sustainability

MDPI

Article

Hybrid Machine Learning-Geostatistical Framework for Sustainable Spatiotemporal Temperature Forecasting

Nouman Iqbal, Sandra De Iaco and Monica Palma *

Department of Economic Sciences, University of Salento, 73100 Lecce, Italy; nouman.iqbal@unisalento.it (N.I.); sandra.deiaco@unisalento.it (S.D.I.)
* Correspondence: monica.palma@unisalento.it

Abstract

Temperature plays a key role in climate systems, with significant implications for environmental sustainability, agricultural productivity, and water resource management. Accurate spatiotemporal forecasting of temperature is, therefore, essential for supporting sustainable development and informing climate adaptation strategies. However, traditional modeling approaches often treat spatial and temporal dimensions separately, thereby limiting their ability to capture the full spectrum of spatiotemporal dependencies. This paper proposes a hybrid approach that integrates a machine learning model with a geostatistical prediction technique to forecast daily mean air temperature over the study area. The best-performing spatiotemporal correlation model is selected among various time series and machine learning models including Holt–Winters, Seasonal Autoregressive Integrated Moving Average (SARIMA), Neural Network Autoregression (NNAR), and Artificial Neural Networks (ANNs). The findings demonstrate that the proposed hybrid approach consistently outperforms traditional, non-integrated methods. Importantly, this study contributes to sustainability by enabling high-resolution temperature forecasting that supports climate-resilient agriculture, efficient water resource allocation, and evidence-based environmental management. The generated predictive maps provide actionable insights for policymakers and stakeholders, enhancing adaptive capacity and promoting sustainable resource management under changing climate conditions.

Keywords: spatiotemporal predictions; sustainable climate modeling; kriging; error metrics; climate adaptation

1. Introduction

Understanding spatiotemporal variability in temperature is essential for managing climate-sensitive systems, particularly in Mediterranean regions such as Apulia, Italy. In this context, temperature dynamics directly influence agricultural productivity, irrigation demand, and water resource allocation. Reliable temperature forecasting, therefore, plays a key role in supporting operational decisions, including irrigation scheduling [1], drought preparedness [2], and frost risk mitigation [3], contributing to climate adaptation and sustainable resource management [4,5], while influencing air pollutant concentrations [6].

Temperature exhibits complex behavior characterized by seasonal cycles, long-term trends, and spatial heterogeneity [7–9]. Traditional modeling approaches often treat spatial and temporal components separately, limiting their ability to capture the joint dependence

Check for updates

Academic Editor: David Pastor-Escuredo

Received: 30 March 2026
Revised: 28 May 2026
Accepted: 8 June 2026
Published: 14 July 2026

structure inherent in environmental processes [10]. Consequently, purely spatial or temporal interpolation methods may fail to represent adequately interactions between space and time, leading to reduced predictive accuracy [11–13].

Spatiotemporal geostatistics provides a rigorous framework for modeling such dependencies by explicitly accounting for covariance structures across both spatial and temporal dimensions. Theoretical advances, such as the development of valid space-time variogram and covariance functions [14–26], even in the complex domain [27–29], or the assessment of their peculiar characteristics, such as the behavior near the origin or at infinity, and the type of separability [30,31], have provided robust tools for modeling stationary and non-stationary processes as well as anisotropy, enabling a more accurate characterization of environmental phenomena. These methods have been widely applied in environmental sciences, including temperature and precipitation modeling, air quality assessment, and hydrological analysis, demonstrating their effectiveness in capturing complex environmental variability [32–39]. These methods have also been successfully applied to sub-daily precipitation mapping in Romania [40], where high resolution spatiotemporal interpolation improved the representation of extreme weather events. Furthermore, the analysis of global temperature station data [41] highlighted the utility of spatiotemporal models in generating continuous, high resolution temperature fields, which are critical for understanding climate variability and changes.

Beyond environmental sciences, spatiotemporal geostatistical and space–time models have also been applied in other disciplines characterized by strong spatial heterogeneity and temporal dynamics, demonstrating the methodological flexibility of these approaches. In spatial epidemiology and public health, space–time kriging and geostatistical modeling have been extensively employed for disease mapping, malaria risk assessment, and the prediction of health outcomes from incomplete or sparsely observed data, enabling spatial and temporal variations in disease transmission to be jointly characterized [42–47]. In the socioeconomic domain, spatiotemporal geostatistical models have been adopted for hedonic real-estate price estimation, explicitly accounting for spatial dependence and temporal variation in property values [48]. Similarly, in criminology and public-safety studies, spatiotemporal cokriging has been applied to crime prediction by integrating crime patterns with auxiliary information derived from social media [49]. Although these applications arise from different domains, they collectively highlight the broad applicability and effectiveness of geostatistical space–time modeling frameworks for analyzing and predicting complex real-world processes characterized by joint spatial and temporal dependence.

More recently, hybrid approaches integrating machine learning with geostatistical methods have attracted increasing attention for their ability to capture non-linear temporal dynamics while preserving spatial dependence structures [50–55]. Applications of these hybrid approaches include urban water demand forecasting [56], air quality prediction and pollutant estimation [57–59], and temperature estimation using machine learning-assisted spatial models [60]. Additional studies have explored combinations of neural networks with autoregressive time series models [61,62], as well as radial basis functions or regression–kriging frameworks for environmental prediction [63]. Despite these advances, existing approaches frequently rely on generic machine learning models, separable covariance structures, or station-level predictions, limiting their ability to capture fully the complex space–time interactions at high temporal resolution.

This study develops a structured spatiotemporal forecasting framework that integrates a Neural Network AutoRegression (NNAR) model with spatiotemporal kriging of residuals based on a non-separable covariance function. The NNAR component captures non-linear temporal dependencies and seasonal patterns through an autoregressive neural network structure, while the geostatistical component models residual spatial and temporal

https://doi.org/10.3390/su18147193

dependence jointly, enabling improved reconstruction of continuous temperature fields.

The novelty of this work lies in four main aspects. First, it combines NNAR-based temporal forecasting with residual correction through non-separable spatiotemporal kriging, allowing simultaneous modeling of non-linear temporal dynamics and space-time dependence. Second, it proposes a unified forecasting interpolation workflow in which residual diagnostics explicitly guide the geostatistical correction stage. Third, the framework is applied to daily temperature prediction at a regional scale, producing continuous high-resolution temperature maps rather than discrete station-level forecasts. Fourth, the approach provides high-resolution temperature outputs at regional scale, with potential use in environmental and climate-related studies.

The proposed framework is evaluated against several baseline models, including Holt–Winters exponential smoothing, Seasonal AutoRegressive Integrated Moving Average (SARIMA), Artificial Neural Networks (ANNs), and NNAR. Model performance is assessed through forecasting accuracy and residual diagnostics to validate the effectiveness of the hybrid approach.

After introducing the fundamental concepts of the spatiotemporal stochastic process and its statistical characteristics, the paper presents the theoretical framework of the proposed modeling approach, and details the methodological formulation and the integration of NNAR modeling with spatiotemporal kriging, in Section 2. The subsequent model diagnostic procedures are described in Section 3, including the evaluation of forecasting accuracy and residual analysis for performance validation. Finally, Section 4 illustrates the case study on temperature prediction in the Apulia region, demonstrating the practical implementation of the proposed hybrid framework and discussing the resulting spatiotemporal patterns and forecasting outcomes.

2. The Theoretical Framework

The paper introduces a hybrid framework that integrates NNAR with spatiotemporal kriging, aiming to combine the temporal predictive strength of neural networks with the spatial interpolation capability of geostatistical methods. This novel approach is carried out after comparing the findings of some of the most widely used traditional forecasting models, such as SARIMA, Holt–Winters, ANNs and NNAR.

Prior to the implementation of the proposed hybrid approach, a review of the main theoretical aspects of each individual model is presented in the sections below.

2.1. SARIMA Model

The SARIMA model, an extension of the ARIMA framework, is specifically designed to handle time series data exhibiting seasonal patterns [64]. The modeling procedure comprises three main stages: identification, parameter estimation, and diagnostic checking. In the identification phase, the time series is examined for stationarity and seasonality using autocorrelation (ACF) and partial autocorrelation (PACF) plots. When the series is non-stationary, differencing is applied to stabilize it. During estimation, the model parameters $(p, d, q) \times (P, D, Q)_s$ are determined, where p, d and q represent the non-seasonal autoregressive (AR), differencing, and moving average (MA) orders, while P, D, and Q represent the corresponding seasonal components with seasonal period s. Finally, diagnostic checking evaluates the model's adequacy by analyzing residuals to ensure they resemble a white noise process. Mathematically, the SARIMA model is expressed as follows. Let X_t be the investigated time series and ϵ_t the white noise error term $(t = 1, 2, \ldots, T;\ T \subseteq \mathbb{R})$, the SARIMA model can be expressed as:

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D X_t = \Theta_Q(B^s)\theta_q(B)\epsilon_t, \qquad t = 1, \ldots, T, \tag{1}$$

where the $\phi_p(B)$, $\Phi_P(B^s)$, and $\theta_q(B)$, $\Theta_Q(B^s)$ are the non-seasonal and seasonal AR and MA polynomials, respectively, with B the backward shift operator. Furthermore, the $(1-B)^d$ and $(1-B^s)^D$ are the non-seasonal and seasonal differencing operators.

2.2. NNAR Model

The NNAR model is a specialized neural network architecture designed for time series forecasting [65–67]. It leverages both seasonal and non-seasonal autoregressive structures, in which past observations and seasonal lags serve as inputs to predict future values. The model employs a feed-forward neural network structure that enables the representation of linear, non-linear, and seasonal dependencies in the data [68,69]. The NNAR model consists of an input layer, a hidden layer, and an output layer, where lagged observations are processed through the hidden layer to generate forecasts and capture complex temporal dependencies.

