RESEARCH ARTICLE

CrossMark

Minimizing the uncertainties of RCMs climate data by using spatio-temporal geostatistical modeling

Venetsanou P.[1] · Anagnostopoulou C.[2] · Loukas A.[3,4] · Lazoglou G.[2] · Voudouris K.[1]

Abstract

The spatio-temporal Kriging approach by using five different covariance models, has been applied into Regional Climate Model (RCM) simulated precipitation and temperature dataset in a coastal area. The results of the spatio-temporal technique were evaluated against the ERA-Interim reanalysis data during the period from 1981 to 2000. The reliability of the spatio-temporal interpolation results were estimated by using both the judgment of the wireframe plots, between the sample and the fitted covariance models, and the statistic metrics. Thus, Taylor diagrams were used and the Mean Square Error (MSE) was calculated. The analysis demonstrates that the sum-metric covariance model is highly superior to the other four covariance models as it is closer to the reanalysis data, having the highest correlation coefficient, as well as, the smallest standard deviation, resulting in the smallest Root Mean Square Error. The spatio-temporal interpolation approach improved the MPI and HadGEM2 climate model dataset. The largest enhancement is pointed out in the interpolated RCM precipitation during winter and autumn. Concerning the temperature, the interpolated MPI temperature data is negligibly improved, whereas the interpolated HadGEM2 temperature is particularly optimized during winter and autumn. The spatio-temporal interpolation technique led to the minimization of the uncertainties of the Regional Climate Models, (RCMs) simulations, and also to the best agreement between them and the ERA-Interim reanalysis data during the period from 1981 to 2000. Nevertheless, the MPI climate model is more reasonable compared to the HADGEM2 for the research area.

Keywords Spatial-temporal kriging method · Spatio-temporal variogram · Sum-metric covariance model · Taylor diagrams · Coastal area

Introduction

Observed or Model estimated climate data is well established and widely used in many scientific fields such as Hydrology, Hydrogeology, Climatology or Agrometeorology. They

Communicated by: H. A. Babaie

✉ Venetsanou P.
  pvenetsn@geo.auth.gr

[1] Department of Geology, Laboratory of Engineering Geology and Hydrogeology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

[2] Department of Geology, Laboratory of Meteorology and Climatology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

[3] Department of Civil Engineering, University of Thessaly, Volos, Greece

[4] Department of Rural and Surveying Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

contribute to modeling purposes for integrated and sustainable natural resources management and adaption in a continuously changing climate. Therefore, high resolution of spatial and temporal climate data is required to be used as input data in climate change impact studies assessing the future climate conditions.

The most widespread and essential tools for understanding, quantifying, and evaluating both the future climate change and constructing climate scenarios are the General Circulation Models (GCMs). Owing to the coarse spatial resolution of the GCMs, the Regional Climate Models (RCMs) have not only been developed in order to represent the local current climate conditions but also estimate more accurately the future climate (IPCC 2007, 2013). Moreover, climate data for climate change impact studies can also be estimated by spatio-temporal interpolation methods. The desired spatio-temporal resolution depends on the research area and the data (Kilibarda et al. 2014; Hengl et al. 2012; Gething et al. 2007).

Recently, various spatial interpolation techniques have considerably been applied into geosciences. These approaches can be divided into three major categories based on both the

interpolation methods and scales of the implementation. The first category involves simple approaches such as Thiessen polygons (Brassel and Douglas 1979), Nearest Neighbor using Thiessen or Voronoi polygons (Sibson 1981), Spline (Hutchinson 1988), Inverse Distance Weighting (Zimmerman et al. 1999) and different types of the well-known Kriging method, for instance, Ordinary Kriging, Universal Kriging, Co-Kriging (Matheron 1962; Cressie 1993; Wackernagel 2003). The second one comprises more complex interpolation techniques, for instance artificial neural networks and fuzzy reasoning method (Friedman 1994; Huang et al. 1998; Wong et al. 2003). Auxiliary data such as satellite imagery and digital elevation models (DEM) are also helpful tools for the interpolation process.

Several studies compare and evaluate different spatial interpolation techniques, indicating the advantages and drawbacks of the above methods and their applications. Yang et al. (2015) compared four spatial interpolation techniques (ANUDEM, Spline, IDW and Kriging) by using rainfall data from regional climate models, in order to estimate daily rainfall data for the future periods. Based on their results, the Inverse Distance Weighting (IDW) method produced slightly accurate predictions than the other three methods. Mair and Fares (2011) point out that the Thiessen polygon is the least promised method, whereas the simple Kriging with varying local means produced smaller error for the interpolation in a mountainous region of an island. Hofstra et al. (2008) indicate that the global Kriging is slightly the best method for climate station data compared to the other methods over Europe from 1961 to 1990.

Nowadays, spatio-temporal techniques have been introduced so as to bridge the spatial gaps in the different time series data making rapid and notable progress. Lasinio et al. (2007) described the spatio-temporal analysis as a prediction of time development of a variable in a given spatial domain. Hence, the spatio-temporal procedures are comprehended more accurately and completely (Heuvelink et al. 2010; Heuvelink et al. 2012; Gräler et al. 2016). In particular, Gräler et al. (2016) applied various spatio-temporal covariance models into the daily rural air quality measurements in Germany. Heuvelink et al. (2012) generated daily temperature maps (1 km × 1 km resolution) by using the time-series MODIS MOD11A2 product Land Surface Temperature images which are publicly available. The observed data is derived from the national network of the meteorological stations in Croatia. The sum-metric and separable covariance models were both used for the spatio-temporal auto-correlation. The results indicate that the application of the spatio-temporal regression-Kriging and the incorporation of time-series of the remote sensing images lead to more accurate temperature maps.

This paper deals with the spatio-temporal Kriging approach in order to quantify the uncertainties and increase the resolution of the RCMs climate data. Hence, different spatio-temporal covariance models were compared and evaluated

so as to estimate the RCM-simulated temperature and precipitation dataset during the period from 1981 to 2000. The ultimate task of this methodology is the adjustment between the climate models data and the observed one.

The present paper is structured as follows: In Section "Data" is demonstrated the description of the research area and the climate data. In Section "Methodology" the methodology about the spatio-temporal Kriging approach and the spatio-temporal covariance models is described. In Section "Results" the results are evaluating by using statistics metrics and discussed ending up which one of the spatio-temporal covariance model represents the climate of the research area more effectively. Finally, the conclusions are drawn in Section "Discussion and Conclusions".

Data

Study area

The research takes place in a small coastal area which is situated in northern Greece, Halkidiki (Fig. 1), extending to the coastal zone of Kassandra Gulf, being part of the Havrias river catchment. It can be characterized as an agricultural and touristic center. According to the GIS analysis, the object of the study, covers an area of 40 km^2. The minimum and the maximum elevation range from 0 to 235 m (mean value 140 m) above the mean sea level, respectively. The morphology is complex due to the transition from the land to the sea as it is depicted in Fig. 1.

Based on Köppen (1954) classification, the climate of the area is typical of Mediterranean (Csa), with warm summers and mild winters, which is highly representative of the Greek climate. The main feature of the area, is the uneven variation of the precipitation throughout the year, with the lowest precipitation occurring in summertime, whereas the highest during winter and autumn. The dry and warm summers in combination with the increased water needs for irrigation and domestic purposes affect the water resources availability rendering the region vulnerable to the anticipated climate change.

