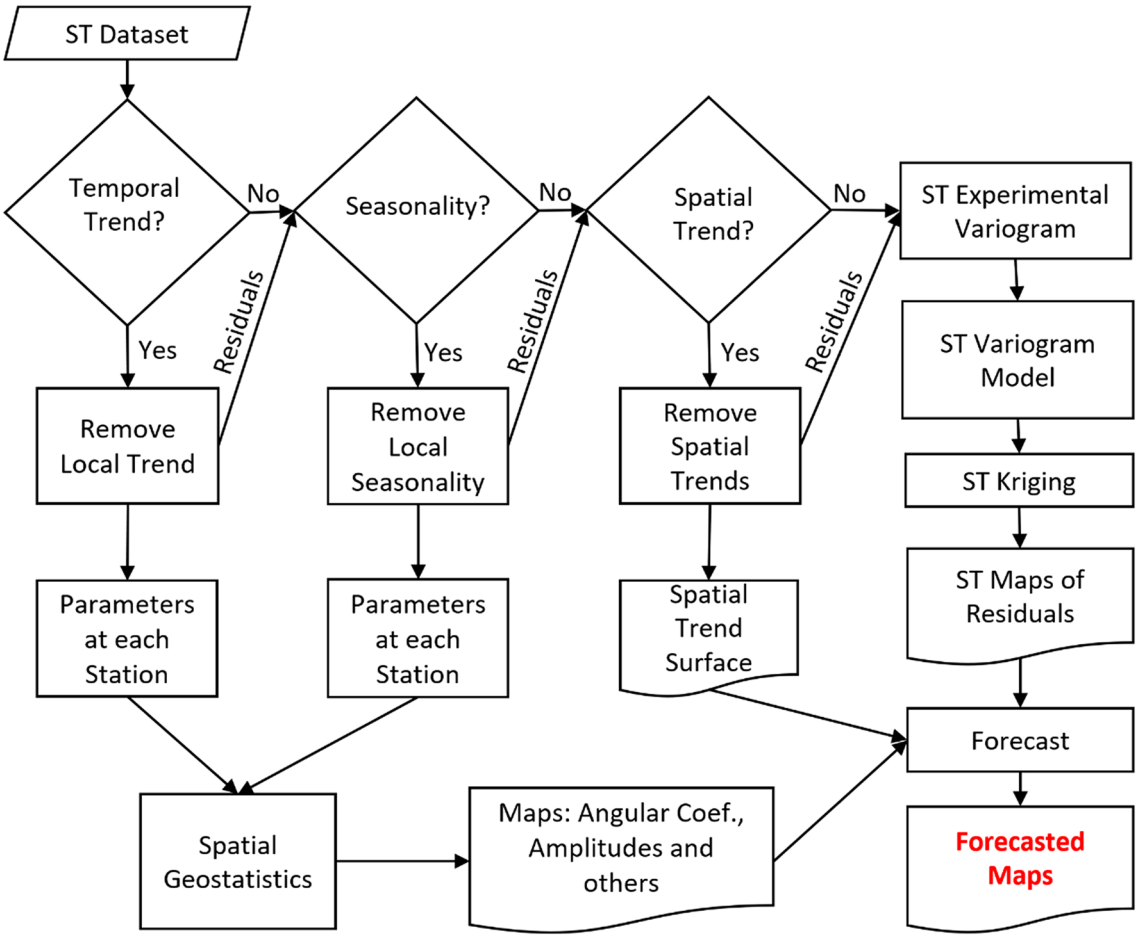

Fig. 1 Flowchart of the spatiotemporal forecast with local temporal drift

[14]. Considering a specific point in space s, the time series is the residual values (res(t)) plus its determinist models, seasonality (sea(t)) and trend (tr(t)), the classical decomposition model:

$$\beta(t) = tr(t) + sea(t) + res(t) \tag{2}$$

A useful technique for estimating $tr(t)$ is the method of least squares, as well as, $sea(t)$ can be modeled as harmonics in Fourier series. The sinusoid of frequency ω can be written as a traditional non-linear function (Eq. 3) and as a linear function (Eq. 4) [56]:

$$sea(t) = R\cos(\omega t + \phi) \tag{3}$$

$$sea(t) = A\cos\omega t + B\sin\omega t \tag{4}$$

where R is the amplitude and ϕ is the phase, $A = R\cos\phi$ and $B = -R\sin\phi$, ω (in radians per unit of time) is the frequency [56]. Furthermore, given any value of A and B, one can solve for R and ϕ (Eq. 3) [56]. Considering A and B as the parameters, the equation is now linear, for fixed ω. So, the calculus is computationally easier if the frequency ω is known and constant [56].

Additionally, the linear regression in the time series context considers that a dependent time series is affected by a set of known and fixed inputs or independent series [57]. One may use a classic linear regression to estimate the trend by fitting the linear model:

$$tr(t) = b_1 + b_2 t \tag{5}$$

where b_i are unknown fixed regression coefficients [57].

The spatiotemporal forecast with local temporal drift depends on two time series considerations. First, the deterministic trends can be extrapolated in order to forecast future steps, so, one can justify the projection by alleging that fundamental trends will commonly change slowly in comparison with the prediction/forecast lead time [58]. Second, the objective is to estimate and define the deterministic elements, linear trend ($tr(t)$) and seasonality ($sea(t)$), expecting that its residual or noise component $res(t)$ are stationary time series [59]. The "residual" $res(t)$ is

the difference between a fitted from an observed value [60].

It is possible to apply the spatiotemporal geostatistics to map the residuals *res*(*t*) at future steps, as well as, to interpolate the *tr*(*t*) and the *sea*(*t*) parameters to all pixels in the domain and deterministically predict the next steps in time. Considering a simple space–time model (Eq. 1) and substituting the components with the ones in the time series analysis (Eqs. 3, 4 and 5), the spatiotemporal process Z(s, t) can be written as:

$$Z(s,t) = U(s,t) + b_1(s) + b_2(s)t + R(s)cos(\omega t + \phi(s)) \quad (6)$$

where, $b_1(s) + b_2(s)t$ is the temporal trend for each point s in space, $R(s)cos(\omega t + \phi(s))$ is the seasonality for each point s in space and $U(s,t)$ is the space–time stationary residuals. Moreover, the parameters from the time series decomposition depend only on its location s in space, these points can be considered as regionalized variables and mapped. The residual is a mixture of the random effect (iid), spatiotemporal interaction and oscillations of other scales.

The spatiotemporal variable $U(s,t)$ can be estimated at every point of the domain and at any desirable time with the spatiotemporal geostatistics. Spatiotemporal continuity is a property that characterizes the relationship between observations at different locations in the space–time framework, such as $U(s_i, t_k)$ and $U(s_j, t_l)$ with i, j = 1,…,N and k, l = 1,…,p [16]. Similarly, to spatial geostatistics, the relationship depends on the spatiotemporal distance between the points $(s_i - s_j; t_k - t_l)$ and its spatiotemporal semivariogram can be defined as the function [16]:

$$\gamma(h,u) = \frac{1}{2}E\left[U(s_i,t_k) - U(s_j,t_l)\right]^2 \quad (7)$$

where (s_i, t_k) and (s_j, t_l) are spatial–temporal points separated by lag h in space and u in time [16]. Most of the theoretical functions to model the spatiotemporal variability are defined as covariance functions. A joint space–time covariance function allows space–time kriging estimation at any spatial position s and any moment t in time [17]. To use the covariance function to estimate, one should consider models that are guaranteed to be valid ones, since any covariance function needs to be nonnegative definite and, in general, the moment estimators of the covariance functions do not satisfy these properties [20]. In addition, to accomplish optimal prediction, $Cov[Z(s_0, t_0), Z(s_n, t_i)]$ for n = 1, …, N, i = 1, …, p, is required, in other words, covariances at unobserved space–time lags must be known [20].

Since the goal of kriging is to predict $U(s_0, t_0)$ from incomplete and noisy data [23]. Considering time as another dimension does not change the form of the kriging estimator nor in the kriging equations [21]. Moreover, apart from characterizing statistical dependencies in space and time, the spatiotemporal covariance can be used for kriging [23]. In that way, if (*s*, *t*) is an unsampled location-time and given $(U(s_1, t_i), \ldots, U(s_N, t_i), i = 1, \ldots, p)$ then U(s, t) could be estimated by [21]:

$$\hat{U}(s,t) = \sum_{i=1}^{p}\sum_{n=1}^{N} \lambda_{n,i} U\left(s_n, t_i\right) \quad (8)$$

where $\lambda_{n,i}$ is the kriging weighs and there are no assumptions about any interrelations between space and the time coordinates of a point [21]. The estimator will interpolate in either space or time and may extrapolate in either space or time [21].

3 Patagonia's land surface temperature case study

Land surface temperature is one of the key parameters in the physics of land surface processes from local through global scales [61] and accesses the effects of climate change [62]. It is the kind of variable that changes rapidly, in both space and time [61, 63, 64]. Therefore, adequate characterization of land surface temperature distribution and its temporal evolution requires measurements with detailed spatial and temporal scales. Given the complexity of surface temperature over land, ground measurements cannot practically provide values over wide areas and it is a spatially inhomogeneous parameter especially in urban areas [65]. With the development of remote sensing from space, satellite data offer the only possibility for measuring temperature over the entire globe with sufficiently high temporal resolution and with complete spatially averaged rather than point values [61]. The land surface temperature derived from satellite observations can be used for a wide range of applications in several fields of study, like agriculture [66, 67], urban climate[65, 68], disasters [69], glaciology [70], land use cover change [71], epidemiology [72] and climate modeling [62]. Different mathematical and statistical methods and satellite data can be used such as combined land surface temperature and vegetation index to detect long term changes in land cover [73]; performed temperature downscaling using principal components of various data [65]; two land surface temperature time series, from geostationary Meteosat Second Generation satellite data and Noah land surface modeling to detect geothermal anomalies and extract the geothermal component of the data [74]; temperature modeled for 15 years (2001–2015) by the geographically weighted regression using imagery of Landsat 7 [75]. Apart from all these examples, the spatio-temporal nature of the remote sensing dataset has not been explored and compared with ground samples.