Let $X_t, t = 1, 2, \ldots, T$, be the time series. The detailed formulation of the NNAR model is given as

$$X_t = \alpha_0 + \sum_{h=1}^{H} \alpha_h \, \sigma\left(\sum_{i=1}^{p} w_{ih} X_{t-i} + \sum_{k=1}^{P} v_{kh} X_{t-ks} + b_h \right) + \varepsilon_t, \tag{2}$$

where p and P denote the non-seasonal and seasonal lag orders with seasonal period s, respectively, and H is the number of hidden neurons. The parameters w_{ih}, v_{kh}, and α_h represent the network weights, while b_h and α_0 are bias terms. Furthermore, $\sigma(\cdot)$ denotes the activation function and ε_t the error term.

2.3. Additive Exponential Smoothing

Exponential smoothing is a widely used technique for analyzing time series data; it assigns greater weight to recent observations while progressively diminishing the influence of older ones. There are three main types of exponential smoothing: single, double, and triple. Single exponential smoothing models the level component of the series, double exponential smoothing additionally incorporates the trend component, while triple exponential smoothing also accounts for seasonality. The latter is commonly referred to as the Holt–Winters method, which comes in two variations: additive and multiplicative exponential smoothing [70,71].

Let X_t be the time series under study, characterized by a trend and a constant seasonal pattern. To forecast the time series at the time point $t+m$, Holt–Winters additive exponential smoothing accounts for the overall level, trend, and seasonal components of the series, i.e.,

$$\begin{aligned} X_{t+m} &= a_t + m b_t + s_{t-s} + \epsilon_t, \\ a_t &= \alpha(X_t - s_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1}), \\ b_t &= \beta(a_t - a_{t-1}) + (1-\beta) b_{t-1}, \\ s_t &= \gamma(X_t - a_t) + (1-\gamma) s_{t-s}. \end{aligned} \tag{3}$$

where a_t, b_t, and s_{t-s} denote the level, trend, and seasonal components, respectively, while α, β, and γ are the associated smoothing parameters, each taking values between 0 and 1.

2.4. ANN Model

Artificial Neural Networks ANNs are widely used machine learning models for time series forecasting and regression tasks [72,73]. Inspired by biological neural networks, ANNs are capable of representing complex non-linear relationships in data. They commonly employ a feed-forward architecture with a single hidden layer for time series forecasting applications [61]. Mathematically, the ANN model can be expressed as

$$X_t = f(X_{t-1}, X_{t-2}, \ldots, X_{t-p}) + \epsilon_t, \tag{4}$$

where p denotes the number of lagged inputs and ϵ_t represents the error term. The lagged input observations are processed through the hidden layer to generate forecasts and represent temporal dependencies in the data. A detailed formulation of the ANN model is given as follows:

$$X_t = \alpha_0 + \sum_{h=1}^{H} \alpha_h \, \sigma\left(b_h + \sum_{i=1}^{p} w_{ih} X_{t-i}\right) + \varepsilon_t, \tag{5}$$

where p is the number of lagged inputs and H is the number of hidden neurons. Furthermore, w_{ih} and α_h are the weights connecting inputs to hidden neurons and hidden neurons to the output, respectively. The terms b_h and α_0 denote the bias parameters of the hidden and output layers, respectively, while $\sigma(\cdot)$ is the activation function and ε_t is the error term.

2.5. Space–Time Kriging

In Geostatistics, phenomena that vary jointly over space and time are represented using a spatiotemporal random function (STRF), denoted by $\boldsymbol{X}(\mathbf{s}, t)$, where $\mathbf{s} \in D \subset \mathbb{R}^d$ $(d \leq 3)$ denotes a spatial location vector, typically $\mathbf{s} = (s_1, s_2)$, and $t \in T \subset \mathbb{R}$ denotes time. The process is commonly decomposed into a deterministic mean component $\mu(\mathbf{s}, t)$, accounting for large-scale variation, and a stochastic residual component $\epsilon(\mathbf{s}, t)$, representing small-scale variability, i.e.,

$$\boldsymbol{X}(\mathbf{s}, t) = \mu(\mathbf{s}, t) + \epsilon(\mathbf{s}, t). \tag{6}$$

The residual process $\epsilon(\mathbf{s}, t)$ is assumed to have zero mean and constant variance, and its dependence structure is characterized by a spatiotemporal covariance function or, equivalently, a variogram, that under the assumption of second-order stationarity, depends only on the spatiotemporal separation between points, namely the lag vector $\mathbf{h} = (\mathbf{h}_s, h_t)$, where $\mathbf{h}_s = (\mathbf{s}' - \mathbf{s})$ denotes the spatial lag vector and $h_t = (t' - t)$ denotes the temporal lag. Hence, the spatiotemporal covariance function and variogram are defined, respectively, as follows:

$$C_{ST}(\mathbf{h}) = \mathrm{Cov}[\epsilon(\mathbf{s} + \mathbf{h}_s, t + h_t), \epsilon(\mathbf{s}, t)], \tag{7}$$

$$2\gamma_{ST}(\mathbf{h}) = \mathrm{Var}[\epsilon(\mathbf{s} + \mathbf{h}_s, t + h_t) - \epsilon(\mathbf{s}, t)]. \tag{8}$$

Both functions provide equivalent representations of spatiotemporal dependence and must satisfy admissibility conditions to ensure a valid covariance or variogram structure. The specification of an appropriate spatiotemporal variogram or covariance model is, therefore, fundamental to spatiotemporal inference [74,75].

Spatiotemporal kriging exploits this dependence structure to produce best linear unbiased predictions (BLUPs) at unobserved space–time locations by optimally combining information from neighboring observations in space and time. This framework is particularly well suited to environmental variables such as rainfall, temperature, air quality, and groundwater levels, which exhibit correlated variability across both spatial and temporal dimensions.

Over the years, various formulations of spatiotemporal covariance and variogram models have been developed to achieve an appropriate balance between flexibility, interpretability, and mathematical soundness, particularly with respect to maintaining positive definiteness. The linear model [76] is one of the earliest approaches, combining spatial and temporal components linearly. The metric model [77] uses a metric space framework to unify spatial and temporal dimensions, while the sum-metric model [38] extends this by incorporating a sum of spatial, temporal, and joint components. The product model [78] depends on a product of separate spatial and temporal functions, and the Cressie–Huang models [14] use spectral methods to construct flexible, non-separable covariance structures. The product-sum model [15] combines the product and sum of spatial and temporal components for greater flexibility, and the Gneiting models [16] introduce non-separable covariance functions capable of capturing complex interactions. Finally, the integrated product-sum model [20] integrates multiple approaches to model non-separable dependencies effectively.

The selection of an appropriate model should be guided by geometric characteristics and theoretical considerations. For example, ref. [79] compared the sum-metric, product-sum, and Gneiting models for groundwater level data, emphasizing the importance of model selection based on data characteristics. Similarly, ref. [80] evaluated the product-sum, Cressie–Huang, and Gneiting variogram models for mapping satellite-observed XCO_2 and compared their performance in variogram fitting and spatiotemporal prediction. These comparative studies demonstrate that the selection of a space–time dependence model should be based on its ability to represent the empirical dependence structure and provide reliable predictions.

Theoretical developments have further strengthened the appeal of the product-sum class. Ref. [81] introduced a general product-sum framework for space–time analysis, demonstrating its flexibility in representing complex dependence structures while maintaining valid covariance properties. Subsequent studies by [82–86] established conditions for strict positive definiteness of product and product-sum covariance functions, providing a rigorous mathematical foundation for their application in spatiotemporal modeling. In particular, ref. [87] formulated spatiotemporal random processes within a linear mixed-model framework and developed computational advances for product, sum, and product-sum covariance functions. Their results further demonstrated the practical relevance of product-sum structures for modeling nonseparable spatial and temporal dependence.

Empirical studies, such as those by [88], demonstrated the superiority of the product-sum model over the Cressie–Huang and Gneiting models in terms of mean square errors and cross-validation results, making it a preferred choice for spatiotemporal variogram modeling.

3. Model Diagnostic

The Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) are among the most widely used accuracy measures for evaluating predictive models in Statistics and Environmental Sciences. The theoretical foundations of these error metrics trace back to the early development of statistical estimation theory. In particular, Fisher's seminal work on maximum likelihood estimation (MLE) [89] provided a formal statistical framework linking error distributions to optimal estimation criteria. Within this framework, minimizing the sum of squared errors, corresponding to RMSE, is optimal when model errors are approximately Gaussian, whereas minimizing the sum of absolute errors, corresponding to MAE, is optimal when errors follow a Laplace distribution. Accordingly, RMSE and MAE have been widely adopted in environmental and climate modeling as standard measures for evaluating predictive accuracy under different error distributions [90–92].

In the present study, RMSE and MAE are used to assess the performance of the adopted models. These metrics are computed at each location across time and then averaged spatially over all locations, thereby capturing both spatial and temporal variations in prediction errors. The formulations are given by

$$RMSE = \sqrt{\frac{1}{S}\sum_{s=1}^{S}\left(\frac{1}{T}\sum_{t=1}^{T}\left(X(s,t) - \widehat{X}(s,t)\right)^2\right)} \tag{9}$$

$$MAE = \frac{1}{S}\sum_{s=1}^{S}\left(\frac{1}{T}\sum_{t=1}^{T}\left|X(s,t) - \widehat{X}(s,t)\right|\right) \tag{10}$$

where $X(s,t)$ represents the random variable at spatial location s and time t, $\widehat{X}(s,t)$ the predicted variable at the spatial location s and time t, T the number of time steps for each spatial location, and S the total number of spatial locations.

The spatiotemporal formulation of RMSE and MAE enables a comprehensive assessment of predictive accuracy across both space and time, while also capturing the heterogeneity of prediction errors across different locations and time periods.

4. Case Study: Air Temperature over the Apulia Region

This section begins with a concise description of the geographical area under study, emphasizing its geological and meteorological characteristics. Then, it introduces the spatiotemporal dataset, which contains 2-m air temperature recorded at 30 locations over a 42-year period. Finally, the spatial and temporal profiles of the analyzed variable are discussed.