Climate data

Observation: Reanalysis data

In the framework of this research, climate data without spatial or temporal gaps is required because filling them in climate data and, in particular in precipitation underlies high uncertainties and risks. Thus, in this investigation, the temperature and precipitation dataset are derived from the ERA-Interim reanalysis data. ERA-Interim is one of the most recent global atmospheric reanalysis database, which is offered by the European Centre for Medium-Range Weather Forecasts
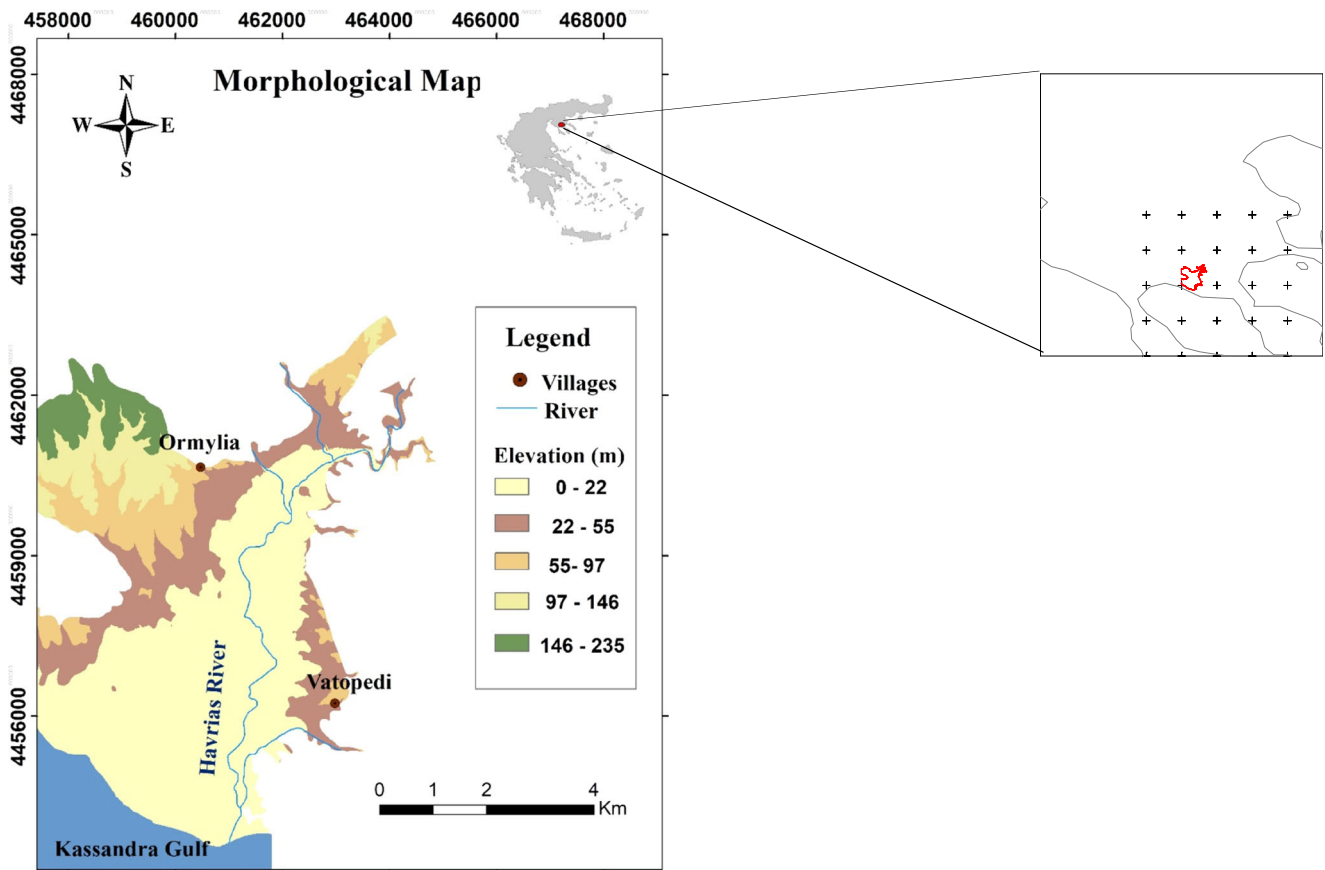
Fig. 1 Morphological map of the research area and the location of the ERA-Interim grid points (cross)

(ECMWF) (http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc). It involves various climate parameters such as precipitation, temperature, wind, radiation, dew point, evaporation, and beginning from 1979 until now (Dee et al. 2011). The provided data concerning this research covers the period from 1981 to 2000 at a daily time step with a spatial resolution 12.5 km × 12.5 km (Fig. 1). The selected grid points include the study area (40 km^2).

Climate models data

The dynamical downscaling was carried out by using the Regional Climate Model version 4 (RegCM4) in the framework of Med-Cordex and Coupled Models Intercomparison Project Phase 5 (CMIP5) programs (www.medcordex.eu). The Regional Climate Model version 4 (RegCM4) is the latest version of the RegCM regional climate model and constitutes an evolution and improvement of its previous version RegCM3 which is demonstrated by Pal et al. (2007). Giorgi et al. (2012) described the Regional Climate Model version 4 (RegCM4) as "a hydrostatic, compressible, sigma-p vertical coordinate model run on an Arakawa B-grid in which wind and thermodynamical variables are horizontally staggered". The spatial resolution of RegCM4 Regional Climate Model is 12 km × 12 km.

In this paper, the Hadley Global Environment Model 2 (HadGME2 –ES) and the MPI Earth System Model running on mixed resolution grid (MPI-ESM-MR), general circulation models (GCMs) with spatial resolution 50 km × 50 km were used as forcing data to the RegCM4 climate model. The general circulation models and the details about their characteristics are figured out in Table 1.

Methodology

Spatio-temporal interpolation

In this research, the spatio-temporal Kriging was implemented in order to provide the same spatial resolution of daily climate model data with the ERA-Interim for the period from 1981 to 2000. For this reason, different types of spatio-temporal covariance models were applied investigating the spatio-temporal interpolation approach and finally ending up the best performance.

The separable, product-sum, metric, sum-metric and simple sum-metric covariance models were tested. The general spatio-temporal covariance function is described in detail by Cressie and Wikle (1998, 2011) and given by

$$C_{st}(\mathrm{h, u}) = \mathrm{Cov}\big(Z(s, t), Z(\acute{s}, \acute{t})\big) \qquad (1)$$

Table 1 The general circulation models (GCMs) and the details about their characteristics

	General Circulation Models (GCMs)	
Model Acronyms	HadGEM2 ES	MPI-ESM-MR
Project	Coupled Models Intercomparison Project Phase 5 (CMIP5)	Coupled Models Intercomparison Project Phase 5 (CMIP5)
Institute	Met Office Hadley Centre (MOHC)	Max Planck Institute for Meteorology (MPI-M)
Type	Coupled Earth System Model	Coupled Earth System Model
Funder	Met Office Handley Centre	Bundesministerium fuer Bildung und Forschung
Main References	Collins et al. 2011, Jones et al. 2011	Popke et al. 2013, Giorgetta et al. 2013
Resolution	50 km × 50 km	50 km × 50 km

where Z is the random function Z = Z(s, t).and the spatio-temporal variogram is described by

$$\gamma_{st} = \text{Cst}(0,0) - \text{Cst}(h,u) \tag{2}$$

The Eq. (1) is referred to a separating spatial distance h and temporal distance u and any pair of points (s, t), (ṡ, ṫ) ∈ S × T with ‖s- ṡ‖ = h and |t- ṫ| = u.

The mathematical formula which describes the separable covariance model (Gräler et al. 2016; Pebesma and Gräler 2017) is given by

$$C_{sep}(h,u) = C_s(h) \times C_t(u) \tag{3}$$

The corresponding variogram is given by

$$\gamma_{sep}(h,u) = sill \times \left(\gamma_s(h) + \gamma_t(u) - \gamma_s(h) \times \gamma_t(u) \right) \tag{4}$$

where γ_s, γ_t are standardised spatial and temporal variograms with separate nugget effects and (joint) sill of 1.

The product-sum covariance function was firstly described by De Cesare et al. (2001) and De Iaco et al. (2001). A slightly modified function for the product-sum covariance model was announced by Gräler et al. (2016) and Pebesma and Gräler (2017) and given by

$$C_{ps}(h,u) = C_s(h) + C_t(u) + k \times C_S(h) \times C_t(u) \tag{5}$$

with k > 0.

Its variogram can be written by

$$\gamma_{ps}(h,u) = (k \times sill_t + 1) \times \gamma_s(h) + (k \times sill_s + 1) \\ \times \gamma_t(u) - k_{\gamma s}(h) \times \gamma_t(u) \tag{6}$$

where γ_s, γ_t are spatial and temporal variograms and the overall sill (sill$_{st}$) is described by

$$sill_{st} = k*sill_s*sill_t + sill_s + sill_t \tag{7}$$

The metric covariance function is described by

$$C_m = C_j\left(\sqrt{h^2 + (k \times u)^2} \right) \tag{8}$$

Its variogram is given by

$$\gamma_m(h,u) = \gamma_j\left(\sqrt{h^2 + (k \times u)^2} \right) \tag{9}$$

where γ_j is any known variogram involving some nugget-effect.

The sum-metric covariance model is a couple of spatial, temporal and a metric model containing an anisotropy parameter k (Bilonick 1988; Snepvangers et al. 2003) is:

$$C_{sm}(h,u) = C_s(h) + C_t(u) + C_j\left(\sqrt{h^2 + (k \times u)^2} \right) \tag{10}$$

Its variogram is represented by

$$\gamma_{sm}(h,u) = \gamma_s(h) + \gamma_t(u) + \gamma_j\left(\sqrt{h^2 + (k \times u)^2} \right) \tag{11}$$

where γ_s, γ_t and γ_j are spatial, temporal and joint variograms with a separate nugget-effect.

The simple sum-metric covariance model constitutes a simplified version of the sum-metric model and the corresponding variogram is written by

$$C_{ssm}(h,u) = nug \times 1_{h>0\lor u>0} + \gamma_s(h) + \gamma_t(u) \\ + \gamma_j\left(\sqrt{h^2 + (k \times u)^2} \right) \tag{12}$$

The application of the spatio-temporal Kriging was carried out in the R environment for statistical computing (R Development Core Team 2013; Pebesma 2004) by using the gstat, sp., spacetime, raster, rgdal, rgeos and xts packages (Pebesma 2012; Gräler et al. 2016 and Pebesma and Gräler 2017). The krigeST function was used for the spatio-temporal Kriging implementation.

The sample variogram for each climate models data, namely for MPI and HadGME2 precipitation and temperature dataset with a spatial resolution of 50 km, was modeled and utilized as an input for the fitting routines of the different models. The sample variogram was estimated through the function variogram (vgmST), with spatial lags of 0.2 degrees

and time lags of one day. The WGS84 coordinate reference system was used.

The Linear, Spherical, the Expontential and Matèrn model were employed to find the best fit model to the sample variogram, testing the whole possible combinations of these models. For each variogram the space and the time parameters remain stable for the two climate models, as well as for the two parameters (precipitation and temperature). On the contrary, the values of the sill, the kappa and the anisotropy (stAni) parameters depend on both the covariance model and the climate parameter.