In contrast with climate modeling, geostatistics has the advantage of being computationally data-driven due to its models depend only on the spatiotemporal covariance of the variable with itself at a settled lag in space and a lag in time. If the variable presents a spatiotemporal continuity, usually, it can be modeled and statistically estimated at any point in space and time. Furthermore, spatiotemporal geostatistics is a useful tool to estimate values in desired locations at a determined time.

In order to study the efficacy of the local temporal drift (kriging the residuals of the local decomposition), 3 geostatistics procedures were applied: (i) spatiotemporal kriging of the original dataset (no decomposition); (ii) spatiotemporal kriging of the global decomposition residuals; (iii) spatiotemporal kriging of the local decomposition residuals. In addition to these 3 geostatistical approaches, the results were compared also with the ARIMA Models Panels (i.e. ARIMA Models considering the influence region of each time series). All 4 methods were applied in the Patagonia region, where the long-term temperature study is particularly interesting for climate change pattern once possible global warming impacts the animal distribution and behavior such as the lizards' spatial distribution [76] and the fish culture quality [77]. Besides that, this region is poorly sampled with high climate diversity, thus, the investigation is done in severe conditions. The comparison with spatiotemporal ordinary kriging (original) and global temporal drift is made for two main reasons: (i) show that the external drift is essential when extrapolating spatiotemporal estimation with geostatistical methods; (ii) compare how much the local trend decomposition is more robust than the global decomposition in an extreme situation. Moreover, the satellite image Land Surface Temperature—LST [78] was considered the true values to calculate the estimation errors. All processing was done in the R project [79], the geostatistics used the 'geoR' [80], 'gstat' [81, 82] and 'spacetime' [83, 84] R packages. The spatiotemporal covariance model was chosen with the help of the 'covatest' [85] R package. All maps are considered as raster images and treated in 'raster' [86] R package.

3.1 Dataset

The dataset used is the Global Summary of The Day [87] and they were obtained in the Climate Data Online (CDO) platform. The downloaded dataset is daily, and it was monthly averaged, considering that at least 10 days to compute an average for a month. Moreover, some weather stations have been discarded due to a very low amount of collected information (days measured) and/or no temperature measurements in the last few years. In practice, the used dataset has at least 25% of measured months, 2 years of continuum measurements and, preferentially, recent data. Another reason to choose this area is testing the methodology in extreme conditions, the selected dataset has only 18 weather stations and it is a very poor sample to cover over one million square kilometers with a high climate diversity. Figure 2 shows the Patagonia weather stations and highlights two stations used as examples. The Patagonia shapefile used was retrieved from the GADM website [88].

The Patagonia weather stations dataset starts monitoring with adequate temporal continuity in 1973 and continues until the present. For a long-term (10 years) comparison, the data set was divided as the calibration period (from January 1973 to December 2007) and the control period (from January 2008 to December 2017). The modeling (decomposition) and forecast were done with the calibration period data, the control period was used just to compare and validate the results' statistics. In order to compute the errors of the method, the forecasted values were compared to LST, which is a satellite image computed to estimate the surface temperature with a resolution of 0.05°. Figure 3 shows the histogram and the boxplot of the calibration period, control period, and the LST. The satellite dataset presents a higher spread temperature range with a higher maximum and a lower minimum, as well as, the control and the calibration periods seem to be similar with a close mean and similar statistical distribution. The main reason for this difference is due to the LST higher territory coverage and some difference between the measured data

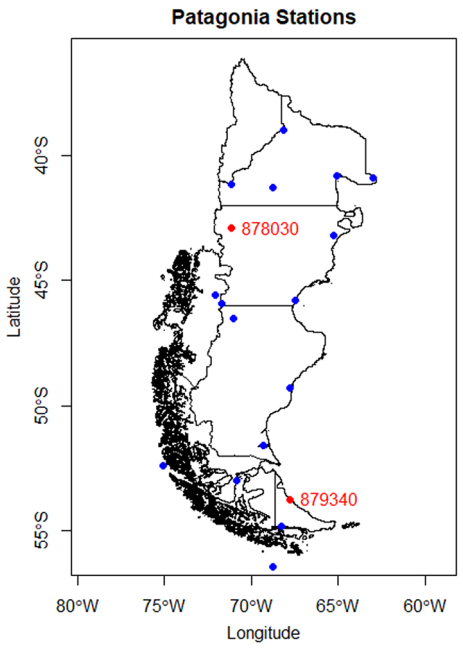

Fig. 2 Patagonia map with the locations of the weather stations in blue and two stations (878,030 and 879,340) used as examples (red)

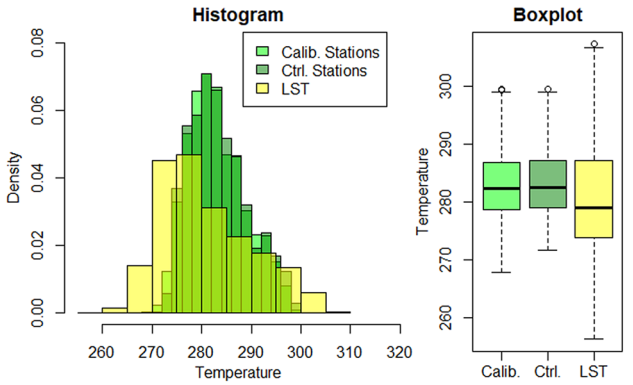

Fig. 3 Histogram and the box plot of the temperature measured at stations and calculated by satellite (LST)

and calculated by the satellite images. LST and the control periods were taken in the same period.

3.2 Decomposition

The calibration period data was decomposed with Eq. 6 that consists of 2 components, one deterministic time series (trend and seasonality) and its residuals ($U(s,t)$). In the global decomposition model, the determinist component as only one temporal model to the entire region, in other words, the model was adjusted to the time series average. That is, the residuals were acquired at each time series location (s) by subtracting the average model from its values. The local decomposition approach takes into account one model at each time series location (s), i.e. the residuals were computed by decomposing each weather station with its time series model. All models were fitted to the dataset using least-squares fitting. Figures 4 and 5 illustrate both decompositions at the selected points of Fig. 2. It is clear that the Global decomposition does not fit well all the temperature at weather stations (i.e. 879,340) due to the regional differences in the time-series pattern. Meanwhile, the local decomposition adjusts the trend and seasonality at each station which signifies that every fit depends on its own local time series. In Figs. 4 and 5 lower panels, it is possible to analyze visually that the global (red dots) and local (blue dots) residuals show no temporal pattern (trend or seasonality). Moreover, five parameters are required to reproduce the deterministic equation, the frequency is constant as twelve months and the other four parameters were recorded as regionalized variables, amplitude (R), intercept (b_1), angular coefficient (b_2) and phase (ϕ).

In order to evaluate the accuracy of the time series models, the Root Mean Square Error (RMSE) was applied. Its computation is expressed as [89]:

$$RMSE = \sqrt{E\left[e_t^2\right]} \tag{9}$$

where e_t is the error (time series sample − time series model) and there are calculated at each time t. The RMSE is largely used due to its theoretical relevance in statistical

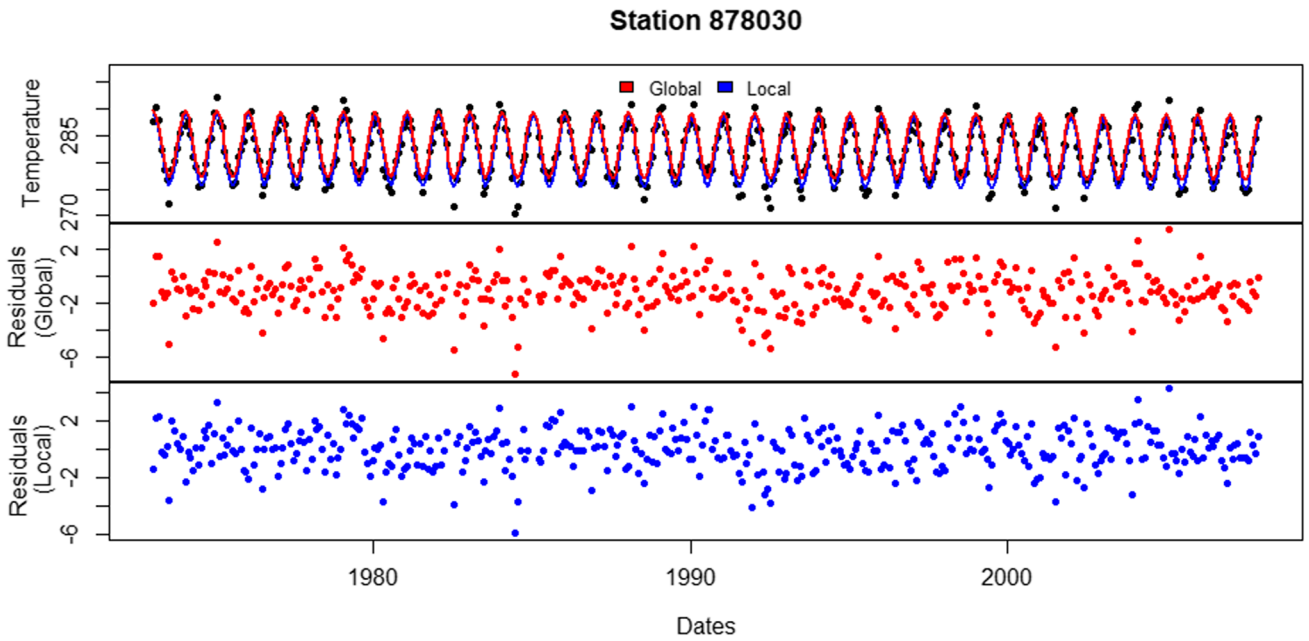

Fig. 4 Temperature measured (black dots) at weather station 878,030, the global (red line) and local (blue line) models and the residuals data with the local (red) and the global (blue) methods