4.1. The Study Area

The Apulia region, situated on the Adriatic coast in south-eastern Italy, spans the geographic coordinates 39.9–41.1 N latitude and 15.0–17.0 E longitude, as illustrated in Figure 1. The region extends approximately 350 km, with an average elevation of about 100 m above sea level [93]. It represents a strategically important area for climate research, particularly owing to its strong dependence on agriculture and natural resources.

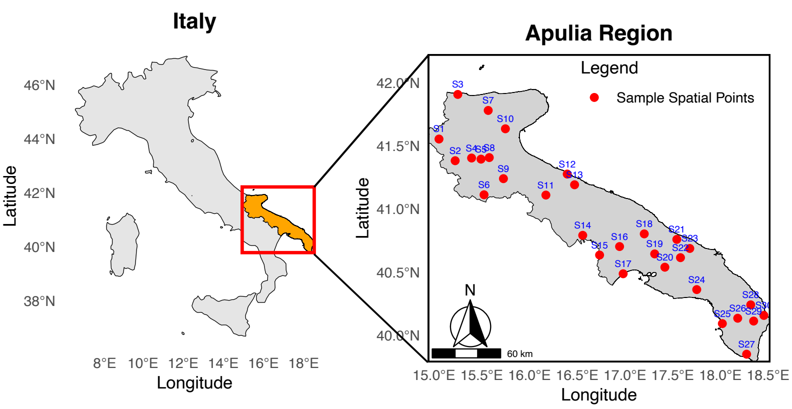


Figure 1. Geographic map of the Apulia region in southeastern Italy, showing the study area and the spatial sampling points used in the analysis. The map highlights the boundaries of the Apulia region and the distribution of the selected locations.

Apulia is characterized by a Mediterranean climate, with hot, dry summers and mild, rainy winters [39]. These climatic conditions play a crucial role in shaping agricultural productivity and water availability across the region. However, ongoing climate change has increased the frequency and intensity of temperature extremes, rendering the region particularly vulnerable to environmental stress [94,95].

Given its economic reliance on climate-sensitive sectors, such as agriculture and water resources, accurate temperature forecasting is essential for supporting efficient resource use, risk mitigation, and adaptive planning [96]. In this context, high-resolution spatiotemporal temperature modeling can provide valuable insights for managing climate-related challenges and improving regional resilience.

4.2. Spatiotemporal Dataset

This study analyses a comprehensive dataset of surface air temperature derived from the NASA POWER (Prediction Of Worldwide Energy Resources) project, which is based on the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) reanalysis conducted by NASA's Global Modeling and Assimilation Office (GMAO). The MERRA-2 assimilation system integrates a wide range of observational data sources, including satellite retrievals, surface measurements, as well as radiosondes, aircraft and marine observations, within a global atmospheric model to produce consistent meteorological fields on a spatial grid of 0.5° × 0.625° and hourly temporal resolution.

For this study, daily mean temperature data have been extracted from the NASA POWER Data Access Viewer (https://power.larc.nasa.gov/data-access-viewer/ (accessed on 10 June 2025)) for thirty randomly distributed sample points across the Apulia region, Italy, which have been selected by ensuring adequate spatial coverage and representativeness of the region's climatic variability. The data set spans a 42-year period (1 January 1982–31 December 2023), providing a long-term and high-resolution basis for climatological assessment. The daily values represent averages derived from hourly MERRA-2 fields, transformed from Coordinated Universal Time (UTC) to solar local time by the POWER processing system.

The dataset has been employed to investigate spatiotemporal dynamics, persistence, variability, and trends in temperature over Apulia, supporting a robust and reproducible analysis of regional climate behavior over recent decades.

4.3. Workflow of the Proposed Approach

The workflow of the proposed hybrid spatiotemporal approach is shown in Figure 2. First, historical temperature observations have been collected and organized by spatial location. For each location, the data have been divided into a training period (1982–2022) and a testing period (2023), enabling model calibration and independent validation.

Subsequently, several temporal forecasting models, including Holt–Winters, NNAR, ANNs, and SARIMA, have been trained independently at each spatial location using the corresponding local temperature time series. This step enables the modeling of spatially differentiated temporal dynamics, including trends, seasonality, and non-linear behavior. Model performance has been then evaluated separately for each location using the RMSE and MAE, and the best-performing model has been selected.

After model selection, the optimal temporal model has been used to decompose the temperature series and extract the residual component. These residuals represent the variability not explained by temporal modeling alone and may exhibit spatial and spatiotemporal dependence. To account for this dependence, the residuals have been subsequently modeled using spatiotemporal kriging.

Finally, the spatiotemporal kriging-based residual estimates have been combined with the forecasts produced by the best-performing temporal model, yielding the final hybrid spatiotemporal predictions.

The proposed methodology integrates spatiotemporal modeling with machine learning and time series techniques to enhance temperature forecasting accuracy. The hybrid approach of the best-selected model combined with kriging leverages the strengths of both temporal and spatial modeling, providing a more comprehensive and reliable predictive framework for Apulia's temperature conditions. This improved predictive capability provides a valuable basis for understanding regional climate patterns and supports more informed planning and management in climate-sensitive sectors.

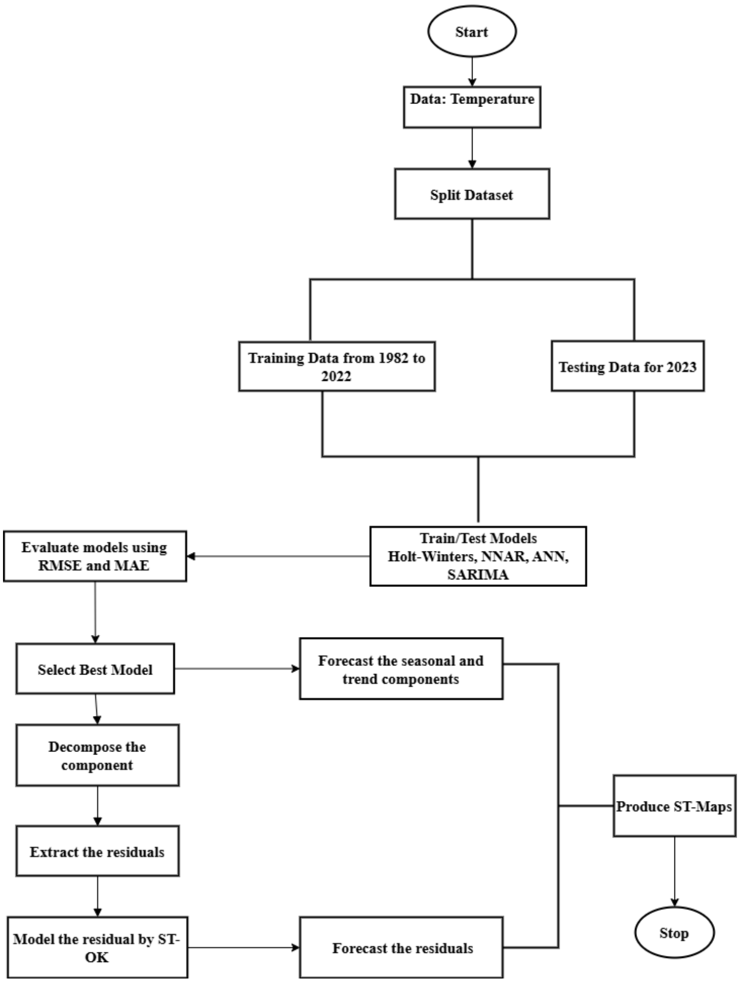


Figure 2. Workflow of the adopted hybrid spatiotemporal approach.

4.4. Exploratory Data Analysis

The summary statistics of daily temperature data from 1982 to 2023, collected across 30 sample spatial points in Apulia, are presented in Table 1. The results are organized by province and land characteristics, allowing a detailed comparison between inland and coastal environments.

The northern inland stations in the province of Foggia (S1–S10) exhibit the coldest conditions and the greatest variability. Minimum temperatures reach −8.09 °C at stations S1 and S2 on 10 January 1985, while similarly low values (−7.79 °C) are observed at S6 and S9 on the same date. These stations are characterized by low first-quartile values (7.69–7.80 °C), median temperatures around 13.4 °C, mean values close to 14.1 °C, and the highest standard deviations in the region (up to 7.75 °C), indicating strong seasonal contrasts typical of continental lowland environments. Third-quartile values remain relatively low (approximately 20.5–20.6 °C). Summer maxima reached 36.05 °C at several inland stations (S4, S5, S8, and S10) on 10 August 1999, confirming the large thermal amplitude characteristic of northern Apulia. Moving southward and toward the Adriatic coast, stations located in the provinces of BAT-Barletta, Andria, Trani (S11–S13) and Bari (S14 and S18) show intermediate thermal characteristics. Minimum temperatures range from −3.39 °C on 10 January 1985 to 2.81 °C on 7 January 2017. First-quartile values increase to approximately 9.1–13.0 °C, median temperatures range between 14.8 and 17.0 °C, and mean values reach up to 17.75 °C. Third-quartile values increase to 22.0–22.8 °C, while standard deviations decrease (5.65–7.63 °C), reflecting the moderating influence of coastal exposure.

Table 1. Summary statistics of the daily temperature measured at the sample spatial points from 1982 to 2023.