Evaluation

The observed (ERA-Interim) and the interpolated MPI and HadGMEM2 climate models precipitation (mm) and temperature (°C) dataset are analyzed and evaluated for assessing the biases between the two databases. For the evaluation of the RCM simulated data against the observed data, the centered RMS difference (E'), the Root Mean Square Error (RMSE), the correlation coefficient, and the standard deviation were used. Taylor diagrams were plotted in order to quantify the uncertainties, compare the five covariance models, judge the best performance of the covariance models variograms in relation to the sample variogram, and finally estimate their reliability. Taylor diagrams (Taylor 2001) provide a comprehensive statistical summary of how closely a dataset approaches the observed data in relation to their correlation, their root-mean-square difference, and their standard deviations. These diagrams are valuable for evaluating various aspects of complex models and for gauging the relative skill of many different models. The standard deviation is given by the radial distances from the reference point to sample and the azimuthal positions describing the correlation coefficient between the two fields. The Root Mean Square Error (RMSE) is described by the semicircles (Taylor 2001).

The correlation coefficient (R), the centered RMS difference (E'), the standard deviation and of the interpolated climate models data (σ_{xmodel}) and the observed data (σ_{xobs}) are calculated by:

$$R = \frac{\frac{1}{N}\sum_{n=1}^{N}\left(x_{model,n}-x_{model}\right)\left(x_{obs,n}-x_{obs}\right)}{\sigma_{xmodel}\sigma_{xobs}} \tag{13}$$

$$E'^2 = \frac{1}{N}\sum_{n=1}^{N}\left[\left(x_{model,n}-x_{model}\right)\times\left(x_{obs,n}-x_{obs}\right)\right]^2 \tag{14}$$

$$\sigma_{xmodel}^2 = \frac{1}{N}\sum_{n=1}^{N}\left(x_{model,n}-x_{model}\right)^2 \tag{15}$$

$$\sigma_{xobs}^2 = \frac{1}{N}\sum_{n=1}^{N}\left(x_{obs,n}-x_{obs}\right)^2 \tag{16}$$

The RMSE is defined by

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(x_{obs,i}-x_{model,i}\right)^2}{n}} \tag{17}$$

where x_{obs} are observed values, namely ERA-Interim precipitation (mm) and temperature (°C) and x_{model} is the interpolated MPI and HadGME2 precipitation (mm) and temperature (°C) data at time/place i.

However, the final choice of the suitable spatio-temporal covariance model should be validated using both the statistics metrics and the judgment of the best performance between them.

Results

Historical MPI and HadGEM2 temperature and precipitation dataset were evaluated against the ERA-Interim reanalysis data, for the period from 1981 to 2000. The largest differences are observed in the HadGEM2 simulations, for both temperature and precipitation data. On the contrary, the discrepancies between the reanalysis and MPI temperature data are negligible. The biases between the climate models data and the ERA-Interim dataset, particularly in precipitation, might be attributed to the complex morphology of the research area. The transition from the land to the sea influences on the climate models signal reducing their ability to simulate effectively the climate parameters, specifically the precipitation (Xoplaki et al. 2004; Tolika et al. 2006).

Analytically, the climate models present wetter climate conditions in comparison to the reanalysis data from 1981 to 2000, without changing the current climate conditions of the research area. According to the Era-Interim reanalysis data, the mean annual rainfall over the area is 510.2 mm (Fig. 2) from 1981 to 2000, while the mean annual precipitation is about 1270 mm for the HadGEM2 simulations and 839 mm for the MPI simulations (not shown). No significant change has been recorded in the precipitation data, from 1981 to 2000 (Fig. 2). According to the reanalysis data, the maximum monthly rainfall is recorded during November and December, while August is the month when the lowest rainfall is presented which are also observed in the MPI and HADGEM2 climate models.

On the other hand, the two RCM-simulated temperatures are much closer to the reanalysis temperatures. The mean annual temperature of the ERA Interim data is 15.6 °C (Fig. 3), the corresponding temperature of the HadGEM2 simulations is 17 °C and 15.9 °C of the MPI simulations (not shown). July and August are the warmest months for both datasets (reanalysis and models), whereas January is the coldest month (Fig. 3). Furthermore, an upward tendency is
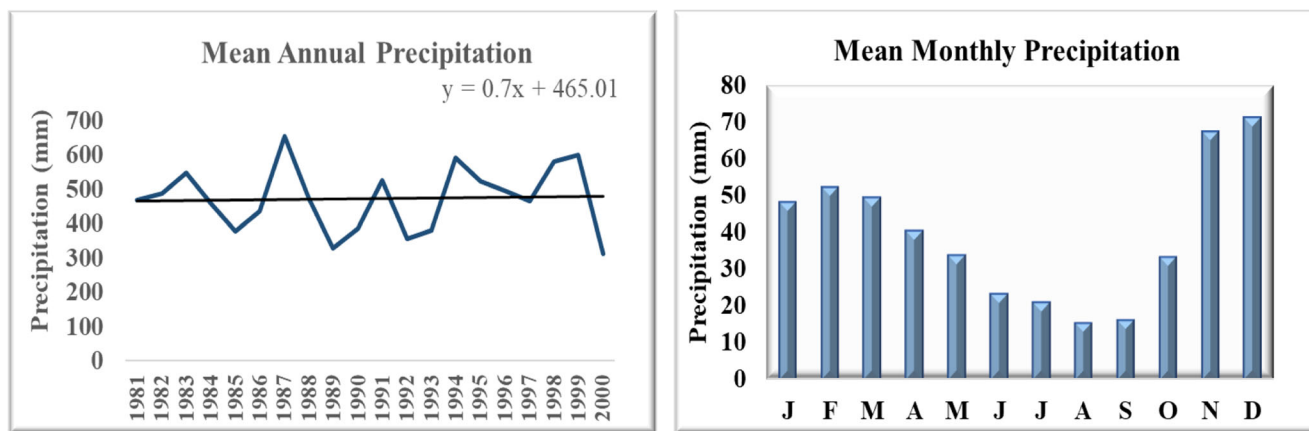
Fig. 2 The mean annual and monthly ERA-Interim precipitation for the research area over the period 1981 to 2000

observed in the reanalysis temperature (Fig. 3) that it is also detected in the RCM-simulated temperatures (not shown).

Evaluation of the Spatio -temporal interpolation models

The separable, metric, product-sum sum-metric and simple-sum-metric covariance models were analyzed and evaluated for the selection of the suitable model for the MPI and HadGEM2 temperature and precipitation dataset. Initially, the the sample variograms between the MPI and HadGEM2 models were compared. The juxtaposition between them showed that HadGME2 sample variograms of both precipitation and temperature dataset are overestimated.

Figures 4 to 5 depict the sample variogram for each climate parameter of the two RCM simulated models (MPI and HadGEM2) in comparison to the spatio-temporal variograms of the five covariance models for a wet-cold period (from October to April). It can be derived from the wireframe plots of the precipitation (Figs. 4(a) to 4(b)) that the least suitable performances are exhibited by the separable and metric co-variance models, while the sum-metric seems to perform

reasonably well for both MPI and HadGEM2 climate data. The other two covariance models seem to respond more satisfactorily in relation to the sample. Concerning the wireframe of temperatures for both RCMs (Figs. 5(a) to 5(b)), the simple-sum-metric model and the metric covariance model present the most apparent differences in relation to the sample variogram.

Since the choice of the suitable covariance model based only on the wireframe plots is quite subjective, the statistical evaluation of the five spatio-temporal covariance models for each MPI and HadGEM2 precipitation and temperature dataset were done by using Taylor diagrams (Figs. 6 and 7).