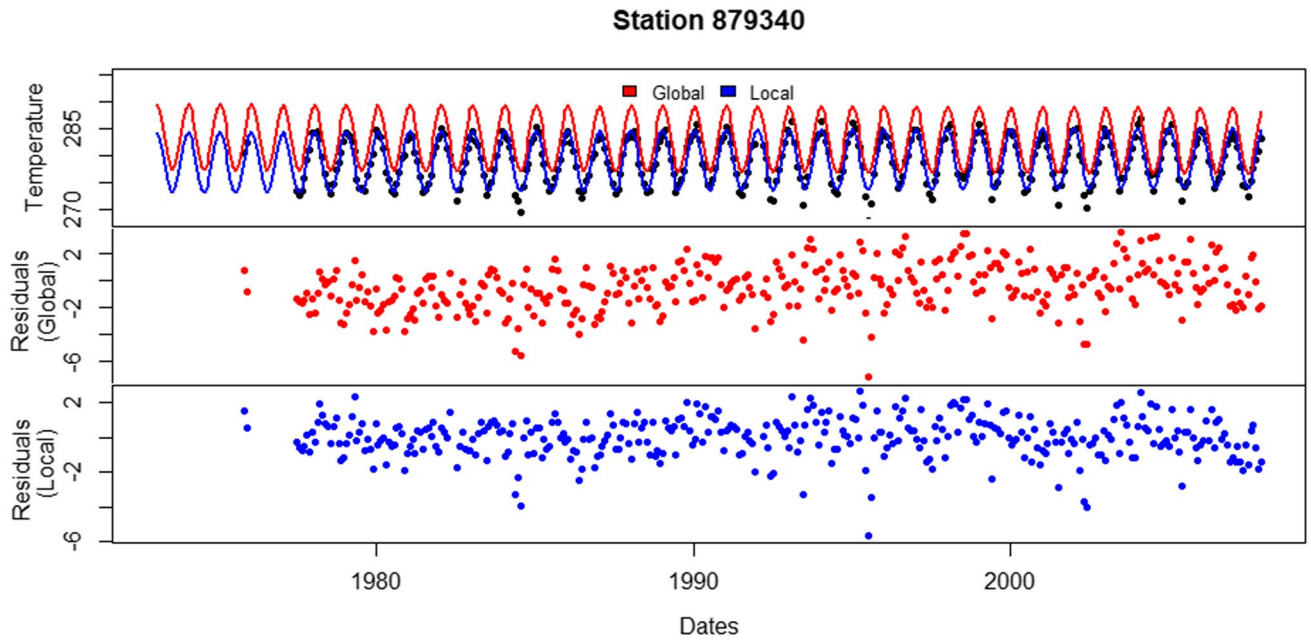

Fig. 5 Temperature measured (black dots) at weather station 879,340, the global (red line) and local (blue line) models and the residuals data with the local (red) and the global (blue) methods

modeling and it has the advantage of being at the same scale of the data [89].

Figure 6 illustrates the statistics with the histogram and the boxplot of the local and the global decomposed residuals' data and its error (RMSE). It is computed one RMSE for each weather station of the calibration period to verify how good are the fitted decompositions models. Thus, it is clear that the local decomposition models fitted better because its residuals are more concentrated near the zero and its RMSE is significantly smaller (RMSE = 1.33) when compared with the global decomposition (RMSE = 3.83).

3.3 Geostatistical analysis

In order to forecast for the entire region, it is necessary to know every parameter of the Fourier analysis (amplitude (R), intercept (b_1), angular coefficient (b_2) and phase (ϕ)) at all locations of the meteorological stations. The parameters are retrieved at each locations from Eq. 6. Moreover, the best way to spatially estimate these parameters is the application of the kriging techniques. If it is not possible to model the variogram, one could use an interpolation method as the inverse of distance, spline, radial basis

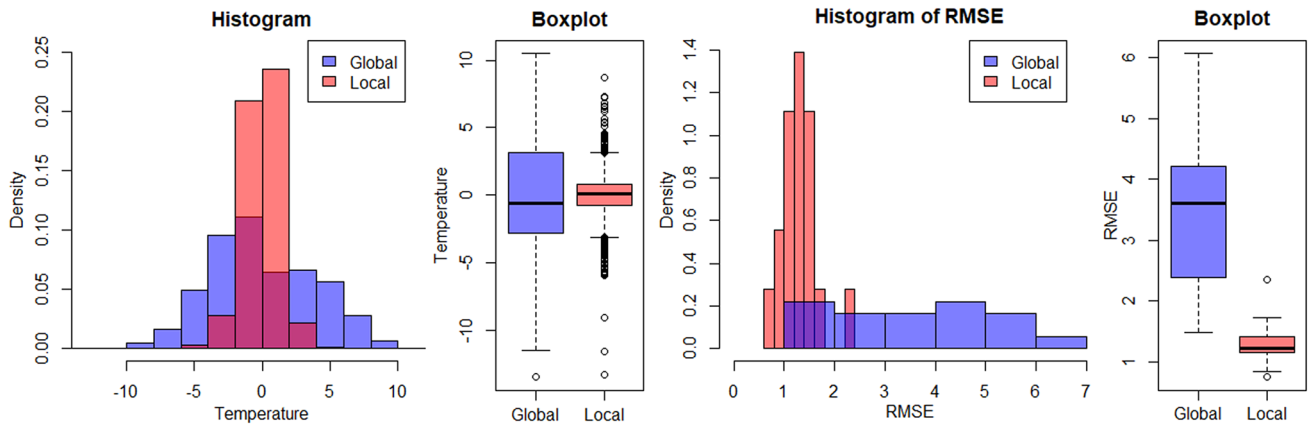

Fig. 6 Histogram and the box plot of the global and local residuals (left) and its errors (RMSE)

function, nearest point or others. Figure 7 points out the station's location and the parameters to be mapped.

Moreover, Fig. 8 shows that only the amplitude and the phase have modelable variograms and the anisotropy could not be recognized. Probably, if the region had a denser dataset, the angular coefficient and the intercept variograms could be structured, as well as, all variograms could present anisotropy. The region has only 18 weather stations and certainly more stations would enrich these models. Whenever the spatial continuity could present some structured variogram its model should be considered and applied for the estimation method. The trends can be identified from the experimental variogram, which keeps increasing above the theoretical sill [19]. Thereby, the variograms of all the time series parameters were verified whether they have any first or second-order spatial trends, and amplitude shows a structured variogram when the spatial first-order polynomial was adjusted and removed. The variogram of the amplitude was modeled with the spherical model with a nugget effect of 0.12, the sill at 0.6 and range equal to 4. As well as, the phase variogram was modeled with the spherical model, no nugget effect, sill at 0.01 and range equals to 5. The semivariogram is the graph of the semivariance (a measure of the spatial discontinuity) by its distances and its parameters are: the nugget effect, that is the variability (variance or covariance) at a very close (near zero) sample distance; the sill, that is the higher variability or where the curve reaches its plateau and; the range, that is the distance that the curve reaches the sill. All variogram parameters were adjusted interactively.

For the purpose of mapping, the chosen cell size is 0.5° in latitude and longitude. And the ordinary kriging used a global search neighborhood. The alternative for the parameters without variogram models (angular coefficient and intercept) was the interpolation by inverse distance weighted with the power of two. Figure 9 illustrates the resultant maps for each parameter. The figure shows a clear trend in the amplitude indicating that the seasonality of the time series temperature is minimized in the Andes, that is, the monthly average temperature there is more stable during the year. The same patters are observed in

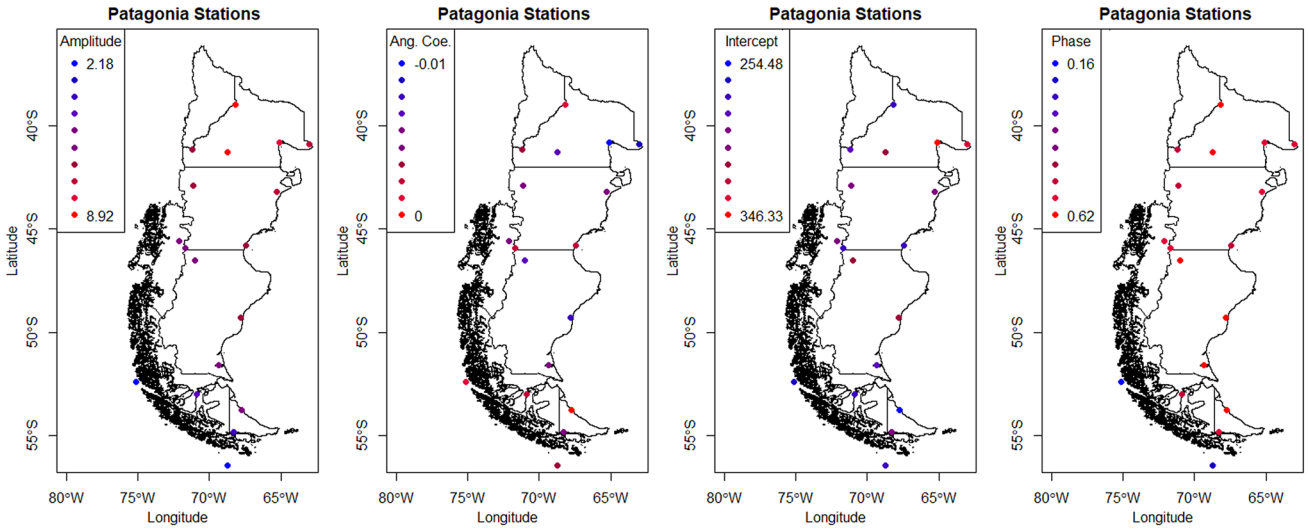

Fig. 7 Time series parameters (amplitude, angular coefficient, intercept and, phase), presented as regionalized variables