Sample Points	Province	Min	1st_Qu.	Median	Mean	3rd_Qu.	Max	SD
S1	Foggia	−8.09	7.69	13.36	14.14	20.61	34.13	7.75
S2	Foggia	−8.09	7.69	13.36	14.14	20.61	34.13	7.75
S3	Foggia	0.36	11.36	16.23	17.03	22.67	33.90	6.56
S4	Foggia	−1.91	9.99	15.74	16.57	23.01	36.05	7.63
S5	Foggia	−1.91	9.99	15.74	16.57	23.01	36.05	7.63
S6	Foggia	−7.79	7.80	13.44	14.13	20.51	33.57	7.68
S7	Foggia	0.64	11.76	16.13	16.95	22.28	32.21	6.06
S8	Foggia	−1.91	9.99	15.74	16.57	23.01	36.05	7.63
S9	Foggia	−7.79	7.80	13.44	14.13	20.51	33.57	7.68
S10	Foggia	−1.91	9.99	15.74	16.57	23.01	36.05	7.63
S11	BAT	−3.39	9.13	14.80	15.62	22.03	34.99	7.63
S12	BAT	0.17	11.29	16.12	16.91	22.53	33.61	6.47
S13	BAT	−3.39	9.13	14.80	15.62	22.03	34.99	7.63
S14	Bari	−3.39	9.13	14.80	15.62	22.03	34.99	7.63
S15	Taranto	−0.64	10.88	15.91	16.81	22.78	34.73	6.93
S16	Taranto	−0.64	10.88	15.91	16.81	22.78	34.73	6.93
S17	Taranto	−0.64	10.88	15.91	16.81	22.78	34.73	6.93
S18	Bari	2.81	12.98	17.00	17.75	22.80	33.43	5.65
S19	Taranto	−0.60	10.83	16.23	17.07	23.39	36.40	7.37
S20	Taranto	−0.60	10.83	16.23	17.07	23.39	36.40	7.37
S21	Brindisi	1.83	11.65	15.56	16.34	21.38	30.15	5.56
S22	Brindisi	−0.60	10.83	16.23	17.07	23.39	36.40	7.37
S23	Brindisi	−0.60	10.83	16.23	17.07	23.39	36.40	7.37
S24	Taranto	−0.60	10.83	16.23	17.07	23.39	36.40	7.37
S25	Lecce	3.09	13.23	17.54	18.31	23.64	32.70	6.04
S26	Lecce	0.90	11.58	15.65	16.30	21.40	29.77	5.66
S27	Lecce	3.09	13.23	17.54	18.31	23.64	32.70	6.04
S28	Lecce	3.09	13.23	17.54	18.31	23.64	32.70	6.04
S29	Lecce	3.09	13.23	17.54	18.31	23.64	32.70	6.04
S30	Lecce	3.38	13.88	17.62	18.22	22.94	32.01	5.20

In the province of Brindisi (S21–S23), located along the southern Adriatic coast, thermal behavior is more heterogeneous. Stations S22 and S23 record the highest maximum temperature of the dataset (36.40 °C) on 24 July 2007, whereas station S21 exhibits a lower maximum (30.15 °C) on 22 July 2023. First-quartile values range from 10.83 to 11.65 °C, median temperatures are approximately 15.6–16.2 °C, and third-quartile values span 21.4–23.4 °C. Standard deviations vary between 5.56 and 7.37 °C, indicating locally variable but generally moderated thermal regimes.

Stations in the province of Taranto (S15–S17, S19, S20, and S24), influenced by the Ionian Sea but including inland locations, show relatively mild winter conditions and pronounced summer extremes. Minimum temperatures are generally close to −0.6 °C and occur during the cold spell of 7 January 2017, while maximum temperatures reach 36.40 °C at stations S19, S20, and S24 on 24 July 2007. These stations are characterized by first-quartile values around 10.8–10.9 °C, median temperatures near 16.1–16.2 °C, third-quartile values exceeding 23.3 °C, and consistently high standard deviations (approximately 6.9–7.4 °C).

The southernmost stations in the province of Lecce (S25–S30), forming the Salento peninsula and strongly influenced by both the Adriatic and Ionian seas, exhibit the warmest and most stable thermal conditions in Apulia. Minimum temperatures remain above 0 °C at all stations, recorded mainly on 7–10 January 2017 and 13 February 2004. First-quartile values exceed 13.2 °C, median temperatures are consistently above 17.5 °C, and mean values reach up to 18.31 °C. Third-quartile values approach 23.6 °C, while standard deviations are the lowest in the region (5.20–6.04 °C), reflecting strong maritime buffering and reduced seasonal variability.

https://doi.org/10.3390/su18147193

Temperature extremes across all provinces are associated with a small number of regional-scale climatic events, most notably the cold outbreak of 10 January 1985, the cold spell of 7 January 2017, and the heatwave of 24 July 2007. Overall, the spatial distribution of summary statistics reveals a coherent north-to-south and inland-to-coast gradient, with colder and more variable conditions in northern inland Apulia and progressively warmer, more stable regimes toward southern coastal and peninsular areas.

The spatial distribution of the mean T2M for the first month of the record (January 1982) and the last month of the record (December 2023) is illustrated in Figure 3. In each panel, the colors shading represents the spatial variability of the mean temperature across Apulia, with cooler tones (purple to blue) indicating lower values and warmer tones (green to yellow) indicating higher values.

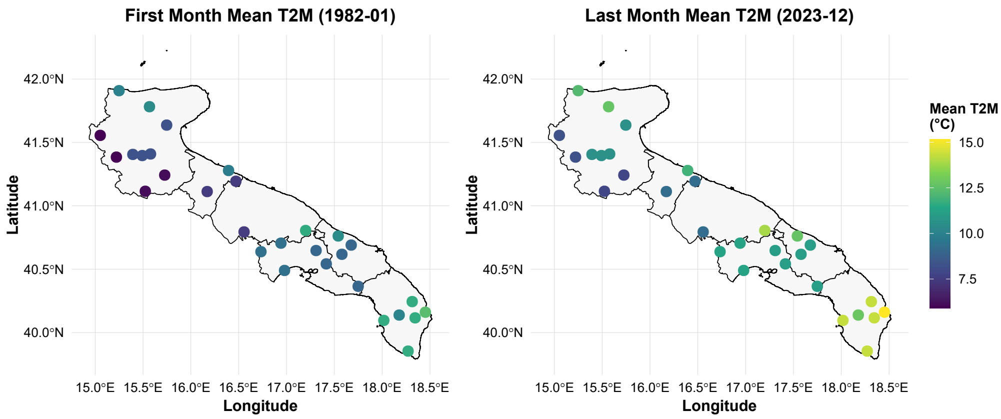
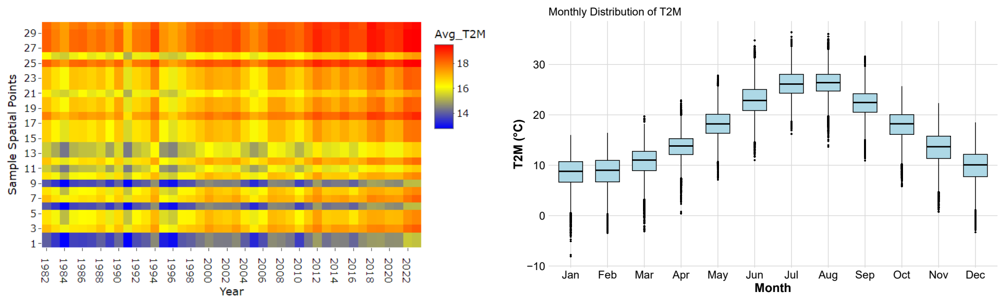


Figure 3. Spatial distribution of mean air temperature (T2M) in Apulia region for January 1982 (**left**) and December 2023 (**right**).

In January 1982, cooler temperatures dominate the northern inland portion of the region, particularly in the province of Foggia, while progressively warmer conditions are observed toward southern and coastal areas. This pattern reflects a pronounced north–south and inland–coastal gradient in winter mean temperature, driven by stronger continental influence inland and the moderating effect of the Adriatic and Ionian seas along the coast.

By December 2023, the spatial pattern shifts toward generally warmer conditions across most stations. The highest mean temperatures are observed in the southern coastal and peninsular areas, especially in the provinces of Taranto and Lecce. Although northern inland areas remain relatively cooler, they still exhibit higher mean temperatures compared with the ones in January 1982, indicating a regional-scale increase in winter-season mean temperature while preserving the underlying spatial gradients. The persistence of spatial contrasts alongside higher absolute temperatures in recent years highlights the coherent nature of warming across the Apulia region.

In Figure 4a the heatmap provides an effective visualization of spatiotemporal variation in temperature, enabling a clear view of regional trends and anomalies. In particular, this map presents the yearly average near-surface air temperature across the Apulia region from 1982 to 2023. The x-axis represents the temporal dimension (years), while the y-axis corresponds to the 30 sampled spatial locations within the study area. The color gradient, ranging from blue (lower temperatures) to red (higher temperatures), highlights a clear warming tendency over the study period.



(**a**) (**b**)

Figure 4. Temperature variability in the Apulia region represented by (**a**) a spatiotemporal heatmap of the mean air temperature (Avg_T2M), and (**b**) the monthly distribution of daily mean air temperature (T2M).

Earlier years are dominated by lower temperatures, particularly for stations in Foggia (S1–S10), whereas more recent years are dominated by higher temperatures. This indicates a progressive increase in temperature.

Figure 4a also highlights spatial heterogeneity across provinces. Stations in Foggia tend to exhibit greater variability and more frequent lower temperatures periods. In contrast, stations in southern provinces such as Lecce (S25–S30) and Taranto (S15–S20, S24) show more stable and consistently higher temperatures. Stations in BAT (S11–S13), Bari (S14, S18), and Brindisi (S21–S23) display intermediate behavior, with moderate variability and temperature levels. These differences suggest that geographic location plays an important role in shaping temperature patterns. Although some degree of interannual variability is evident, with occasional relatively cooler years, the overall pattern suggests a consistent warming trend across the majority of spatial sampling points.

Complementing the spatiotemporal heatmap, Figure 4b illustrates the monthly distribution of air temperature (T2M) across all spatial locations and years using box plots. This representation provides insights into the seasonal variability of temperature within the region. The results reveal a clear seasonal cycle, with the lowest median temperatures occurring during the winter months (December–February). Temperatures gradually increase during spring, reaching their maximum values during the summer months, particularly in July and August. The interquartile ranges are generally wider during the warmer months. This implies greater variability in temperature during this period. Furthermore, the presence of outliers indicates the occurrence of occasional extreme temperature events. As a whole, the box plot highlights the pronounced seasonal pattern of temperature variation in Apulia. This pattern is characterized by a cyclical increase from winter to summer followed by a gradual decline toward the end of the year.