According to Taylor diagram, for the precipitation dataset (Fig. 6), the worst performances are observed in the metric (blue) and simple sum-metric (yellow) models for both RCM models. Even though, the correlation coefficient of the MPI for these two covariance models is higher than 0.6, the significant high RMSE and STDEV values indicate the unsuitable character of these models. The lowest correlation coefficient of the separable model (black) as well as the relatively high value of the STDEV of the product-sum model (red), demonstrate that these models cannot be considered proper for the
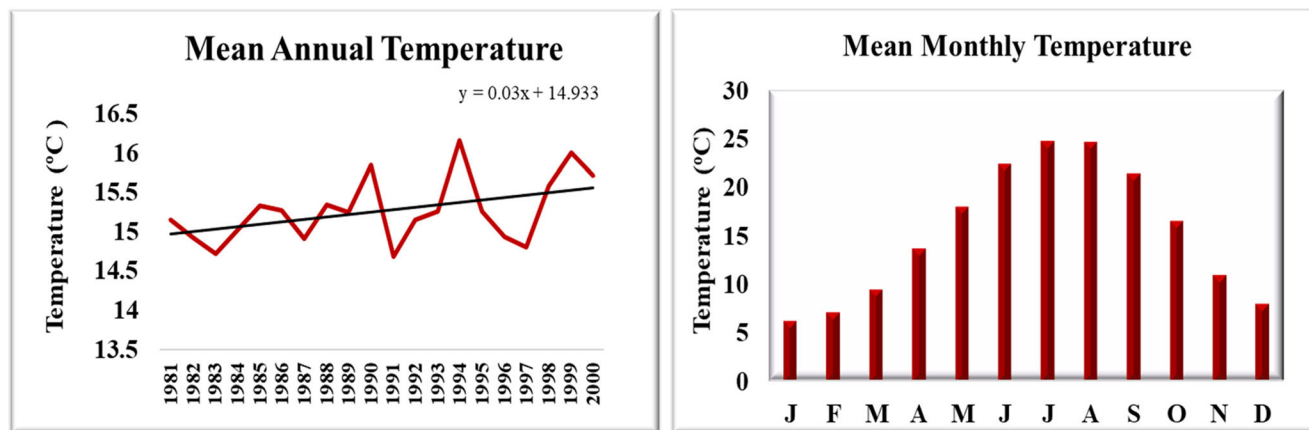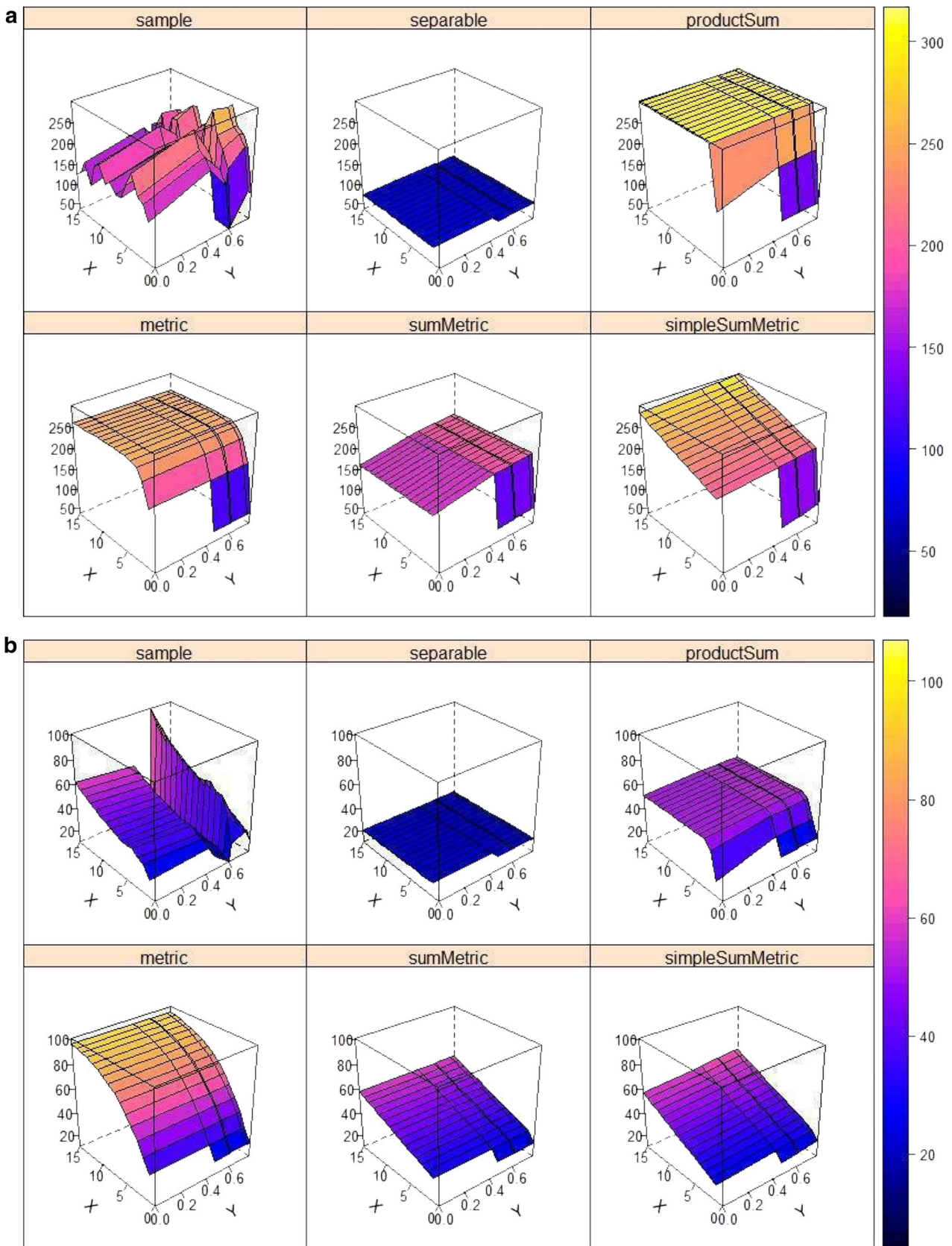

Fig. 3 The mean annual and monthly ERA-Interim temperature for the research area over the period 1981 to 2000

Fig. 4 **a** The wireframe plots of the sample variogram and the five fitted covariance models variograms for HadGEM2 precipitation data in wet - cold period, where X= time in days and Y= distance in degrees. **b** The wireframe plots of the sample variogram and the five fitted covariance models variograms for MPI precipitation data in wet - cold period, where X= time in days and Y= distance in degrees
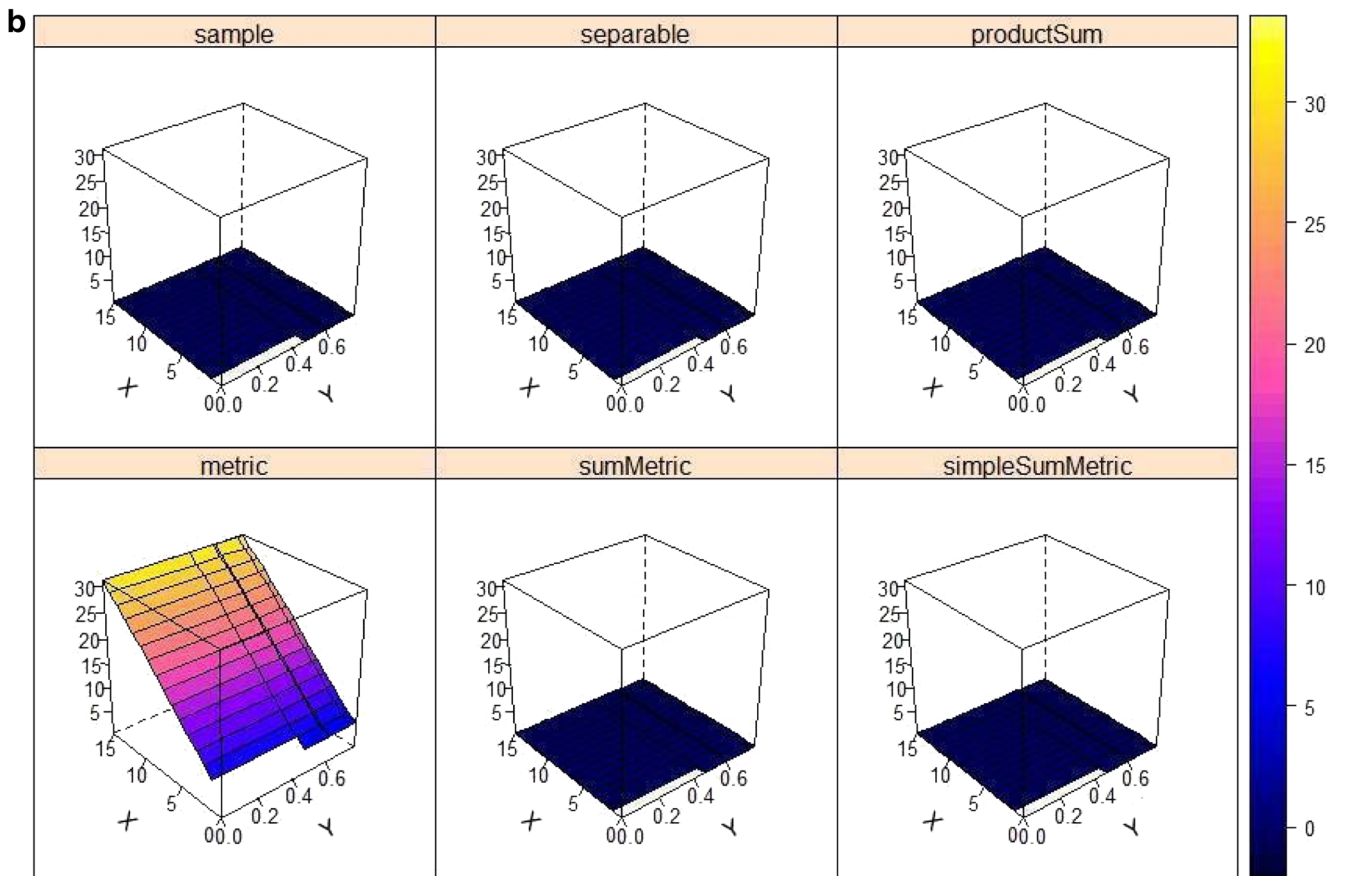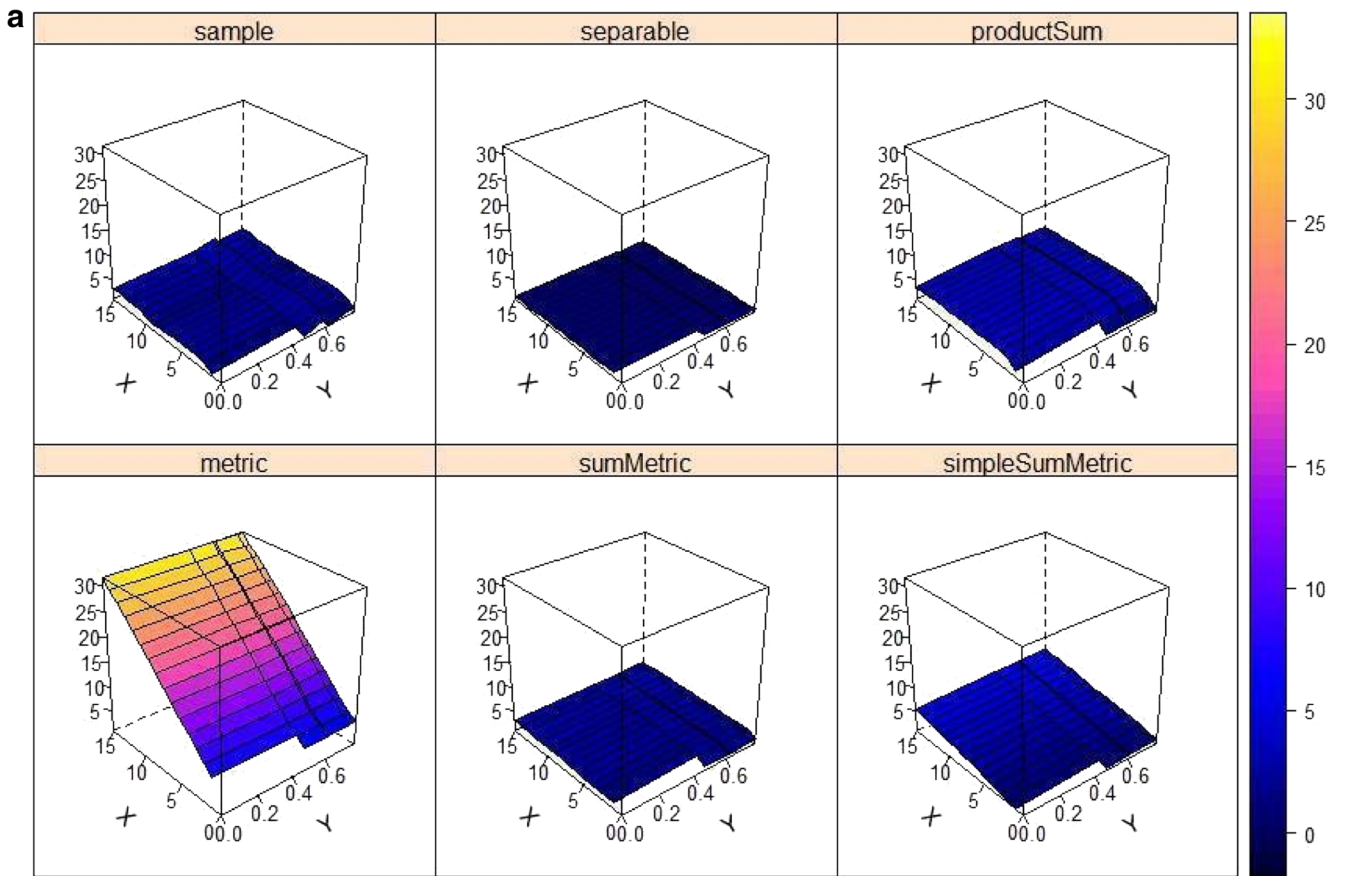
◀ **Fig. 5** **a** The wireframe plots of the sample variogram and the five fitted covariance models variograms for HadGEM2 temperature data in wet-cold period, where X=time in days and Y=distance in degrees. **b** The wireframe plots of the sample variogram and the five fitted covariance models variograms for MPI temperature data in wet-cold period, where X= time in days and Y= distance in degrees