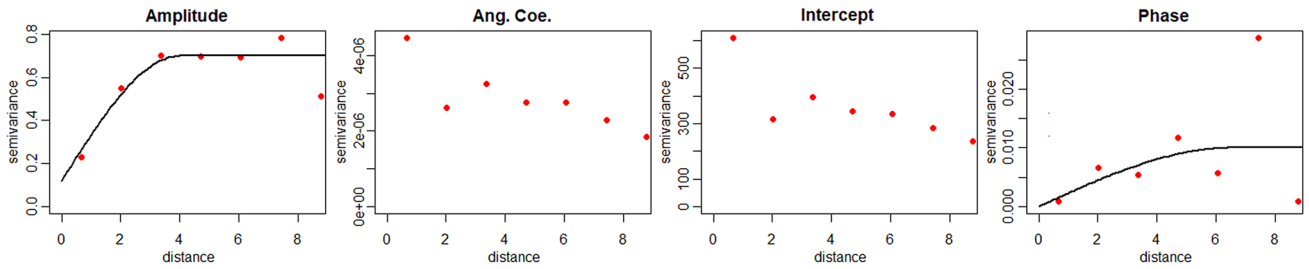

Fig. 8 Experimental semivariograms (red dots) and its models (lines) of the time series parameters as a regionalized variable

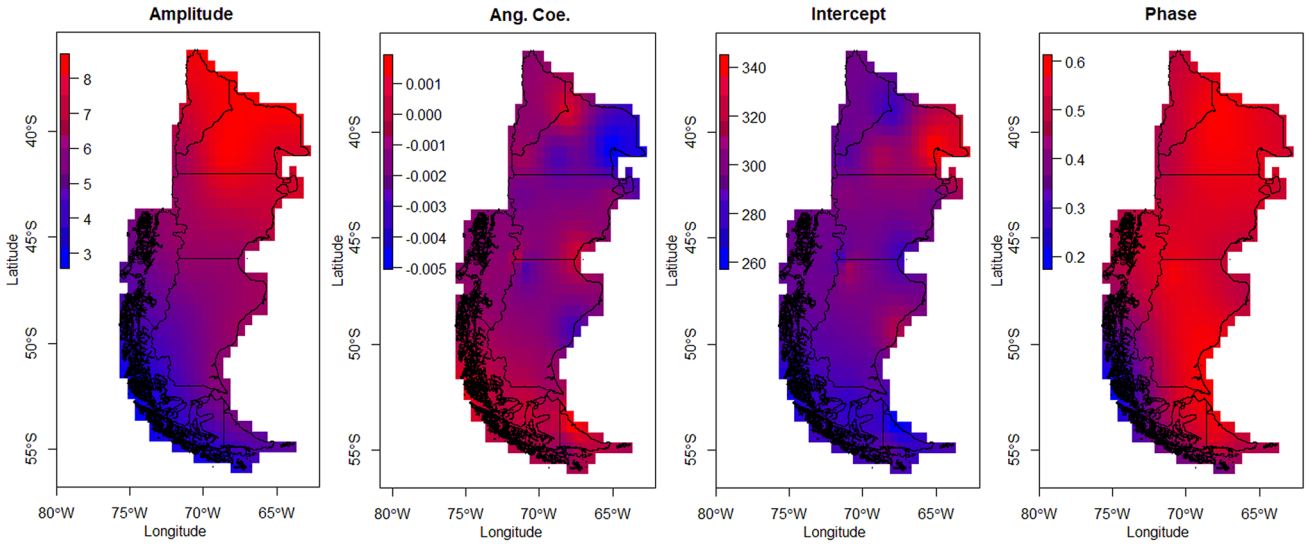

Fig. 9 Spatial interpolated time series parameters

the phase, illustrating that the sine wave that represents the seasonality is in a different phase in the pacific coast.

The spatiotemporal variogram was calculated for the original and the two decomposed datasets. To choose the adequate spatiotemporal model, the test of separability [90, 91] was applied for all datasets and every result does not reject the null hypothesis at 0.05 level of significance. The separability characteristics mean that the spatiotemporal covariance can be defined by a purely spatial and purely temporal covariance [16]. The function is separable if $C(h, u)/C(h,0) = C(0, u)/C(0,0)$, for all h and u and, since the $C(0,0)$ is constant, the equation can be simplified as $C(h, u) \sim C(h, 0)C(0, u)$ [16, 92]. The Product Covariance Model [21] was used to model every experimental variogram in this study. The authors explain that the product of the covariances separate the dependence on the two and it is the simplest way to model a variogram in space–time. The method multiplies the spatial and the temporal covariances (Eq. 10) and, it can also be written as the variogram function (Eq. 11). Both functions are defined [21] as:

$$C(h, u) = C_s(h)C_t(u), \quad (h, u) \in R^n \times T \tag{10}$$

$$\gamma(h, u) = C_s(0)\gamma_t(u) + C_t(0)\gamma_s(h) - \gamma_s(h)\gamma_t(u), \quad (h, u) \in R^n \times T \tag{11}$$

where C_s is the spatial covariance, C_t is the temporal covariance, γ_s is the spatial variogram, γ_t is the temporal variogram [21]. Thus, the spatiotemporal variogram model of all datasets was detached between the spatial component and the temporal component. This separability makes the spatial anisotropy easy to access, that is, one can model the 2D (or 3D if working with a three-dimension spatial dataset) variogram with anisotropy in the spatial component $\gamma_s(h)$ or $C_s(h)$. The variogram model of the original data has no nugget effect, the joint sill at 22, the spatial range at 8° and the temporal at 45 months. The variogram model of the global decomposition residuals shows no nugget effect, the joint sill at 14, the spatial component with the same sill and range at 8°, and the temporal component with the sill at 2.52 and the range at 12 months. Last but not least, the variogram model of the local decomposition residuals exhibits no nugget effect, the joint sill at 2, the spatial part with the sill at 1.02 and the range at 6°, and the temporal part with the sill at 1.6 and the range at 12 months. Although the spatial components seem to have a spatial trend, the polynomials analysis does not indicate any improvement in the spatial variance. Some authors [14, 93–95] explain that the variogram should be fitted manually due to the decisions that the user must make and the automatic algorithms can make mistakes, such as anisotropy direction, nugget effect, structures (quantity and type), guarantee the positive definite function and consider the points as independent observations. For that reason, all analysis, fitting and adjustments in the spatiotemporal variogram models were done manually with exhausting tests to guarantee that the final results follow the best practices in geostatistics. All the spatiotemporal variogram parameters are the same as the spatial, however, the range considers the distance in the space and in the time components.

The original variogram shows a strong temporal seasonality that is partially maintained in the global decomposition residuals. Meanwhile, the local decomposition residuals modeled and removed almost all the trend and seasonality so well that leads the experimental variogram to an almost pure nugget effect in the temporal

component. The three experimental variograms and their respective models are represented in Fig. 10.

The spatiotemporal ordinary kriging as computed to all three datasets with their respective variogram models. The cell size is 0.5° in latitude and longitude in space and every month for the next 10 years (January 2008 to December 2017). Finally, both decomposed datasets were transformed back with its original deterministic model (trend and seasonality) with the Eq. 6. The global decomposition method used averaged parameters and the local decomposition approach used the parameters beforementioned (Fig. 9).

Another comparison is considering the influence area of each time series. That is, the area (panels or sub-regions) of each time series are made of its closest grid points. Thus, ARIMA Models can forecast one value for the entire panels at each time. The ARIMA Models forecast was executed by automatic ARIMA ('forecast' R package) [96, 97]. The ARIMA Models were introduced by [98]; and are explained by several authors [58, 99–105]. The ARIMA Models and its forecast are presented in Fig. 11 at the weather station 978,030 and 879,340. It is clear that the seasonality is well defined in these models.

3.4 Results and comparison

In order to determine the most accurate and precise method, the results are compared with the temperature measured at the weather stations (calibration and control periods) and the temperature computed by satellite images. Table 1 display their statistics and it is easy to check that all approaches maintained the stations average, meanwhile, they smoothed the results, that is, the minimum is higher, and the maximum is lower in all cases. Consequently, the standard deviation tends to be lower as the range of values is narrow. The ARIMA Panels and the local decomposition method smoothed less than the global method and the smoothness of the spatiotemporal ordinary kriging is brutal with the temperature 283.3 K in all quartiles.

Similarly, the average and the standard deviation of each month was calculated and compared among all methods (Fig. 12). It is straightforward that the spatiotemporal kriging tends to the values' averages at the variogram range and does not represent the temporal oscillation, either the spatial distribution. Both decomposition methods represented properly the average temperature of the control period. Moreover, the standard deviation graph illustrates that the kriging and the global decomposition approaches go to zero due to the values is the average after the influence of the variogram's range, that is, only one global average is considered at all spatial points for each month. Besides that, the local decomposition approach repeated the stations' pattern with smaller values, probably due to the smooth effect. It is straightforward that the temporal average and standard deviation of ARIMA Panels could follow the dataset (weather stations temperature) patterns, once its forecast (Fig. 11) and statistics (Table 1) shows no smoothing. The mean temperature range of the stations is narrower, and the standard deviation is systematically lower than the respective results of LST. The statistical difference between the stations and LST is due to insufficient sampling.