4.5. Temperature Modeling and Fitting Validation

The primary objective of this study is to identify the most effective stochastic model, among the various models previously discussed for temperature prediction. The analyses proposed in this study have been performed using established RStudio (version 2023.09.1) libraries, namely `statistics` for the Holt–Winters model, `forecast` for the SARIMA and NNAR models, and `nnet` for the feed-forward ANN. The elaborations have been carried out using R functions, as detailed in Appendix A. Each model was trained using daily temperature data and validated over a one-year test period to ensure temporal consistency and robust comparison.

The Holt–Winters additive model has decomposed the temperature time series into level, trend, and seasonal components through recursive exponential smoothing controlled by the parameters α (level), β (trend), and γ (seasonal). These parameters have been estimated automatically from the training data by minimizing the sum of squared one-step-ahead forecast errors, with each parameter constrained between 0 and 1. Across all stations, the estimated level smoothing parameter α ranges between 0.86 and 0.87, the trend smoothing parameter β converges to zero, and the seasonal smoothing parameter γ tends to one, indicating consistent parameter behavior throughout the study region.

The additive formulation has been selected because the seasonal temperature structure remained approximately stable over time, with no evidence of systematic changes in seasonal amplitude. This stability is supported by the previously presented Figure 4b, which shows consistent temporal and spatial temperature patterns during the study period, indicating that long-term changes primarily manifest as shifts in the mean level rather than changes in the seasonal structure. As a result, the Holt–Winters model effectively captures the long-term trend and annual oscillation but lacks flexibility in representing complex, non-linear temperature variations. Furthermore, it tends to underperform when irregularities or abrupt temperature changes occur. Although conceptually simple and computationally efficient, its limitations become apparent when compared with models capable of learning non-linear dependencies.

The SARIMA model has been configured as SARIMA$(3, 0, 0)(0, 1, 0)_{[365]}$ to account for both autoregressive and seasonal structures in the daily temperature data. The non-seasonal component has included three autoregressive terms ($p = 3$), no differencing ($d = 0$), and no moving average component ($q = 0$), while the seasonal component has included one seasonal differencing term ($D = 1$) with a periodicity of 365 days. This specification has allowed the model to effectively represent short-term autocorrelation and remove recurring yearly effects in the data. The model parameters have been estimated using maximum likelihood optimization, and diagnostic tests confirmed that the residuals were approximately a white noise process, indicating a satisfactory model fit. However, due to its inherent linear structure, the SARIMA model is unable to capture non-linear and multi-scale temperature variations, which are characteristic of meteorological data influenced by complex environmental processes.

The ANN model has been designed to capture non-linear relationships among lagged temperature values. Using the `nnet` package, the ANN architecture has incorporated 14 lagged input neurons, a hidden layer with 5 neurons, and a single output neuron corresponding to the predicted temperature value. The hidden layer has employed a sigmoid activation function to introduce non-linearity, and a linear activation function has been applied at the output layer to conform to the continuous nature of temperature data. Model training has been performed using the gradient descent backpropagation algorithm, iterating up to 500 epochs to minimize the loss function. The ANN has demonstrated strong ability to model short-term fluctuations and non-linear dependencies; however, this model has not explicitly included seasonal lag structures. Consequently, it has been less effective in reproducing the annual periodicity observed in the data, which has slightly limited its long-term forecast accuracy compared to NNAR.

The NNAR model has combined the flexibility of neural networks with the memory structure of autoregressive processes. The selected configuration, NNAR$(8,1,5)_{[365]}$, consists of eight non-seasonal lagged inputs ($p = 8$), one seasonal lag ($P = 1$) representing the annual cycle with a periodicity of 365 days, and five neurons in the hidden layer. The model has been implemented using the `nnetar` function from the `forecast` package. A sigmoid activation function has been applied in the hidden layer, while a linear activation function has been used in the output layer, allowing the model to represent both non-linear relationships and continuous-valued outputs. To enhance robustness and reduce sensitivity to random weight initialization, an ensemble of 20 independently trained networks has been averaged for each spatial location.

https://doi.org/10.3390/su18147193

The specification of NNAR(8,1,5)$_{[365]}$ has been determined using time-series diagnostics rather than information criteria such as AIC or BIC, as these are not directly suitable for neural network autoregressive models. The ACF and PACF plots of the original daily temperature series have indicated a clear annual seasonal pattern, supporting the inclusion of one seasonal lag ($P = 1$) with a frequency of 365 days. This behavior has been further confirmed by the monthly distribution of air temperature (T2M) shown in Figure 4b, where box plots across all spatial locations and years have demonstrated a stable cyclical pattern, with higher temperatures during summer months and lower values during winter months.

Following seasonal adjustment, the ACF and PACF of the transformed series have been examined to identify the appropriate non-seasonal lag order. The autocorrelation structure gradually diminishes after approximately lag 8, suggesting that it $p = 8$ provides a suitable representation of the short-term temporal dependence.

The number of hidden neurons ($k = 5$) has been selected according to the standard heuristic implemented in the `nnetar()` framework, where the hidden-layer size is approximately half of the total number of input nodes plus one. With eight non-seasonal lags and one seasonal lag, the model contains nine lagged inputs, resulting in an optimal hidden layer of approximately five neurons.

This modeling strategy has ensured that the NNAR structure can be supported by both statistical diagnostics and the inherent seasonal characteristics of the dataset. Consequently, the model can represent short-term variability, annual seasonality, and non-linear temperature dynamics more effectively than the competing approaches.

The quantitative evaluation of the model performance is presented in Table 2. The results have confirmed the superiority of the NNAR model, which achieved the lowest RMSE (2.3535) and MAE (1.8332) and the highest R^2 (0.8861), accounting for approximately 88.6% of the variance in the temperature data. The SARIMA model has recorded RMSE = 2.9371, MAE = 2.3565, and R^2 = 0.8087, demonstrating moderate accuracy but reduced flexibility in handling non-linear dynamics. The ANN has achieved RMSE = 2.4774, MAE = 1.9999, and R^2 = 0.7774, confirming its effectiveness for non-linear learning but its limitation in modeling long-term seasonality. The Holt–Winters additive model has exhibited the weakest performance with RMSE = 3.8525, MAE = 3.1905, and R^2 = 0.6915, indicating that its simplicity could not adequately represent complex temperature variability. The NNAR has been identified as the best-performing stochastic model. This model effectively captures both long-term trends and seasonality variations in temperature data, making it the optimal choice for future predictions.

Table 2. Performance metrics of different models for measuring average temperature.

Model	RMSE	MAE	R^2
Holt–Winters	3.8525	3.1905	0.6915
SARIMA	2.9371	2.3565	0.8087
ANN	2.4774	1.9999	0.7774
NNAR	**2.3535**	**1.8332**	**0.8861**

4.6. Spatiotemporal Temperature Predictions

In the previous section, the best-performing model has been identified. To generate spatially and temporally consistent temperature predictions, spatiotemporal kriging has been subsequently applied to model the residuals of the temperature data, thereby accounting for the spatial and temporal dependencies inherent in the dataset. Within this framework, the deterministic component of temperature variability, denoted as $T(\mathbf{s}, t)$, has been first estimated using the NNAR model, while the stochastic component has been forecasted through kriging to capture the remaining spatiotemporal structure. Hence, after estimating the deterministic component, the residuals $\epsilon(\mathbf{s}, t)$ have been obtained by removing both trend and seasonality from the observed temperature values, i.e.,

$$\epsilon(\mathbf{s}, t) = X(\mathbf{s}, t) - T(\mathbf{s}, t), \tag{11}$$

where $X(\mathbf{s}, t)$ represents the observed temperature at location $\mathbf{s}$ and time t. The residual term $\epsilon(\mathbf{s}, t)$, therefore, encapsulates the portion of the temperature variability that is not explained by the trend or seasonality components.

These residuals often exhibit spatial and temporal autocorrelation arising from unresolved local climatic factors and short-term variations. To capture these dependencies, spatiotemporal ordinary kriging (ST-OK) has been adopted, treating the residuals as a second-order stationary and isotropic process. The spatial and temporal dependencies have been characterized through empirical spatiotemporal variograms, which quantify the dependence structure across both dimensions.

In this framework, residuals are assumed to be stationary and isotropic, implying consistent behavior across space and time and the product-sum variogram model [81] is chosen to describe the spatiotemporal correlation behavior characterizing the residual component, i.e.,

$$\gamma_{ST}(\mathbf{h}_s, h_t) = \gamma_{ST}(\mathbf{h}_s, 0) + \gamma_{ST}(\mathbf{0}, h_t) - k \cdot \gamma_{ST}(\mathbf{h}_s, 0) \cdot \gamma_{ST}(\mathbf{0}, h_t) \tag{12}$$

where

- $\mathbf{h}_s = \mathbf{s}' - \mathbf{s}$ denotes the spatial lag vector,
- $h_t = t' - t$ denotes the temporal lag,
- $\gamma_{ST}(\mathbf{h}_s, 0)$ and $\gamma_{ST}(\mathbf{0}, h_t)$ are the spatial and temporal marginal variograms of the residual process, respectively,
- k is the interaction parameter estimated from the sample variogram as follows

$$k = \frac{s_s + s_t - s_{st}}{s_s \cdot s_t}, \tag{13}$$

with s_s, s_t, s_{st} the spatial, temporal, and spatiotemporal marginal sills, respectively. Parameter k is estimated according to this condition:

$$\max(s_s, s_t) \leq s_{st} < s_s + s_t \tag{14}$$

required for the negative definiteness of the variogram.