reproduction of the precipitation. Overall, the sum-metric model is close to the observed precipitation data, having the highest correlation coefficient (0.7 in HadGEM2 and 0.8 in MPI) in both RCM climate models (Fig. 6). It is also pointed out that, the sum-metric model displays the smallest standard deviation in both climate models, resulting in the smallest RMS error in comparison to the other four covariance models.

As regards to the temperature, the results are more satisfactory compared to the precipitation (Fig. 7). Specifically, the covariance models display very lower RMSE and STDEV than the precipitation. It is also worth mentioning that no one of the covariance model outperforms the boundaries of the diagram. In particular, the separable model (black) presents a negative correlation coefficient. The largest RMSE and STDEV values are exhibited and in this circumstance, by the metric (blue) and the simple sum-metric (yellow) models. On

the contrary, the product-sum and sum-metric covariance models display more reasonable results than the other models. However, the correlation coefficient of the sum-metric model is slightly superior to the product-sum model.

For the comprehensive statistical evaluation, the Mean Square Error (MSE) was calculated. Table 2 displays the results regarding the MSE for each general circulation climate model (GCMs) and the corresponding spatio-temporal variogram models. The MSE should display the lowest error in order for a covariance model to be considered appropriate. As it can be derived from the Table 2, the lowest error is presented by the sum-metric model, while the largest error is exhibited by the metric model in both MPI and HADGEM2 climate models data.

The evaluation and combination of the spatio-temporal variograms results and the statistics metrics of the five covariance models indicate that the least satisfactory results are presented by the metric covariance model, while the best model is the sum-metric. In conclusion, the sum-metric is highly superior to the other four spatio-temporal covariance models, for both MPI and HadGEM2 climate data. After the identification of the suitable covariance model, the daily MPI and HadGEM2 results were interpolated for the period from
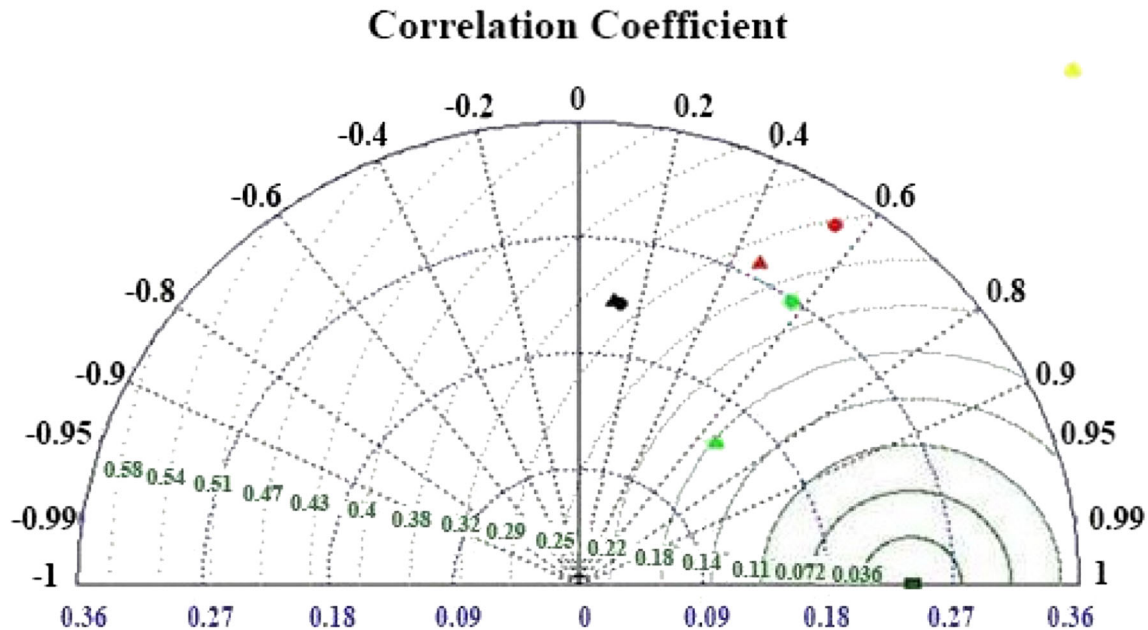
Fig. 6 Taylor diagram displaying a statistical comparison between MPI (triangle) and HADGEM2 (circle) climate precipitation data and separable (black), product-sum (red), metric (blue), sum-metric (green) and simple sum-metric (yellow) covariance models. Black lines indicate correlation coefficient, green lines indicate the centered RMS difference and blue line indicates standard deviation
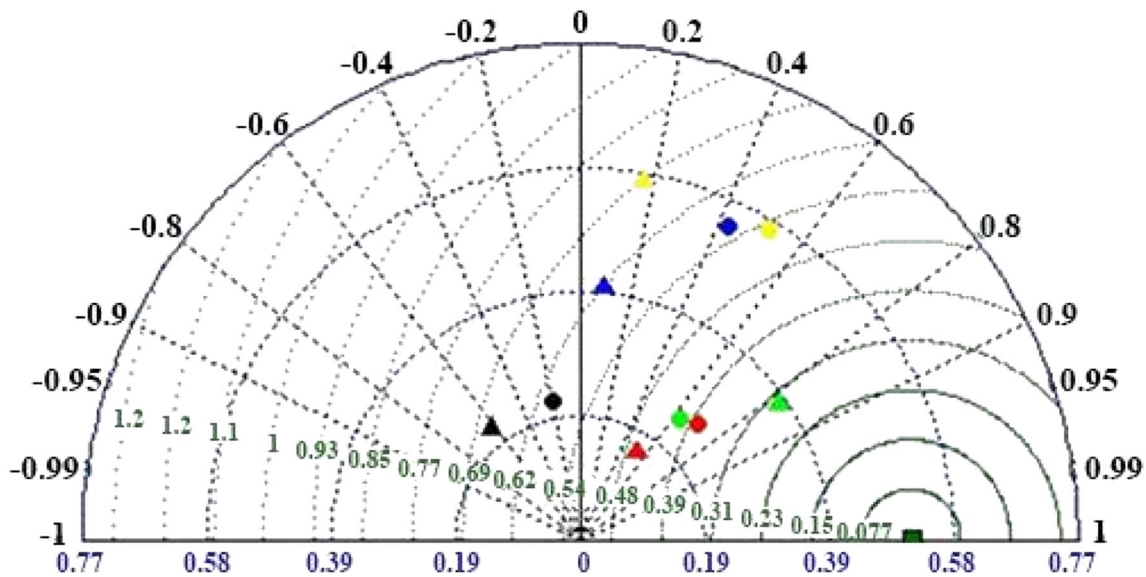
Fig. 7 Taylor diagram displaying a statistical comparison between MPI (triangle) and HADGEM2 (circle) climate temperature data and separable (black), product-sum (red), metric (blue), sum-metric (green) and simple sum-metric (yellow) covariance models. Black lines indicate correlation coefficient, green lines indicate the centered RMS difference and blue lines indicate standard deviation

1981 to 2000 by using the sum-metric covariance model. The spatial resolution of the interpolated climate data was defined at 12.5 km × 12.5 km which corresponds to the spatial resolution of the ERA-Interim reanalysis data.