Another results' comparison is to measure its inaccuracy, one robust way of doing this is computing the RMSE (Eq. 9). This calculation requires the true values at every point and while there are no real values at every single pixel, the temperature calculated by the satellite image LST was used to compute the estimation errors. The LST resolution (0.05°) dataset was adjusted by aggregating the pixels to be at the same resolution of the cell sizes used in the interpolation (0.5°). Firstly, the RMSE of the control period (all weather stations) is 3.13, which means that the temperature taken at the weather stations is different from the calculated in LST. Besides that, considering all points in the grid at all study period (January 2008 to December 2017), the prediction methods: spatiotemporal kriging, with global and local decompositions and ARIMA Panels have RMSE, respectively, 9.42, 7.05, 5.44 and 5.30. Thus, the comparison of all methods shows that the ARIMA Panels and the local decomposition method have a significantly lower error and it is slightly above the control periods' error.

Withal, the error can also be assessed spatially (compute the error of the temporal values at each pixel) or temporally (compute the error the spatial values at each month). Figure 12 shows the spatial error, in other words, the RMSE was calculated at each pixel. That is, the error of every single pixel comparing the estimated value and the LST value of all months (January 2008 to December 2017). The figure exposes that, spatially, the local and the ARIMA Panels approaches have an inferior error because they are visually more blueish, and the Andes region presents the higher values in all methods. Furthermore, a large part of the error in the stations and consequently in all methods is caused by scarce information where the temperature oscillation is higher, the Andes (Fig. 13).

The temporal error can be measured each month comparing all the estimated values with the LST on that date. It is an adjustment of Eq. 9 that calculates the spatial points error at a fixed time. The graph in Fig. 14 clarifies the improvement of the local decomposition and the ARIMA Panels approaches, their errors are almost always lower than the global decomposition approach and the regular kriging procedure, as well as the temporal patterns, reflect the one in the control period.

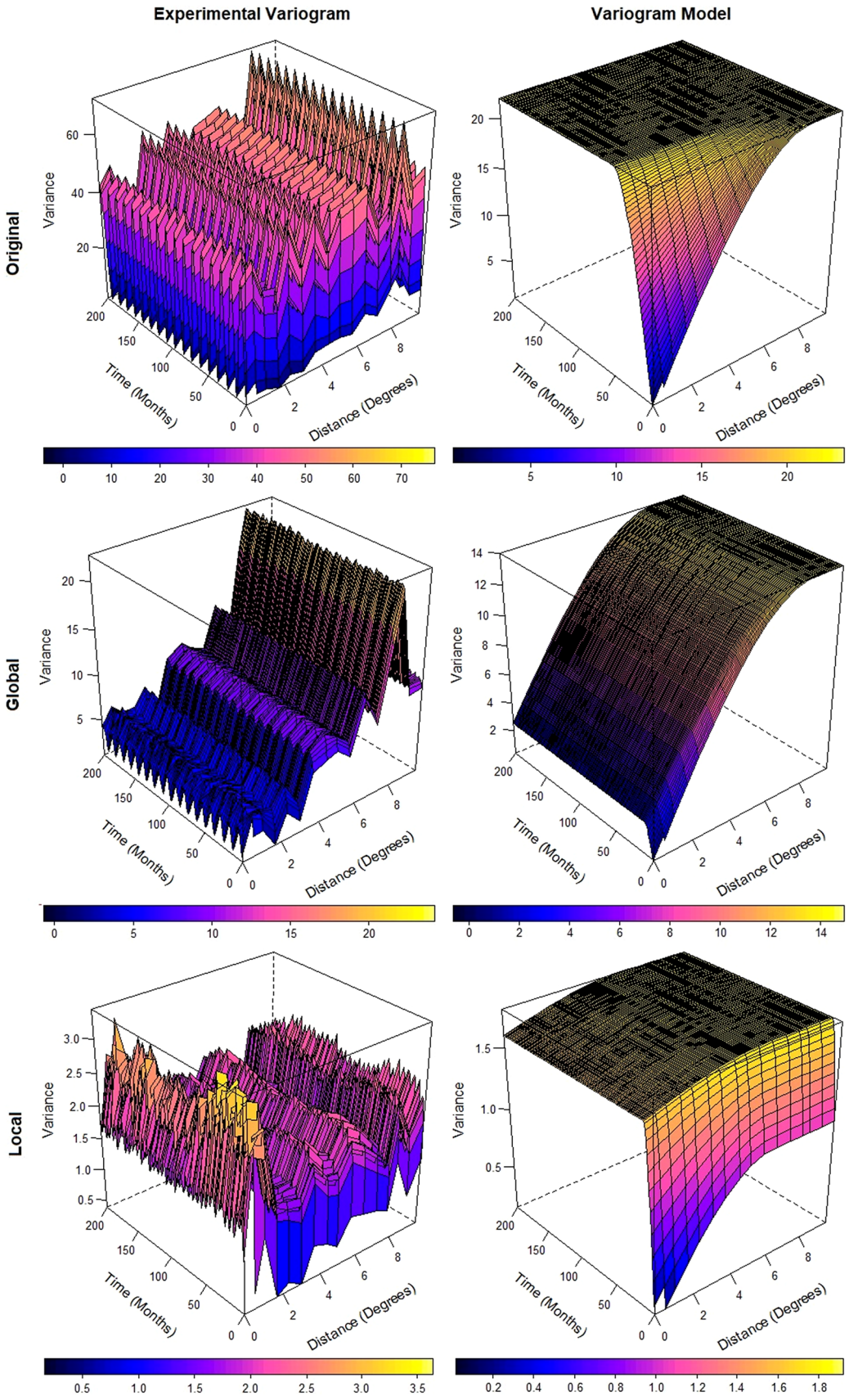

Fig. 10 Experimental and model variogram of the original and residuals from global and local decompositions

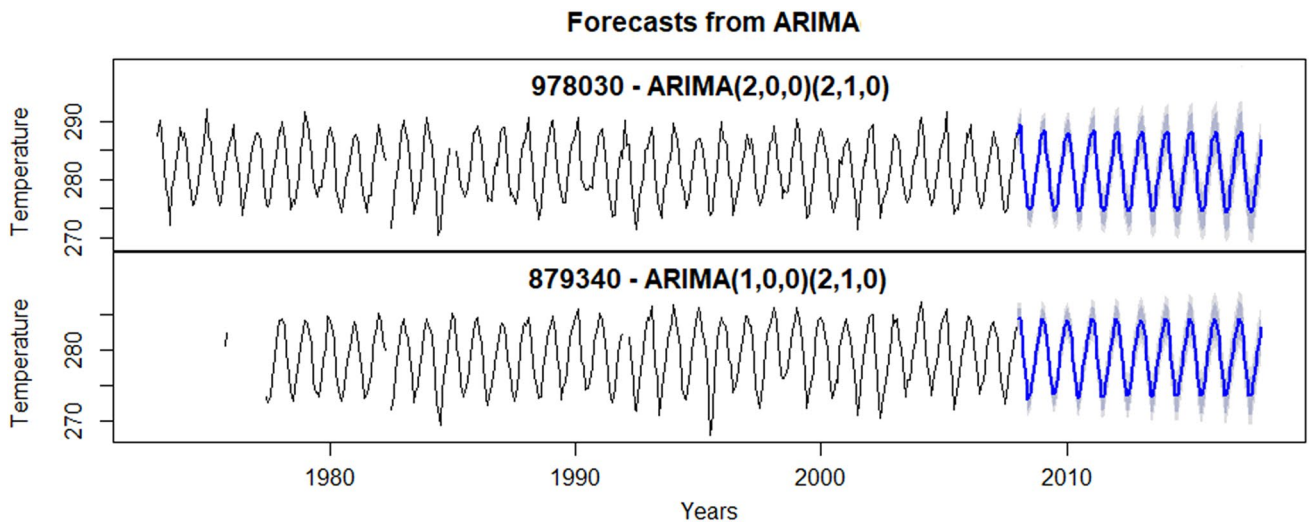

Fig. 11 ARIMA models with forecast (blue lines) at weather stations (978,030 and 879,340)

Table 1 Statistical analysis of the stations, LST, kriged, local and global estimations

	Min	1st	Median	Mean	3rd	Max	Std. Dev
Calib. stations	267.9	278.8	282.4	283.1	287	299.6	5.86
Ctrl. stations	271.8	279.1	282.6	283.5	287.2	299.5	5.78
LST	256.3	273.9	279.0	280.9	287.2	307.3	9.17
Krig	281.5	283.3	283.3	283.6	283.3	294.7	1.18
Global	274.7	278.9	282.7	282.7	286.5	295.9	4.39
Local	273.7	278.6	282.1	282.8	286.2	296.9	4.97
ARIMA Panels	272.9	278.9	282.2	283.1	287.1	297.3	5.61

4 Conclusion

The local temporal drift approach demonstrated to be an efficient way to map the deterministic parameters and estimate a spatiotemporal variable at any location at any time even in a poorly sampled region. Despite this, the map of the decomposition parameters was achieved by ordinary kriging and inverse distance weighted and the results seem to be satisfactory for an area with a few stations. To ensure the second-order stationary for the entire area, the local decomposition removed, deterministically, not only its trend but also its cyclicity. The temporal drift at each spatial location (station) guarantees the best fit model for the complete dataset and not mixing space and time dimension. The global decomposition method considers that the temporal oscillation is the same for the entire area, which may not be the case, especially, when forecasting massive regions with several temporal patterns. However, the global external drift is easier to set and computationally less demanding. That is, if one is working in a small area that the time series parameters are similar to all spatial points, the global decomposition method may present acceptable results. In general, the ARIMA Panels presented results very close to the local decomposition approach, that is, low smoothness and low errors, as well as, expected temporal patterns in the average temperature and standard deviation. The major advantage of the ARIMA Panels is that this method does not need a spatial correlation model, so it is good for a poorly sampled region as Patagonia. On the other hand, this method cannot generate well-distributed maps (only one value per sub-region) and the spatial continuity may be lost in long-term studies due to the independence of each time series forecast.