This formulation reflects the fact that spatial and temporal dependencies jointly influence the residual structure rather than acting independently. For temperature data, observations recorded at nearby locations tend to remain more similar over short temporal intervals, while this similarity gradually decreases as spatial or temporal separation increases. Consequently, the interaction parameter enables the model to represent space-time variability more flexibly. To characterize these dependence structures, the spatial and temporal marginal variograms are modeled using appropriate parametric variogram functions [10].

For the present case study, the marginal spatial variogram, which describes how residual variability changes across locations, is presented in Figure 5a, while the marginal temporal variogram, characterizing residual dependence over time, is shown in Figure 5b. In these figures, the horizontal axis represents spatial distance and time lag, respectively, whereas the vertical axis corresponds to the semivariogram values. Both marginal variograms were fitted using exponential models. Specifically, the spatial variogram is given by $\widehat{\gamma}(\mathbf{h}_s, 0) = \exp(\|\mathbf{h}_s\|; 30)$ in $\mathbb{R}^d$, whereas the temporal variogram is defined as $\widehat{\gamma}(\mathbf{0}, h_t) = \exp(|h_t|; 2)$ in $\mathbb{R}$. The fitted models indicate a practical spatial range of approximately 30 spatial units and a practical temporal range of about 2 days, beyond which the residual dependence becomes negligible. These findings provide evidence of substantial spatial and temporal correlation in the residual process and, therefore, justify the adoption of a spatiotemporal modeling framework.

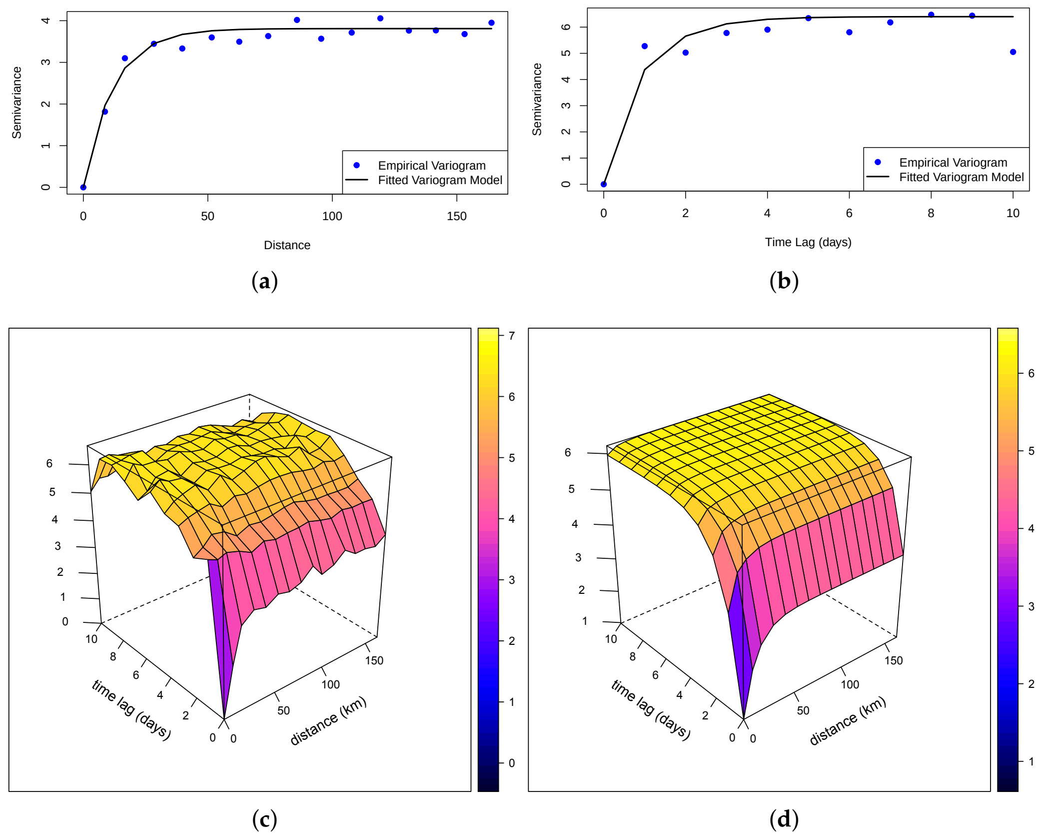


Figure 5. Spatiotemporal variogram surface (**c**,**d**) along with the spatial (**a**) and temporal (**b**) marginal variograms, displaying fitted models for the NNAR residuals.

The spatiotemporal variogram surface, shown in Figure 5c, provides an empirical representation of spatiotemporal dependence structure, showing how spatial and temporal correlation changes jointly across spatial distance and time lag.

To model this joint dependence, a product-sum variogram model in Equation (12) is fitted, and the resulting surface is shown in Figure 5d. The fitted variogram captures the main features of the sample variogram, including the decay of dependence over spatial and time lag as well as their interaction.

From Figure 5d, the fitted model produces a spatial sill $s_s = 0.2001$ and a temporal sill $s_t = 2.6511$. These values represent the total variability attributed to the spatial and temporal processes, respectively. The fitted model also includes an interaction parameter, $k = 6.3168$, which quantifies the degree of spatiotemporal coupling beyond the additive contribution of the spatial and temporal marginals. A positive value k indicates non-negligible interaction. This implies that spatial dependence varies with temporal separation and vice versa. It means they jointly influence the overall dependence structure rather than the two components acting independently. It is worth noting that the estimated sill parameters and the fitted value k satisfy the admissibility condition of the product-sum model presented in Equation (14). This ensures that the resulting spatiotemporal variogram $\gamma_{ST}(\mathbf{h}_s, h_t)$ is valid and positive definite.

After modeling the spatiotemporal variogram, ST-OK was carried out to predict the residuals at unobserved space-time locations (s_0, t_0). As is well known, ST-OK is a weighted interpolation technique that uses observed residuals at nearby locations to estimate the residuals at the target point [10,97]. In this application, residuals have been predicted six days ahead at each observed spatial location using ST-OK based on the fitted product-sum variogram model discussed above. Finally, the kriged residuals have been combined with the NNAR estimated trend and seasonal components to generate spatially and temporally coherent six day ahead temperature forecasts for 2024 across all sampled spatial locations.

Furthermore, to validate the model's ability, a two-day rolling forecast window has been applied to minimize forecast errors through residual correction, and a short-term temperature forecasting framework has been developed that iteratively refines predictions. The methodology has utilized an initial six-day subset of historical residuals, which have been subsequently incorporated into the forecasting process. After each iteration, new residuals have been computed by subtracting the observed temperature values from the corrected forecasts. These updated residuals have been merged into the dataset, ensuring dynamic model adaptation to evolving errors. The effectiveness of this approach has been assessed using relative Root Mean Square Error (r-RMSE). The distribution of r-RMSE values across sample spatial locations is presented in Figure 6 and indicates robust predictive performance, with a median r-RMSE of 0.1625, a minimum of 0.0808, and a maximum of 0.2483. These results demonstrate the model's ability to minimize forecast errors through residual correction and iterative refinement.

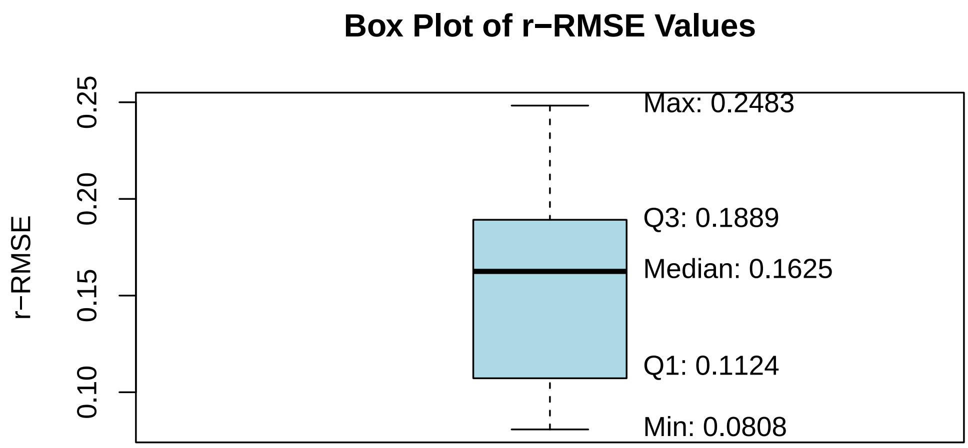


Figure 6. Box plot of r-RMSE of predicted temperature distribution.

Finally, temperature predictions have been spatially interpolated and clipped to the Apulia regional boundary to maintain geospatial precision. The predicted temperatures from 1–6 January 2024, are illustrated in Figure 7. The maps reveal a distinct north-to-south temperature gradient, with cooler temperatures in northern interior regions and warmer conditions in southern and coastal areas. The predicted temperature range (7.45 °C to 15.27 °C) aligns with established climatic patterns in Apulia, further validating the accuracy of the proposed methodology. The smooth spatial transitions and temporal consistency across consecutive days highlight the effectiveness of integrating NNAR-based forecasting with spatiotemporal kriging. This approach enables reliable short-term temperature prediction while preserving spatial coherence.