## Spatio-temporal interpolated climate data

Applying the spatio-temporal Kriging by using the sum-metric model, the MPI and HadGEM2 climate data results are improved for the period from 1981 to 2000. Figures 8 to 9 illustrate the interpolated MPI (I_MPI) and HadGEM2 (I_HadGEM2) precipitation data compared to the ERA-interim reanalysis data. Shadow area indicates the mean value ± standard deviation of the ERA-interim data. Evaluating the

interpolated climate data against the initial MPI and HadGEM2 precipitation and temperature dataset, discrepancies are pointed out.

In case of the precipitation data, mostly during the wet period, when the climate models present larger differences in relation to the reanalysis data, the most essential improvements are recorded for both MPI and HadGEM2 climate models. Analytically, based on the spatio-temporal Kriging procedure, the I_MPI and I_HadGEM2 precipitation decrease is estimated equal to 17 mm and 45 mm, in autumn, respectively. As it can be derived from the diagram (Fig. 8), the I_MPI and I_HadGEM2 precipitation dataset are close to the mean value of the ERA-Interim reanalysis data, minimizing the uncertainty, for the period from 1981 to 2000.

Table 2 Mean Square Error (MSE) for each general circulation climate model (GCMs) and the corresponding spatio-temporal variogram covariance models

Variables	Climate models	Period	Covariance models				
			Separable	ProductSum	Metric	SumMetric	Simple SumMetric
Precipitation (mm)	*MPI*	Wet-Cold	135	38.6	217.8	**25.7**	215.4
		Dry-Warm	2.9	0.8	310.9	**0.3**	2.7
	HadGEM2	Wet-Cold	330	85.4	10,759.7	**30.3**	8312.8
		Dry-Warm	3231.2	2430.9	18,060.4	**304.8**	17,049.9
Temperature (°C)	*MPI*	Wet-Cold	225.9	11.7	423.9	**8.9**	84.5
		Dry-Warm	13.7	21.1	225.6	**2.5**	17.9
	HadGEM2	Wet-Cold	13.8	56.1	209.8	**6.6**	30.6
		Dry-Warm	27.9	144.1	338.5	**14.7**	41.5
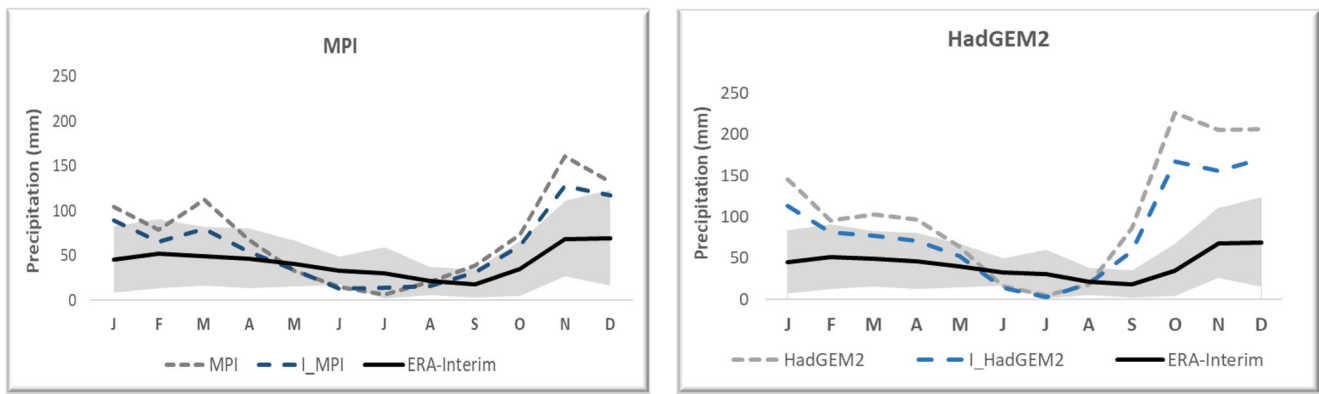
Fig. 8 The interpolated MPI (I_MPI) and HadGEM2 (I_HadGEM2) precipitation data compared to the ERA-interim reanalysis data. Shadow area indicates the mean value ± standard deviation of the ERA-interim data

Nevertheless, during the dry period when the precipitation is minimized, both of the dataset are close to each other.

In regard to the temperature, the dataset is closer to the reanalysis data (Fig. 9). Although, the sum-metric model enhances the temperature during the months in which it is overestimated by the climate models. According to the spatio-temporal Kriging, the I_MPI temparature is neglibly improved, mainly in winter and autumn. In contrast, concering the I_HadGEM2 temperature is reduced by 2.7 °C in winter and 1.1 °C in autumn. Besides, there is no obvious improvement in summer.

The boxplots (Figs. 10 and 11) clearly demonstate that the I_MPI precipitation and temperature dataset approach the ERA-Interim reanalysis data better than the I_HadGEM2 climate data, for the period from 1981 to 2000. In particular, the MPI climate model slightly overestimates the precipitation during winter, autumn and spring (Fig. 10) However, a slight underestimation of the precipitation is observed by the MPI climate model during summer.

On the contrary, according to the I_HadGEM2 precipitation outputs, a significant overestimation is recorded,

especially during autumn and winter (Fig. 10). However, during summer, the HadGEM2 model presents similarities to the MPI simulations.

Regarding the I_MPI temperature, the similarities to the ERA-Interim temperature data are apparent, as illustrated in the boxplots (Fig. 11). However, the I_MPI temperature is slihgtly overestimated in winter. The I_HadGEM2 temperature also presents an overestimation in winter and summer.

Discussion and conclusions

In the present research, the spatio-temporal Kriging technique by using five different covariance models, has been implemented to MPI and HadGEM2 climate models precipitation and temperature dataset. The spatio-temporal Kriging results were evaluated in comparison to the ERA-Interim reanalysis data for the period 1981 to 2000, investigating which one of the spatio-temporal covariance model represents the local climate conditions of the research area more effectively. The accuracy and the reliability of the spatio-temporal
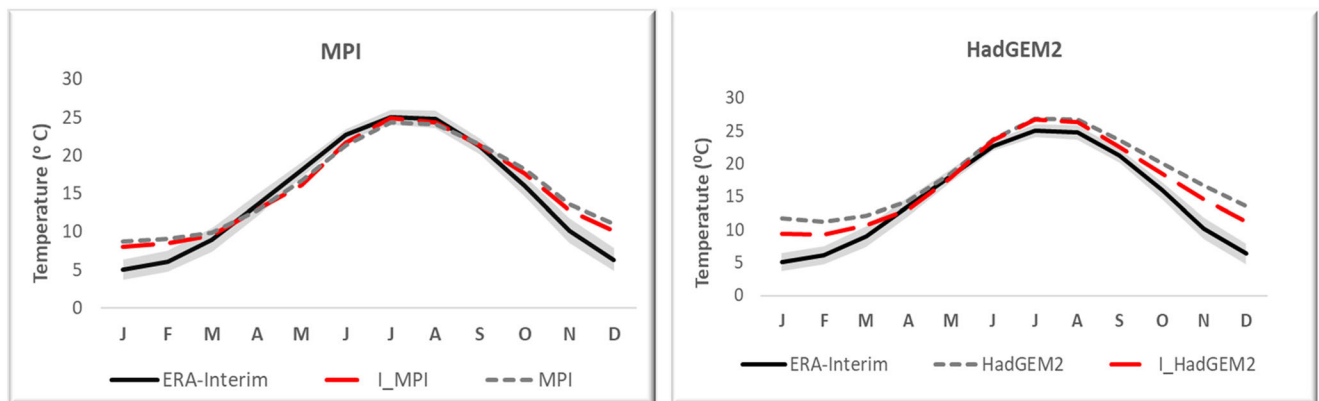