The proposed method is notably good when the st-kriging can be used to estimate the spatial distribution of the residual, in other words, estimation considers the residual spatial correlation while the variogram does not reach the sill. Nevertheless, the ten-year analysis indicates that the local decomposition approach can also be applied to long-term forecast due to three main reasons: (i) the average temperature is very close to the observed in the control period; (ii) the standard deviation follows the same pattern of the control period; (iii) the error is inferior of

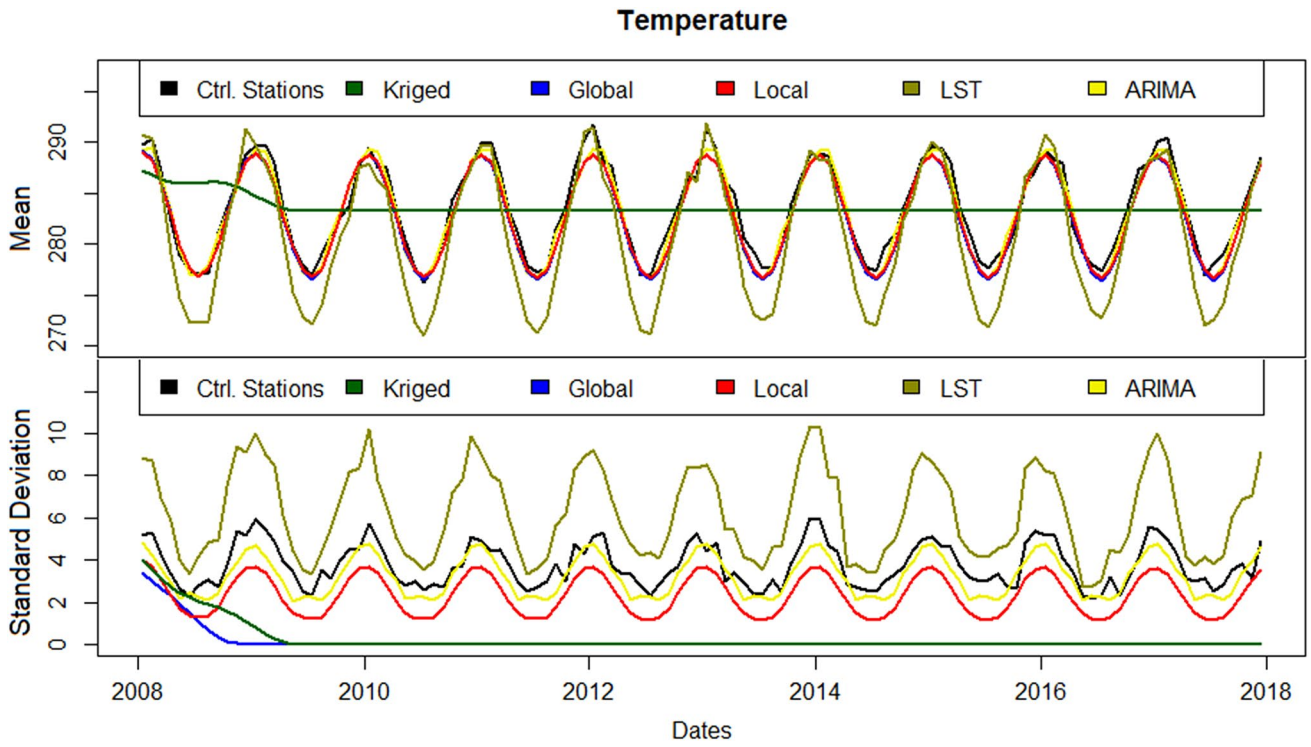

Fig. 12 Mean and the standard deviation of the temperature at each month

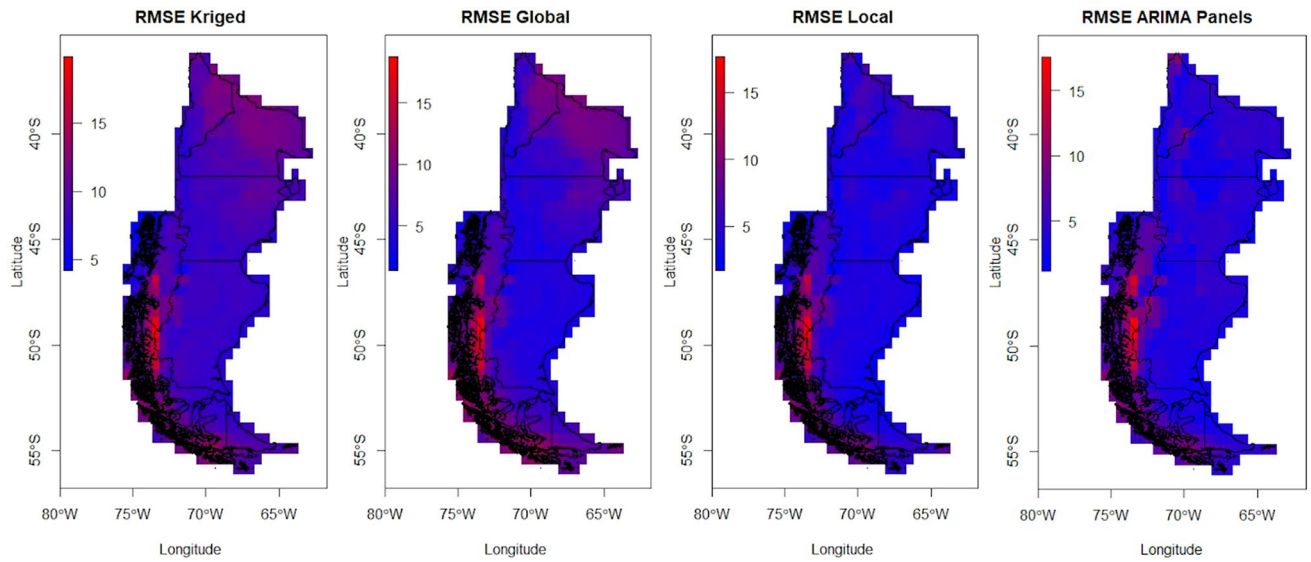

Fig. 13 Error (RMSE) at each estimated pixel

the st-kriging and global decomposition method and it presents a similar pattern of the control periods' error. Although the temporal correlation of the case study is not long enough to cover the entire forecast period, the results beyond the temporal (variogram) range may not be influenced by the kriging of the residuals. That means that after the temporal (variogram) range the results are just a model of its trend and seasonality. In order to keep the space–time covariance in the long-term forecast, one may apply a sequential stochastic simulation instead of kriging. Another point is that it is expected that this method should perform better when the covariance model is not

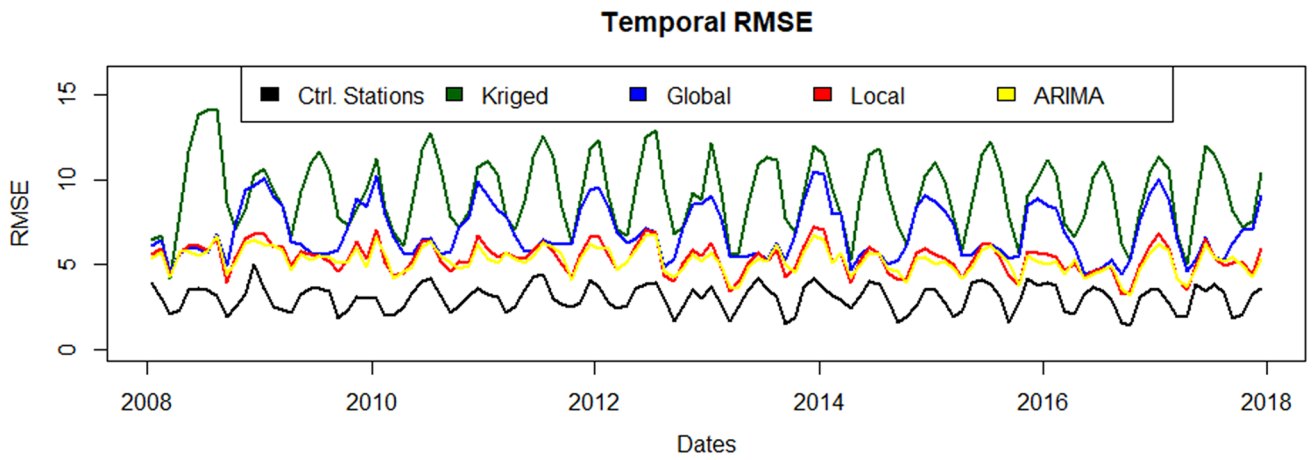

Fig. 14 Error (RMSE) calculated at each month with the three methods and the original values at stations

separable, that is, the spatiotemporal framework could not be entirely separable. Thus, the variogram model of the residuals should represent the intrinsic space–time continuity.