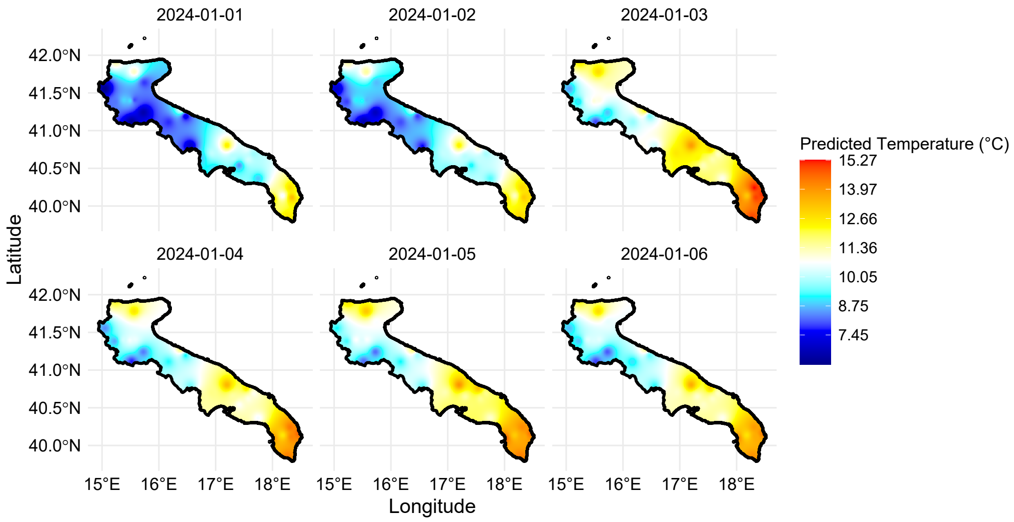


Figure 7. Predicted temperature maps of the Apulia region from 1 January 2024, to 6 January 2024.

5. Conclusions

This study demonstrated the effectiveness and applicability of spatiotemporal geostatistics and its associated tools for temperature modeling. A stochastic model for temperature prediction in the Apulia region of Italy was developed by integrating the NNAR model with spatial kriging, enhancing forecasting accuracy by simultaneously capturing temporal trends and spatial dependencies. The advantage of this hybrid approach lies in its ability to model large-scale temperature trends, seasonal variations, and localized residual fluctuations.

Among the four candidate models (Holt–Winters, SARIMA, ANNs, and NNAR), the NNAR model emerged as the most effective temporal model, achieving the highest predictive accuracy with the lowest RMSE (2.3535) and MAE (1.8332), along with the highest R^2 (0.8861). NNAR effectively captures short-term fluctuations, seasonal patterns, and non-linear relationships, outperforming traditional time series methods such as SARIMA, which primarily captures linear seasonal dependencies, and ANNs, which focus on non-linear relationships but lacks autoregressive components.

To account for spatial dependencies, kriging has been employed to interpolate temperature variations across different locations. Additionally, a generalized product-sum variogram model has been selected to describe the spatiotemporal correlation structure of the residuals. The adequacy of the model was assessed through graphical analysis of the marginal spatial and temporal variograms, where the empirical variogram surface was fitted and the corresponding spatiotemporal dependence structure was evaluated. The comparison between the hybrid spatiotemporal NNAR-kriging model and purely temporal models confirmed that incorporating spatial dependencies significantly enhances predictive performance. Future research may further extend the proposed framework through comparisons with advanced deep learning approaches such as LSTM/GRU, XGBoost, Random Forest, and Transformer-based architectures. The integration of such methods with multivariate geostatistical ones [98] may provide additional capability for capturing complex non-linear and long-range spatiotemporal dependencies in environmental forecasting applications.

Beyond methodological improvements, the results highlight the practical relevance of high-resolution spatiotemporal temperature forecasting in climate-sensitive regions such as Apulia. Accurate temperature predictions can support more efficient agricultural planning, improved water resource allocation, and better management of climate-related risks. In particular, the ability to capture localized temperature variability provides valuable information for adapting to increasing climate variability and extreme events.

Importantly, the proposed framework is data-driven and not region-specific in structure. This supports its potential applicability to geographical areas beyond the Apulia region. However, successful generalization depends on the availability and quality of input data, the spatial density of observations, and the similarity of climatic conditions. Regions with markedly different environmental characteristics may require recalibration of model parameters and variogram structures.

Regarding scalability, the framework performs adequately on moderately large datasets. However, the computational cost of the kriging component increases substantially with the number of spatial locations. For very large spatiotemporal datasets, such as those comprising thousands to millions of observations, conventional kriging methods are no longer feasible. This is due to its $\mathcal{O}(n^3)$ time complexity and $\mathcal{O}(n^2)$ memory requirements. The temporal dimension further compounds these demands.

Future research will focus on cross-regional model validation and will also explore more computationally efficient approaches. These include sparse covariance approximations and low-rank representations for large spatial datasets [99–101], covariance tapering techniques [102–104], complex-valued covariance models [105,106] as well as hybrid machine learning–geostatistical frameworks that combine the strengths of statistical learning and spatial dependence modeling [51,107,108]. Such strategies are expected to improve computational scalability while preserving predictive accuracy in large-scale, high-resolution environmental applications.

Overall, despite the computational challenges noted above, the proposed hybrid framework offers a reliable tool for understanding and predicting temperature dynamics, contributing to more informed decision-making and strengthening adaptive capacity in regions facing environmental and climatic challenges.

Author Contributions: N.I.: Writing, review and editing, Writing original draft, Methodology, Formal analysis, Data curation, Conceptualization. M.P.: Writing, review and editing, Writing original draft, validation, Conceptualization. S.D.I.: Validation and Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the ICSC–National Research Center in High Performance Computing, Big Data and Quantum Computing, supported by the European Union–NextGenerationEU, Project PNRR-HPC (Project code: CN00000013; CUP: C83C22000560007).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors thank the Editor and the referees for their constructive suggestions during the reviewing process.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. R Code

Listing A1. R script used to generate the descriptive statistics of the spatiotemporal temperature datat.

```r
# Load necessary libraries
library(sf)
library(ggplot2)
library(cowplot)
library(rnaturalearth)
library(rnaturalearthdata)
```

```
library(dplyr)
library(ggspatial)
library(grid)
library(patchwork)
library(viridis)
library(leaps)
library(plotly)

italy_shape <- ne_countries(scale = "medium", country = "Italy", returnclass = "sf")

setwd("C:/Users/Nouman␣Iqbal/Downloads/ITA_adm")

apulia_shape <- st_read("./ITA_adm1.shp") %>%
  filter(NAME_1 == "Apulia")

data <- read.csv("E:/e␣folder/Phd␣research␣work/longitude␣and␣latitude␣RN/data␣T2m/data␣T2M.csv")

points_sf <- data %>%
  select(Latitude, Longitude) %>%
  distinct() %>%
  st_as_sf(coords = c("Longitude", "Latitude"), crs = st_crs(apulia_shape)) %>%
  st_join(apulia_shape, join = st_within) %>%
  filter(!is.na(NAME_1)) %>%
  arrange(st_coordinates(.)[, 1]) %>%
  mutate(Label = paste0("S", row_number()))

italy_map <- ggplot() +
  geom_sf(data = italy_shape, fill = "gray90", color = "black") +
  geom_sf(data = apulia_shape, fill = "orange", color = "black") +
  annotate("rect", xmin = 14.93497, xmax = 18.52069,
           ymin = 39.78959, ymax = 42.22653,
           color = "red", fill = NA, size = 1) +
  labs(title = "Italy", x = "Longitude", y = "Latitude") +
  theme_minimal()

apulia_map <- ggplot() +
  geom_sf(data = apulia_shape, fill = "lightgray", color = "black") +
  geom_sf(data = points_sf,
          aes(geometry = geometry,
          color = "Sample␣Spatial␣Points"),
          size = 2) +
  geom_text(data = points_sf %>%
              mutate(x = st_coordinates(.)[,1],
                     y = st_coordinates(.)[,2]),
            aes(x = x, y = y, label = Label),
            vjust = -1, color = "blue", size = 2)

combined_map <- ggdraw() +
  draw_plot(italy_map, 0, 0.5, 1, 0.5) +
  draw_plot(apulia_map, 0.3, 0, 0.7, 0.5)