Fig. 9 The interpolated MPI (I_MPI) and HadGEM2 (I_HadGEM2) temperature data compared to the ERA-interim reanalysis data. Shadow area indicates the mean value ± standard deviation of the ERA-interim data
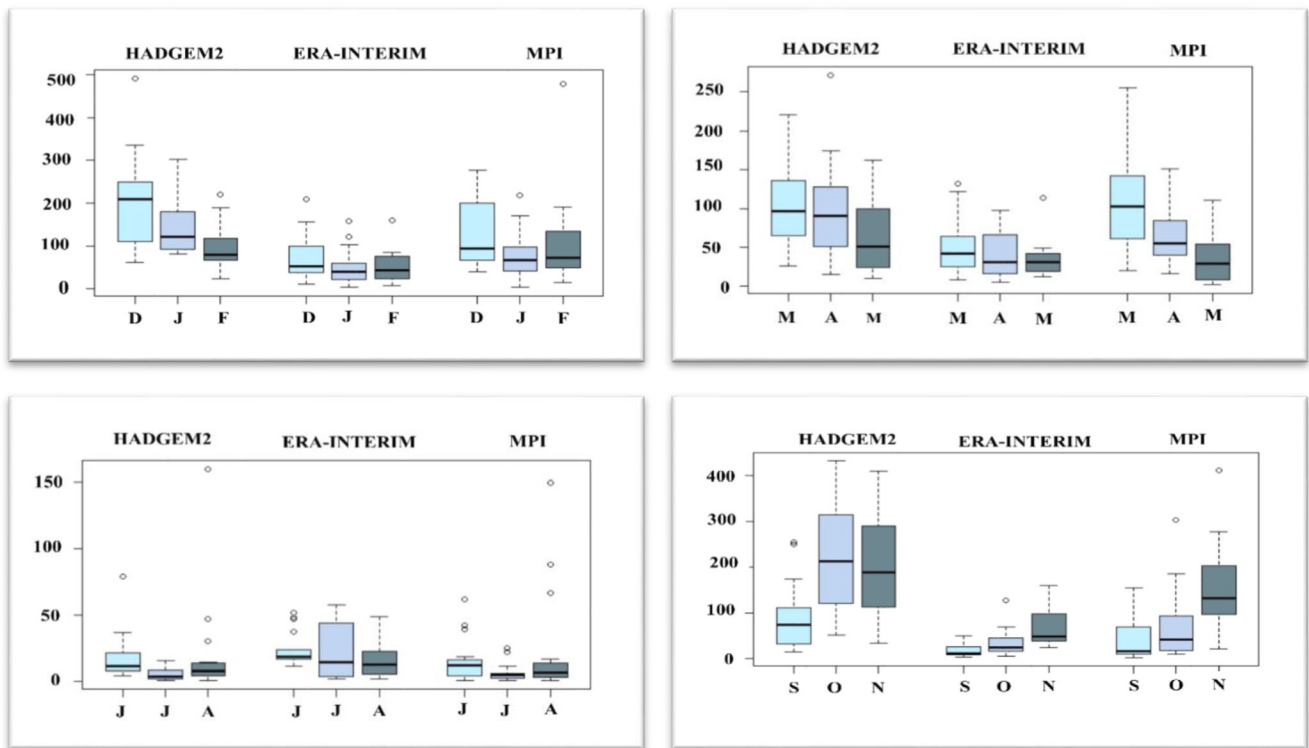
Fig. 10 The boxplots illustrating the seasonal comparison between the interpolated HADGEM2, MPI and ERA-Interim precipitation (mm) dataset for the period from 1981 to 2000

interpolation results were assessed by using the judgment of the wireframe plots between the sample and the fitted

covariance models, Taylor diagram, and the Mean Square Error (MSE).
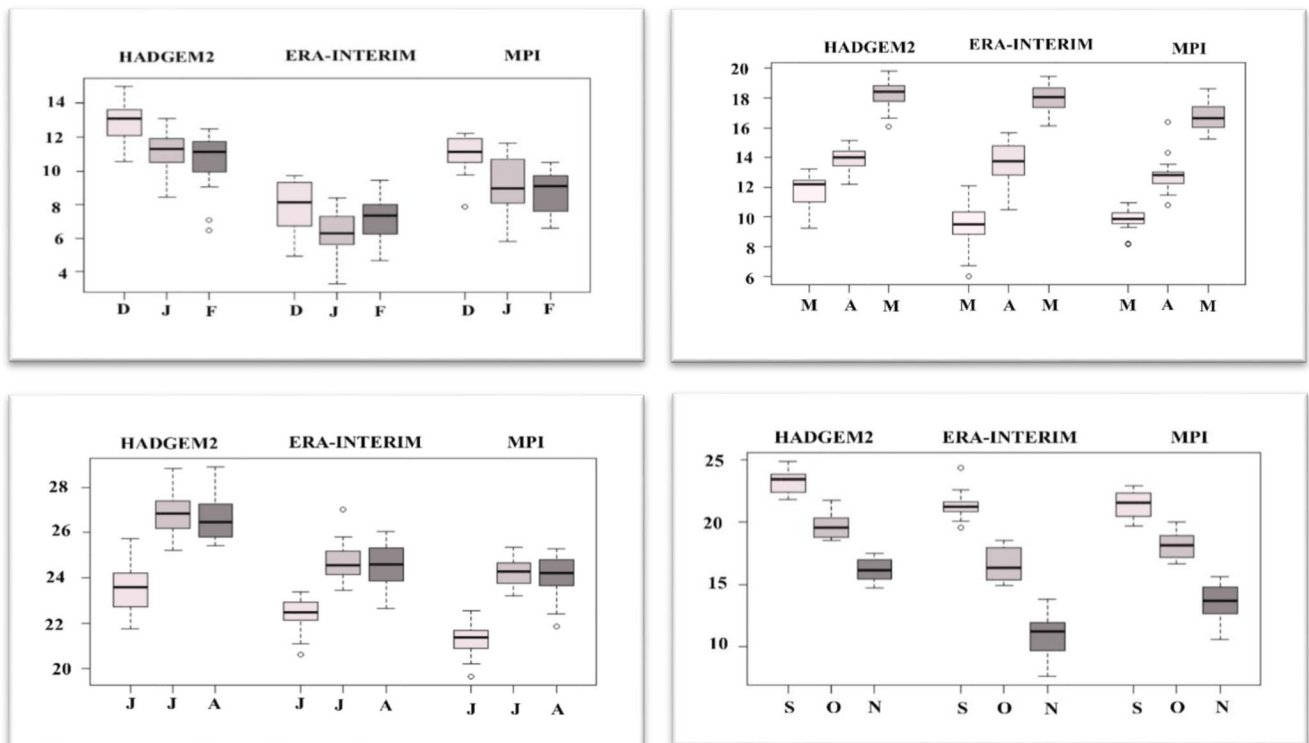
Fig. 11 The boxplots illustrating the seasonal comparison between the interpolated HADGEM2, MPI and ERA-Interim reanalysis temperature (°C) dataset from the period from 1981 to 2000

Judging the performance between the sample and the fitted covariance models is revealed that the sum-metric model is the most suitable among the covariance models. In the contrast, the worst performance is exhibited by the metric covariance model in both MPI and HadGEM2 simulations.

Moreover, Taylor diagrams confirm that the sum-metric covariance model approaches the MPI and HadGEM2 climate data more reasonably than the other four covariance models. Specifically, the sum-mertic model presents the highest correlation coefficient, the smallest standard deviation which lead to the smallest RMS error in both climate models outputs, in all seasons. Consequently, the performance of the sum-metric is superior to the other spatio-temporal covariance models, as it produces more effective results for both MPI and HadGEM2 precipitation and temperature outputs. According to the authors' results, Heuvelink et al. (2012) suggested the sum-metric covariance model as the spatio-temporal model with the least uncertainties. Furthermore, Kilibarda et al. (2014) applied the sum-metric in order to predict the mean, maximum, and minimum daily temperature, globally.

The MPI and HadGEM2 climate models accurately represent the present climate. Demicran et al. (2017) showed that HadGEM2-ES and MPI-ESM-MR's climate models underestimate the temperature, whereas the precipitation is overestimated by MPI-ESM-MR's model in winter, while HadGEM2-ES overestimates it in autumn and spring compared to the observed data in Eastern Mediterranean and particularly, in Turkey from 1971 to 2000. The observed uncertainties of the climate outputs are depending on the GCM driver, the area of interest and the season.

The statistical evaluation indicates that the I_MPI climate model outputs are closer to the reanalysis precipitation and temperature climate data. Consequently, the MPI climate model represents more reasonably and accurately the climate of the research area. On the contrary, the I_HadGEM2 model notably overestimates the climate parameters even after the Kriging interpolation.

Analytically, the spatio-temporal interpolation method led to the optimization of the MPI and HadGEM2 temperatute and precipitation outputs, conversing the seasonal pattern of the reasearch area. Regarding the precipitation data, the major improvement occurs during winter and autumn for both climate models. The uncertainties of the I_MPI precipitation data are minimized by 14% and 20% during winter and autumn, respectively, approaching the ERA-Interim precipitation data better. Regarding the I_HadGEM2 precipitation, during the corresponding periods, the enhancement comes up to 18% and 26%, respectively, without reaching satisfactorily the ERA-Interim reanalysis data. The I_MPI temperature data is negligibly improved, but the I_HadGEM2 temperature is significantly optimized in particular during winter and autumn. Finally, the applying spatio-temporal Kriging technique provides climate data adjusting the ERA-Interim reanalysis data for the period from 1981 to 2000, quantifying the precision of the interpolation.

The analysis of the spatio-temporal Kriging procedure indicates that the selection of the appropriate spatio-temporal model, mostly depends on the data, the research area, and the desired spatio-temporal resolution (Aalto et al. 2013). Additionally, the reliability and the accuracy of spatio-temporal interpolation should be based on both statistical measures and the judgment of the best performance of the fitted covariance models.

Nevertheless, this research has some limitations. The main drawback is the absence of continuous and long time series data from meteorological stations in the research area. Additionally, the study area is a small coastal one, therefore, the transition from the land to the sea and vice versa, effects on the climate models signal reducing their reliability.