Acknowledgements The authors thank everyone involved in collecting, processing, storing and distributing the data used in this study. The MOD11C3 data product was retrieved from the online, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access/data_pool. Data from weather stations were obtained from the National Oceanic and Atmospheric Administration (NOAA) National Center for Environmental Information (NCEI). Besides, the author is also grateful to those involved in the R and GADM projects and all collaborators of the R packages. Last but not least, the authors are grateful to CAPES (Coordination for the Improvement of Higher Education Personnel/Brazil) and IGc/USP (Institute of Geosciences/University of São Paulo) for the first author's Ph.D. scholarship.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Robeson SM (1997) Statistical consideration. In: Thompson RD, Perry A (eds) Applied climatology: principles and practice. Routledge, New York, pp 22–35
2. Stull RB (2000) Meteorology for scientists and engineers, 2nd edn. Brooks/Cole, Pacific Grove
3. Lau W (2006) ESSAY: model interpretation of climate signals: an application to Asian monsoon climate. In: Bridgman H, Oliver J (eds) The global climate system: patterns, processes, and teleconnections. Cambridge University Press, Cambridge, pp 281–308
4. Lorenz EN (1963) Deterministic nonperiodic flow. J Atmos Sci. https://doi.org/10.1175/1520-0469(1963)020%3c0130:DNF%3e2.0.CO;2
5. Lorenz EN (1996) Predictability: a problem partly solved. In: Proceedings ECMWF seminar on predictability, vol. I, reading, United Kingdom, ECMWF, pp 1–18
6. Xia Y, Leung H, Chan H (2006) A prediction fusion method for reconstructing spatial temporal dynamics using support vector machines. IEEE Trans Circuits Syst II. https://doi.org/10.1109/TCSII.2005.854585
7. Dhanya CT, Kumar DN (2010) Nonlinear ensemble prediction of chaotic daily rainfall. Adv Water Resour. https://doi.org/10.1016/j.advwatres.2010.01.001
8. Wang S, Li G, Iskandarani M, Hénaff ML, Knio OM (2018) Verifying and assessing the performance of the perturbation strategy in polynomial chaos ensemble forecasts of the circulation in the Gulf of Mexico. Ocean Model. https://doi.org/10.1016/j.ocemod.2018.09.002
9. Das M, Ghosh SK (2019) FB-STEP: a fuzzy Bayesian network based data-driven framework for spatio-temporal prediction of climatological time series data. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2018.08.057
10. Henderson-Sellers A, McGuffie K (1997) Climate models. In: Thompson RD, Perry A (eds) Applied climatology: principles and practice. Routledge, New York, pp 36–50
11. Matheron G (1963) Principles of geostatistics. Econ Geol 58:1246–1266
12. Matheron G (1971) The theory of regionalized variables and its applications. École Nationale Supérieure des Mines de Paris, Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau
13. Jenkins GM, Watts DG (1968) Spectral analysis and its applications. Holden-Day, San Francisco
14. Chilès JP, Delfiner P (1999) Geostatistics: modeling spatial uncertainty. Wiley, New York
15. Snepvangers JJJC, Heuvelink GBM, Huisman JA (2003) Soil water content interpolation using spatio-temporal kriging with external drift. Geoderma. https://doi.org/10.1016/S0016-7061(02)00310-5
16. Montero J-M, Fernández-Avilés G, Mateu J (2015) Spatial and spatio-temporal geostatistical modeling and kriging. Wiley, Chichester
17. Kyriakidis P, Journel AG (1999) Geostatistical space-time models: a review. Math Geol. https://doi.org/10.1023/A:1007528426688
18. De Iaco S, Posa D, Cappello C, Maggio S (2019) Isotropy, symmetry, separability and strict positive definiteness for

covariance functions: a critical review. Spatial Stat. https://doi.org/10.1016/j.spasta.2018.09.003
19. Deutsch CV (2002) Geostatistical reservoir modeling. Oxford University Press, New York
20. Sherman M (2011) Spatial statistics and spatio-temporal data: covariance functions and directional properties. Wiley, Chichester
21. De Cesare L, Myers DE, Posa D (1997) Spatial-temporal modeling of SO2 in Milan District. In: Baafi EY, Schofield NA (eds) Geostatistics wollongong 96. Springer, New York, pp 1031–1042
22. De Cesare L, Myers DE, Posa D (2001) Product-sum covariance for space-time modeling: an environmental application. Environmetrics. https://doi.org/10.1002/1099-095X(200102)12:1%3c11:AID-ENV426%3e3.0.CO;2-P
23. Cressie N, Wikle CK (2011) Statistics for spatio-temporal data. Wiley, Hoboken
24. De Iaco S, Posa D (2012) Predicting spatio-temporal random fields: Some computational aspects. Comput Geosci. https://doi.org/10.1016/j.cageo.2011.11.014
25. De Iaco S, Palma M, Posa D (2015) Spatio-temporal geostatistical modeling for French fertility predictions. Spatial Stat. https://doi.org/10.1016/j.spasta.2015.10.002
26. Menezes R, Piairo H, Garcia-Soidán P, Sousa I (2016) Spatial–temporal modellization of the NO_2 concentration data through geostatistical tools. Stat Methods Appl. https://doi.org/10.1007/s10260-015-0346-3
27. Monteiro A, Menezes R, Silva ME (2017) Modelling spatio-temporal data with multiple seasonalities: the NO_2 Portuguese case. Spatial Stat. https://doi.org/10.1016/j.spasta.2017.04.005
28. Sølna K, Switzer P (1996) Time trend estimation for a geographic region. J Am Stat Assoc. https://doi.org/10.1080/01621459.1996.10476927
29. Adamowski K, Bocci C (2001) Geostatistical regional trend detection in river flow data. Hydrol Process. https://doi.org/10.1002/hyp.1045
30. Paci L, Gelfand AE, Holland DM (2013) Spatio-temporal modeling for real-time ozone forecasting. Spat Stat. https://doi.org/10.1016/j.spasta.2013.04.003
31. Krzyszczak J, Baranowski P, Zubik M, Kazandjiev V, Georgieva V, Cezary S, Siwek K, Kozyra J, Nieróbca A (2018) Multifractal characterization and comparison of meteorological time series from two climatic zones. Theoret Appl Climatol. https://doi.org/10.1007/s00704-018-2705-0
32. Majumder S, Kanjilal PP (2019) Application of singular spectrum analysis for investigating chaos in sea surface temperature. Pure Appl Geophys. https://doi.org/10.1007/s00024-019-02140-4
33. Rigal A, Azais J-M, Ribes A (2019) Estimating daily climatological normals in a changing climate. Clim Dyn. https://doi.org/10.1007/s00382-018-4584-6
34. Bendre M, Manthalkar R (2019) Time series decomposition and predictive analytics using MapReduce framework. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2018.09.017
35. Isensee J, Datseris G, Parlitz U (2019) Predicting spatio-temporal time series using dimension reduced local states. J Nonlinear Sci. https://doi.org/10.1007/s00332-019-09588-7
36. Liu H, Zhan Q, Yang C, Wang J (2019) The multi-timescale temporal patterns and dynamics of land surface temperature using Ensemble Empirical Mode Decomposition. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2018.10.252
37. Deng Q, Fu Z (2018) Comparison of methods for extracting annual cycle with changing amplitude in climate series. Clim Dyn. https://doi.org/10.1007/s00382-018-4432-8
38. Kilibarda M, Hengl T, Heuvelink GBM, Graler B, Pebesma E, Tadić MP, Bajat B (2014) Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. J Geophys Res. https://doi.org/10.1002/2013JD020803
39. Kilibarda M, Tadić MP, Hengl T, Luković J, Bajat B (2015) Global geographic and feature space coverage of temperature data in the context of spatio-temporal interpolation. Spat Stat. https://doi.org/10.1016/j.spasta.2015.04.005
40. Wang H, Pardo-Igúzquiza E, Dowd PA, Yang Y (2017) Optimal estimation of areal values of near-land-surface temperatures for testing global and local spatio-temporal trends. Comput Geosci. https://doi.org/10.1016/j.cageo.2017.06.002
41. Hudson G, Wackernagel H (1994) Mapping temperature using kriging with external drift: theory and an example from Scotland. Int J Climatol 14:77–91
42. López C, Álvarez A, Hernández-García E (2000) Forecasting confined spatiotemporal chaos with genetic algorithms. Phys Rev Lett. https://doi.org/10.1103/PhysRevLett.85.2300
43. Montero-Lorenzo J-M, Fernández-Áviles G, Mondéjar-Jiménez J (2013) A spatio-temporal geostatistical approach to predicting pollution levels: The case of mono-nitrogen oxides in Madrid. Comput Environ Urban Syst. https://doi.org/10.1016/j.compenvurbsys.2012.06.005
44. Gladish DW, Wikle CK (2014) Physically motivated scale interaction parameterization in reduced rank quadratic nonlinear dynamic spatio-temporal models. Environmetrics. https://doi.org/10.1002/env.2266
45. Russo A, Soares AO (2014) Hybrid model for urban air pollution forecasting: a stochastic spatio-temporal approach. Math Geosci. https://doi.org/10.1007/s11004-013-9483-0
46. Gasch CK, Hengl T, Graler B, Meyer H, Magney TS, Brown DJ (2015) Spatio-temporal interpolation of soil water, temperature, and electrical conductivity in 3D + T: the Cook Agronomy Farm data set. Spat Stat. https://doi.org/10.1016/j.spasta.201504.001
47. Raissi M, Karniadakis GE (2018) Hidden physics models: Machine learning of nonlinear partial differential equations. J Comput Phys. https://doi.org/10.1016/j.jcp.2017.11.039
48. Varouchkis EA, Hristopulos DT (2017) Comparison of spatiotemporal variogram functions based on a sparse dataset of groundwater level variations. Spat Stat. https://doi.org/10.1016/j.spasta.2017.07.003
49. Richardson RA (2017) Sparsity in nonlinear dynamic spatiotemporal models using implied advection. Environmetrics. https://doi.org/10.1002/env.2456
50. McDermott PL, Wikle CK (2017) An ensemble quadratic echo state network for non-linear spatio-temporal forecasting. Statistics. https://doi.org/10.1002/sta4.160
51. Bruno F, Guttorp P, Sampson P, Cocchi D (2009) A simple non-separable, non-stationary spatiotemporal model for ozone. Environ Ecol Stat. https://doi.org/10.1007/s10651-008-0094-8
52. Martínez WA, Melo CE, Melo OO (2017) Median Polish Kriging for space–time analysis of precipitation. Spat Stat. https://doi.org/10.1016/j.spasta.2016.10.003
53. Behm S, Haupt H, Schmid A (2018) Spatial detrending revisited: modelling local trend patterns in NO_2-concentration in Belgium and Germany. Spat Stat. https://doi.org/10.1016/j.spasta.2018.04.004
54. Hristopulos DT, Agou VD (2019) Stochastic local interaction model with sparse precision matrix for space–time interpolation. Spat Stat. https://doi.org/10.1016/j.spasta.2019.100403
55. Zammit-Mangion A, Wikle CK (2020) Deep integro-difference equation models for spatio-temporal forecasting. Spat Stat. https://doi.org/10.1016/j.spasta.2020.100408
56. Bloomfield P (1976) Fourier analysis of time series: an introduction. Willey, New York

57. Shumway RH, Stoffer DS (2011) Time series analysis and its application: with R examples, 3rd edn. Springer Science+Business Media, New York
58. Cowpertwait PSP, Metcalfe AV (2009) Introductory time series with R. Springer Science+Business Media, New York
59. Brockwell PJ, Davis RA (2002) Introduction to Time Series and Forecasting, 2nd edn. Springer, New York
60. Kedem B, Fokianos K (2002) Regression models for time series analysis. Wiley, Hoboken
61. Li ZL, Tang BH, Wu H, Ren H, Yan G, Wan Z, Trigo IF, Sobrino JA (2013) Satellite-derived land surface temperature: current status and perspectives. Remote Sens Environ. https://doi.org/10.1016/j.rse.2012.12.008
62. Hooker J, Duveiller G, Cescatti A (2018) A global dataset of air temperature derived from satellite remote sensing and weather stations. Sci Data. https://doi.org/10.1038/sdata.2018.246
63. Vauclin M, Vieira SR, Bernard R, Hatfield JL (1982) Spatial variability of surface temperature along two transects of a bare soil. Water Resour Res. https://doi.org/10.1029/WR018i006p01677
64. Prata AJ, Caselles V, Coll C, Sobrino JA, Ottlé C (1995) Thermal remote sensing of land surface temperature from satellites: current status and future prospects. Remote Sens Rev. https://doi.org/10.1080/02757259509532285
65. Zakšek K, Oštir K (2012) Downscaling land surface temperature for urban heat island diurnal cycle analysis. Remote Sens Environ. https://doi.org/10.1016/j.rse.2011.05.027
66. Sousa D, Small C (2019) Mapping and monitoring rice agriculture with multisensor temporal mixture models. Remote Sens. https://doi.org/10.3390/rs11020181
67. Magarreiro C, Gouveia CM, Barroso CM, Trigo IF (2019) Modelling of wine production using land surface temperature and FAPAR—the case of the Douro Wine Region. Remote Sens. https://doi.org/10.3390/rs11060604
68. Bonafoni S, Keeratikasikorn C (2018) Land surface temperature and urban density: multiyear modeling and relationship analysis using MODIS and landsat data. Remote Sens. https://doi.org/10.3390/rs10091471
69. Pavlidou E, Van der Meijde M, Van der Werff H, Hecker C (2019) Time series analysis of land surface temperatures in 20 earthquake cases worldwide. Remote Sens. https://doi.org/10.3390/rs11010061
70. Marchand N, Royer A, Krinner G, Roy A, Langlois A, Vargel C (2018) Snow-covered soil temperature retrieval in Canadian arctic permafrost areas using a land surface scheme informed with satellite remote sensing data. Remote Sens 1:11. https://doi.org/10.3390/rs10111703
71. Liu T, Yu L, Bu K, Yan F, Zhang S (2018) Seasonal local temperature responses to paddy field expansion from rain-fed farmland in the cold and Humid Sanjiang Plain of China. Remote Sens. https://doi.org/10.3390/rs10122009
72. Kestens Y, Brand A, Fournier M, Goudreau S, Kosatsky T, Maloley M, Smargiassi A (2011) Modelling the variation of land surface temperature as determinant of risk of heat-related health events. Int J Health Geogr. https://doi.org/10.1186/1476-072X-10-7
73. Julien Y, Sobrino JA, Mattar C, Ruescas AB, Jimenez-Munoz JC, Soria G, Hidalgo V, Atitar M, Franch B, Cuenca J (2011) Temporal analysis of normalized difference vegetation index (NDVI) and land surface temperature (LST) parameters to detect changes in the Iberian land cover between 1981 and 2001. Int J Remote Sens. https://doi.org/10.1080/01431161003762363
74. Romaguera M, Vaughan RG, Ettema J, Izquierdo-Verdiguier E, Hecker CA, van der Meer FD (2018) Detecting geothermal anomalies and evaluating LST geothermal component by combining thermal remote sensing time series and land surface model data. Remote Sens Environ. https://doi.org/10.1016/j.rse.2017.10.003
75. Alibakhshi Z, Ahmadi M, Asl MF (2020) Modeling biophysical variables and land surface temperature using the GWR model: case study—Tehran and its satellite cities. J Indian Soc Remote Sens. https://doi.org/10.1007/s12524-019-01062-x
76. Bonino MF, Azócar DLM, Schulte JA, Cruz FB (2015) Climate change and lizards: changing species' geographic ranges in Patagonia. Reg Environ Change. https://doi.org/10.1007/s10113-014-0693-x
77. Báez VH, Aigo JC, Cussac VE (2011) Climate change and fish culture in Patagonia: present situation and perspectives. Aquac Res. https://doi.org/10.1111/j.1365-2109.2011.02804.x
78. Wan Z, Hook S, Hulley G (2015) MOD11C3 MODIS/terra land surface temperature/emissivity monthly L3 Global 0.05Deg CMG V006. NASA EOSDIS Land Process DAAC. https://doi.org/10.5067/MODIS/MOD11C3.006
79. R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/. Accessed 10 Apr 2018
80. Ribeiro Jr PJ, Diggle PJ (2016) geoR: Analysis of geostatistical data. R package version 1.7–5.2. https://CRAN.R-project.org/package=geoR. Accessed 1 July 2019
81. Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. Comput Geosci. https://doi.org/10.1016/j.cageo.2004.03.012
82. Gräler B, Pebesma E, Heuvelink GBM (2016) Spatio-temporal interpolation using gstat. The R Journal. https://doi.org/10.32614/RJ-2016-014
83. Pebesma E (2012) spacetime: spatio-temporal data in R. J Stat Softw. https://doi.org/10.18637/jss.v051.i07
84. Bivand RS, Pebesma E, Gomez-Rubio V (2013) Applied spatial data analysis with R, 2nd edn. Springer, New York
85. De Iaco S, Cappello C, Posa D (2018) covatest: tests on properties of space-time covariance functions. R package version 1.0.0. https://CRAN.R-project.org/package=covatest. Accessed 1 July 2019
86. Hijmans, RJ (2017) raster: Geographic data analysis and modeling. R package version 2.6–7. https://CRAN.R-project.org/package=raster. Accessed 1 July 2019
87. GSOD (2018) Climate data online (CDO). https://www7.ncdc.noaa.gov/CDO/cdoselect.cmd?datasetabbv=GSOD&countryabbv=&georegionabbv. Accessed 10 February 2018
88. Global Administrative Areas (2018) GADM database of Global Administrative Areas (version 2.0). www.gadm.org. Accessed 16 July 2018
89. Hyndman RJ, Koehler AB (2006) Another look at measures of forecast accuracy. Int J Forecast. https://doi.org/10.1016/j.ijforecast.2006.03.001
90. De Iaco S, Palma M, Posa D (2016) A general procedure for selecting a class of fully symmetric space–time covariance functions. Envirometrics. https://doi.org/10.1002/env.2392
91. Cappello C, De Iaco S, Posa D (2018) Testing the type of non-separability and some classes of space–time covariance function models. Stoch Env Res Risk Assess. https://doi.org/10.1007/s00477-017-1472-2
92. Li B, Genton MG, Sherman M (2007) A nonparametric assessment of properties of space–time covariance functions. J Am Stat Assoc. https://doi.org/10.1198/016214507000000202
93. Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York
94. Goovaerts P (1997) Geostatistics for national resources evaluation. Oxford University Press, New York
95. Armstrong M (1998) Basic linear geostatistics. Springer, Berlin

96. Hyndman RJ, Khandakar Y (2008) Automatic time series forecasting: the forecast package for R. J Stat Softw 26:1–22
97. Hyndman R (2017) 'forecast': Forecasting functions for time series and linear models. R package (version 8.2). https://pkg.robjhyndman.com/forecast Accessed 14 May 2018
98. Box GEP, Jenkins GM (1970) Time series analysis: forecasting and control. Holden-day, San Francisco
99. Newbold P, Agiakloglou C, Miller J (1993) Long-term inference based on short-term forecasting models. In: Rao TS (ed) Developments in time series analysis. Chapman & Hall, London, pp 9–25
100. Janacek G (2001) Practical time series. Hodder Education Publishers, London
101. Chatfield C (2003) The analysis of time series: an introduction. CRC Press Company, Boca Raton
102. Prado R, West M (2010) Time series: modeling, computation, and inference. CRC Press, Boca Raton
103. Brockwell PJ, Davis RA (2010) Introduction to time series and forecasting, 2nd edn. Springer, New York
104. Woodward WA, Gray HL, Elliott AC (2011) Applied time series analysis. CRC Press, Boca Raton
105. Box GEP, Jenkins GM, Reinsel GC, Ljung GM (2015) Time series analysis: forecasting and control, 5th edn. Wiley, Hoboken