The above methodology, can be applied into any coastal areas with similar characteristics in order reasonable climate model dataset to be produced. Finally, the interpolated climate model outputs can be used as inputs in climate change impact studies. This process should be carried out with prudent so as to the current climate pattern of the research area be maintained.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Aalto J, Pirinen P, Heikkinen J, Venäläinen A (2013) Spatial interpolation of monthly climate data for Finland: comparing the performance of kriging and generalized additive models. Theor Appl Climatol 112: 99–111. https://doi.org/10.1007/s00704-012-0716-9

Bilonick RA (1988) Monthly hydrogen ion deposition maps for the northeastern U.S. from July 1982 to September 1984. Atmos Environ 22(9):1909–1924. https://doi.org/10.1016/00046981(88)90080-7

Brassel KE, Douglas R (1979) A procedure to generate thiessen Polygons. Geogr Anal 11(3):290–303

Collins WJ, Bellouin N, Doutriaux-Boucher M, Gedney N, Halloran P, Hinton T, Hughes J, Jones CD, Joshi M, Liddicoat S, Martin G, O'Connor F, Rae J, Senior C, Sitch S, Totterdell I, Wiltshire A, Woodward S (2011) Development and evaluation of an earth-system model HadGEM2. Geosci Model Dev Discuss 4:997–1062. https://doi.org/10.5194/gmdd-4-9972011

Cressie C (1993) Statistics for Spatial Data. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc, New York

Cressie C, Wikle K (1998) The variance-based cross-variogram: you can add apples and oranges. Math Geol 30:789–799

Cressie C, Wikle K (2011) Statistics for Spatio-temporal data. Wiley

De Cesare L, Myers D, Posa D (2001) Estimating and modeling space-time correlation structures. Statist Probab Lett 51(1):9–14. https://doi.org/10.1016/S0167-7152(00)00131-0

De Iaco S, Myers D, Posa D (2001) Space-time analysis using a general product-sum model. Statist Probab Lett 52(1):21–28. https://doi.org/10.1016/S0167-7152(00)00200-5

Dee P, Uppala MS, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Holm EV, Isaksen L, Kallberg P, Kohler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thepaut JN, Vitart F (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. R Meteorol Soc 137:553–597

Demicran M, Gürkan H, Eskioǵlu O, Arabaci H, Coşkun M (2017) Climate change projections for Turkey: three models and two scenarios. Turkish J Water Sci Manage:22–43

Friedman JH (1994) An overview of predictive learning and function approximation. In "from statistics to neural networks", Springer, Berlin, Germany, 136:1–61

Gething P, Atkinson P, Noor A, Gikandi P, Hay S, Nixon M (2007) A local space-time kriging approach applied to a national outpatient malaria data set. Comput Geosci 33(10):1337–1350. https://doi.org/10.1016/j.cageo.2007.05.006

Giorgetta MA, Jungclaus J, Reick CH, Legutke S, Bader J, Böttinger M, Brovkin V, Crueger T, Esch M, Fieg K, Glushak K, Gayler V, Haak H, Hollweg HD, Ilyina T, Kinne S, Kornblueh L, Matei D, Mauritsen T, Mikolajewicz U, Mueller W, Notz D, Pithan F, Raddatz T, Rast S, Redler R, Roeckner E, Schmidt H, Schnur R, Segschneider J, Six KD, Stockhause M, Timmreck C, Wegner J, Widmann H, Wieners KH, Claussen Marotzke J, Stevens B (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. J Adv Model Earth Syst 5:572–597. https://doi.org/10.1002/jame.20038

Giorgi F, Coppola E, Solmon F, Mariotti L, Sylla MB, Bi X, Elguindi N, Diro GT, Nair V, Giuliani G, Turuncoglu UU, Cozzini S, Güttler I, O'Brien TA, Tawfik AB, Shalaby A, Zakey AS, Steiner AL, Stordal F, Sloan LC, Brankovic C (2012) RegCM4: model description and preliminary tests over multiple CORDEX domains. Climate Res 52:7–29. https://doi.org/10.3354/cr01018

Gräler B, Gerharz L, Pebesma E (2016) Spatio-temporal interpolation using gstat. The R Journal 8(1):204–218

Hengl T, Heuvelink GBM, Perˇcec Tadic M, Pebesma EJ (2012) Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. Theor Appl Climatol 107:265–277

Heuvelink GBM, Heuvelink GBM, Griffth DA (2010) Space-time geostatistics for geography: a case study of radiation monitoring across parts of Germany. Geog Anal 42(2):161–179. https://doi.org/10.1111/j.1538-4632.2010.00788.x

Heuvelink GBM, Griffith DA, Hengl T, Melles SJ (2012) Sampling design optimization for space-time kriging. John Wiley, Oxford, pp 207–230. https://doi.org/10.1007/s00704-011-0464-2

Hofstra N, Haylock M, New M, Jones P, Frei C (2008) Comparison of six methods for the interpolation of daily, European climate data. J Geophys Res 113:D21110. https://doi.org/10.1029/2008JD010100

Huang Y, Wong P, Gedeon T (1998) Spatial interpolation using fuzzy reasoning and genetic algorithms. J Geogr Inf Decis Anal 2:204–214

Hutchinson MF (1988) Calculation of hydrologically sound digital elevation models, third international symposium on spatial data handling at Sydney. Australia 3(1):120–127

IPCC (2007) Climate change 2007. Synthesis Report

IPCC (2013) Climate change 2013. Synthesis Report

Jones CD, Hughes JK, Belloin N, Hardiman C, Jones GS, Knight J, Liddicoat S, O'Connor FM, Andres RJ, Bell C, Boo KO, Bozzo A, Butchart N, Cadule P, Corbin KD, Doutriaux-Boucher M, Friedlingstein P, Gornall J, Gray L, Halloran PR, Hurtt G, Ingram

WJ, Lamarque JF, Law RM, Meinshausen M, Osprey S, Palin EJ, Chini LP, Raddatz T, Sanderson MG MG, Sellar AA, Schurer A, Valdes P, Wood N, Woodward S, Yoshioka M, Zerroukat M (2011) The HadGEM2-ES implementation of CMIP5 centennial simulations. Geosci Model Dev 4:543–570. https://doi.org/10.5194/gmd-4-543-2011

Kilibarda M, Hengl T, Heuvelink GBM et al (2014) Spatio-temporal interpolation of daily temperatures for global land areas at1 km resolution. J Geophys Res Atmos 119:2294–2313. https://doi.org/10.1002/2013JD020803

Köppen W (1954) Classification of climates and world patterns. In: Trewartha GT (ed) An Introduction to Climate. McGraw-Hill, New York, pp 225–226

Lasinio GJ, Sahu SK, Mardia KV (2007) Modeling rainfall data using a Bayesian Kriged-Kalman model. In: Upadhyay SK, Singh U, Dey DK (eds) Bayesian statistics and its applications. Tunbridge Wells, UK, Anshan, pp 61–86

Mair A, Fares A (2011) Comparison of rainfall interpolation methods in a mountainous region of a tropical island. J Hydrol Eng 16:371–383

Matheron, G (1962). Traité de Geostatistique Appliquée. Memoires du Bureau de Recherches Geologiques et Minieres, Tome I(14). Paris: Editions Technip (in French)

Pal JS, Giorgi F, XBi E et al (2007) Regional climate modeling for the developing world: the ICTP RegCM3 and RegCNET. Bull Am Meteorol Soc 88:1395–1409

Pebesma E (2004) Multivariable Geostatistics in S: the gstat package. Comput Geosci 30(7):683–691. https://doi.org/10.1016/j.cageo.2004.03.012

Pebesma E (2012) Spacetime: Spatio-temporal data in R. J Stat Softw 51(7):1–30. https://doi.org/10.18637/jss.v051.i07

Pebesma E, Gräler B (2017) Introduction to Spatio-temporal Variography. Institute for Geoinformatics University of Münster: 1–11

Popke D, Stevens B, Voigt A (2013) Climate and climate change in a radiative-convective equilibrium version of ECHAM6. J Adv Model Earth Syst 5:1–14. https://doi.org/10.1029/2012MS000191

R Core Team (2013) A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria., see https://www.R-project.org/

Sibson R (1981) A brief description of natural neighbour interpolation, in interpreting multivariate data. John Wiley and Sons, Chichester UK, pp 21–36

Snepvangers J, Heuvelink G, Huisman J (2003) Soil water content interpolation using spatio-temporal kriging with external drift. Geoderma 112(3–4):253–271. https://doi.org/10.1016/S0016-7061(02)00310-5

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res 106(D7):7183–7192

Tolika K, Maheras P, Vafiadis M, Flocas HA, Arseni-Papadimitriou A (2006) An evaluation of a general circulation model (GCM) and the NCEP-NCAR reanalysis data for winter precipitation in Greece. Int J Climatol 26:935–955

Wackernagel H (2003) Multivariate Geostatistics: an introduction with applications, 3rd edn. Springer, Berlin, Heidelberg

Wong KW, Wong PM, PM TDGTD, Fung CC (2003) Rainfall prediction model using soft computing technique. Soft Comput 7(6):434–438

Xoplaki E, González-Rouco JF, Luterbacher J, Wanner H (2004) Wet season Mediterranean precipitation variability: influence of large-scale dynamics and trends. Clim Dyn 23:63–78

Yang X, Xie X, Liu Ji F, Wang L (2015) Spatial interpolation of daily rainfall data for local climate impact assessment over greater Sydney region. Hindawi publishing corporation advances in meteorology, 1-12. https://doi.org/10.1155/2015/563629

Zimmerman D, Pavlik C, Ruggles A, Armstrong MP (1999) An experimental comparison of ordinary and universal kriging and inverse distance weighting. Math Geol 31(4):375–390