```

```r
print(combined_map)

#R script used for spatial data processing and station statistics in the Apulia~region

italy_shape <- ne_countries(scale = "medium", country = "Italy", returnclass = "sf")

setwd("C:/Users/Nouman Iqbal/Downloads/ITA_adm")

apulia_shape <- st_read("./ITA_adm2.shp") %>%
  filter(NAME_1 == "Apulia")

data <- read.csv("E:/e folder/Phd research work/longitude and latitude RN/data T2m/data T2M.csv")

points_sf <- data %>%
  select(Latitude, Longitude) %>%
  distinct() %>%
  st_as_sf(coords = c("Longitude", "Latitude"), crs = st_crs(apulia_shape)) %>%
  st_join(apulia_shape, join = st_within) %>%
  filter(!is.na(NAME_2)) %>%
  arrange(st_coordinates(.)[, 1]) %>%
  mutate(Label = paste0("S", row_number()))

data <- data %>% mutate(date = as.Date(date))

data_sf <- data %>%
  st_as_sf(coords = c("Longitude", "Latitude"),
           crs = st_crs(apulia_shape),
           remove = FALSE)

data_joined <- st_join(data_sf, apulia_shape, join = st_within) %>%
  filter(!is.na(NAME_2))

data_joined_tbl <- data_joined %>% st_drop_geometry()

stations <- data_joined_tbl %>%
  distinct(Longitude, Latitude, NAME_2) %>%
  arrange(Longitude, Latitude) %>%
  mutate(Label = paste0("S", row_number()))

data_labeled <- data_joined_tbl %>%
  left_join(stations, by~= c("Longitude", "Latitude", "NAME_2"))

station_summary <- data_labeled %>%
  group_by(Label, NAME_2, Longitude, Latitude) %>%
  summarise(
    n_days  = n(),
    start   = min(date, na.rm = TRUE),
    end     = max(date, na.rm = TRUE),
    mean    = mean(T2M, na.rm = TRUE),
    sd      = sd(T2M, na.rm = TRUE),
```

```
    min      = min(T2M, na.rm = TRUE),
    q25      = quantile(T2M, 0.25, na.rm = TRUE),
    median   = median(T2M, na.rm = TRUE),
    q75      = quantile(T2M, 0.75, na.rm = TRUE),
    max      = max(T2M, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  arrange(Label)

station_summary


#R script used to compute monthly mean T2M and visualize first and last month temperature distribution in~Apulia

data_labeled2 <- data_labeled %>%
  mutate(
    date = as.Date(date),
    year_month = floor_date(date, "month")
  )

monthly_mean <- data_labeled2 %>%
  group_by(Label, NAME_2, Longitude, Latitude, year_month) %>%
  summarise(T2M_month_mean = mean(T2M, na.rm = TRUE), .groups = "drop")

first_last_month <- monthly_mean %>%
  group_by(Label) %>%
  summarise(
    first_month = min(year_month),
    last_month  = max(year_month),
    .groups = "drop"
  )

first_last_vals <- monthly_mean %>%
  inner_join(first_last_month, by~= "Label") %>%
  filter(year_month %in% c(first_month, last_month)) %>%
  mutate(period = ifelse(year_month == first_month, "First month", "Last month")) %>%
  select(Label, NAME_2, Longitude, Latitude, period, year_month, T2M_month_mean)

stations_fl_sf <- st_as_sf(first_last_vals, coords = c("Longitude", "Latitude"), crs = 4326)

lims <- range(stations_fl_sf$T2M_month_mean, na.rm = TRUE)

base_theme <- theme_minimal() +
  theme(
    legend.position = "right",
    plot.title = element_text(face = "bold"),
    axis.title = element_blank()
  )

p_first <- ggplot() +
```

```
  geom_sf(data = apulia_shape, fill = "white", colour = "black", linewidth = 0.3) +
  geom_sf(
    data = subset(stations_fl_sf, period == "First month"),
    aes(color = T2M_month_mean),
    size = 3
  ) +
  scale_color_viridis_c(
    name = "Mean T2M (°C)",
    limits = lims
  ) +
  labs(
    title = paste0("First month mean T2M (", format(min(first_last_vals$year_month), "%Y-%m"), ")")
  ) +
  base_theme

p_last <- ggplot() +
  geom_sf(data = apulia_shape, fill = "white", colour = "black", linewidth = 0.3) +
  geom_sf(
    data = subset(stations_fl_sf, period == "Last month"),
    aes(color = T2M_month_mean),
    size = 3
  ) +
  scale_color_viridis_c(
    name = "Mean T2M (°C)",
    limits = lims
  ) +
  labs(
    title = paste0("Last month mean T2M (", format(max(first_last_vals$year_month), "%Y-%m"), ")")
  ) +
  base_theme

(p_first | p_last) + plot_layout(guides = "collect")

#R script used to generate an interactive heatmap of yearly average temperature across spatial sampling~points

data$date <- as.Date(data$date)
data$Year <- format(data$date, "%Y")

yearly_data <- data %>%
  group_by(Year, Longitude) %>%
  summarize(Average_Temperature = mean(T2M, na.rm = TRUE)) %>%
  ungroup()

longitude_mapping <- yearly_data %>%
  select(Longitude) %>%
  distinct() %>%
  arrange(Longitude) %>%
  mutate(Longitude_Label = seq_along(Longitude))

yearly_data <- yearly_data %>%
```

```r
  left_join(longitude_mapping, by = "Longitude")

plot_ly(
  yearly_data,
  x = ~Year,
  y = ~Longitude_Label,
  z = ~Average_Temperature,
  type = "heatmap",
  colorscale = list(
    list(0, "blue"),
    list(0.5, "yellow"),
    list(1, "red")
  )
) %>%
layout(
  title = "",
  xaxis = list(title = "Year"),
  yaxis = list(
    title = "Sample Spatial Points",
    tickvals = longitude_mapping$Longitude_Label,
    ticktext = longitude_mapping$Longitude_Label
  )
)

#R script used to generate the monthly boxplot distribution of T2M

data$Month <- factor(
  data$MO,
  levels = 1:12,
  labels = c("Jan","Feb","Mar","Apr","May","Jun",
             "Jul","Aug","Sep","Oct","Nov","Dec")
)

data_t2m <- data %>%
  select(Month, T2M)

ggplot(data_t2m, aes(x = Month, y = T2M)) +
  geom_boxplot(
    fill = "lightblue",
    outlier.size = 0.6
  ) +
  theme_bw() +
  theme(
    axis.text.x = element_text(size = 10),
    axis.title = element_text(size = 11)
  ) +
  labs(
    x = "Month",
    y = "T2M (°C)",
    title = "Monthly Distribution of T2M"
  )
```

Listing A2. R script used to perform modelling, forecasting, and performance evaluation of daily temperature data across spatial locations.

```
#ARIMA modleing
# Load necessary libraries
library(dplyr)
library(forecast)
library(tseries)
library(nnet)
data <- read.csv("E:/e folder/Phd research work/longitude and latitude RN/data T2m/data T2M.csv")

data$date <- as.Date(data$date)

# Remove leap day
data <- data %>% filter(!(format(date, "%m-%d") == "02-29"))

unique_latitudes <- unique(data$Latitude)

residuals_analysis_list <- list()
performance_metrics_list <- list()
forecast_valid_list <- list()

for (lat in unique_latitudes) {

  lat_data <- data %>% filter(Latitude == lat)
  lat_data <- lat_data %>% arrange(date)

  ts_data <- ts(lat_data$T2M, frequency = 365, start = c(1982, 1))

  n.valid <- 365
  n.train <- length(ts_data) -~n.valid

  train_data <- window(ts_data, end = c(1982 + (n.train - 1) / 365, 1))
  valid_data <- window(ts_data, start = c(1982 + n.train / 365, 1))

  arima_model <- Arima(train_data,
                       order = c(3,0,0),
                       seasonal = list(order = c(0,1,0), period = 365))

  arima_forecast <- forecast(arima_model, n.valid)
  forecast_valid <- arima_forecast$mean

  residuals <- c(residuals(arima_model), valid_data - forecast_valid)

  rmse_train <- sqrt(mean(residuals(arima_model)^2, na.rm = TRUE))
  mae_train <- mean(abs(residuals(arima_model)), na.rm = TRUE)

  rmse_valid <- sqrt(mean((valid_data - forecast_valid)^2, na.rm = TRUE))
```

```
  mae_valid <- mean(abs(valid_data - forecast_valid), na.rm = TRUE)

  metrics <- data.frame(
    Latitude = lat,
    RMSE_Train = rmse_train,
    MAE_Train = mae_train,
    RMSE_Valid = rmse_valid,
    MAE_Valid = mae_valid,
    Train_Length = n.train,
    Valid_Length = n.valid
  )

  validation_dates <- lat_data$date[(n.train + 1):length(ts_data)]

  residuals_analysis_list[[as.character(lat)]] <- data.frame(
    date = lat_data$date,
    Latitude = lat,
    residuals = as.vector(residuals)
  )

  performance_metrics_list[[as.character(lat)]] <-~metrics

  forecast_valid_list[[as.character(lat)]] <- data.frame(
    date = validation_dates,
    Latitude = lat,
    actual = as.vector(valid_data),
    forecast_valid = as.vector(forecast_valid),
    residuals = as.vector(valid_data - forecast_valid)
  )
}

residuals_analysis_df <- bind_rows(residuals_analysis_list, .id = "Latitude")
performance_metrics_df <- bind_rows(performance_metrics_list)
forecast_valid_df <- bind_rows(forecast_valid_list, .id = "Latitude")


#R script used to implement Holt--Winters exponential smoothing for temperature forecasting and model performance evaluation across spatial~locations


hw_forecast_list <- list()
residuals_analysis_list <- list()
performance_metrics_list <- list()
hw_valid_list <- list()

for (lat in unique_latitudes) {

  lat_data <- data %>% filter(Latitude == lat)

  lat_data <- lat_data %>% arrange(date)

```

```
ts_data <- ts(lat_data$T2M, frequency = 365, start = c(1982, 1))

n.valid <- 365
n.train <- length(ts_data) -~n.valid

train_data <- window(ts_data, end = c(1982 + (n.train - 1) / 365, 1))
valid_data <- window(ts_data, start = c(1982 + n.train / 365, 1))

hw_model <- HoltWinters(train_data, seasonal = "additive")

hw_forecast <- as.vector(forecast(hw_model, h = n.valid)$mean)

hw_fitted <- c(as.vector(hw_model$fitted[, "xhat"]), hw_forecast)

if (length(hw_fitted) < length(ts_data)) {
  hw_fitted <- c(hw_fitted, rep(NA, length(ts_data) - length(hw_fitted)))
}

residuals <- ts_data -~hw_fitted

residuals_train <- residuals[1:n.train]
residuals_valid <- valid_data -~hw_forecast

rmse_train <- sqrt(mean(residuals_train^2, na.rm = TRUE))
mae_train <- mean(abs(residuals_train), na.rm = TRUE)

ss_total_train <- sum((train_data - mean(train_data, na.rm = TRUE))^2, na.rm = TRUE)
ss_res_train <- sum(residuals_train^2, na.rm = TRUE)
r2_train <- 1 - (ss_res_train / ss_total_train)

rmse_valid <- sqrt(mean((valid_data - hw_forecast)^2, na.rm = TRUE))
mae_valid <- mean(abs(valid_data - hw_forecast), na.rm = TRUE)

ss_total_valid <- sum((valid_data - mean(valid_data, na.rm = TRUE))^2, na.rm = TRUE)
ss_res_valid <- sum((valid_data - hw_forecast)^2, na.rm = TRUE)
r2_valid <- 1 - (ss_res_valid / ss_total_valid)

metrics <- data.frame(
  Latitude = lat,
  RMSE_Train = rmse_train,
  MAE_Train = mae_train,
  R2_Train = r2_train,
  RMSE_Valid = rmse_valid,
  MAE_Valid = mae_valid,
  R2_Valid = r2_valid,
  Train_Length = n.train,
  Valid_Length = n.valid
)
```

```

  validation_dates <- lat_data$date[(n.train + 1):length(lat_data$
     date)]

  hw_forecast_list[[as.character(lat)]] <- data.frame(
    date = lat_data$date,
    Latitude = lat,
    hw_fitted = hw_fitted
  )

  residuals_analysis_list[[as.character(lat)]] <- data.frame(
    date = lat_data$date,
    Latitude = lat,
    residuals = residuals
  )

  performance_metrics_list[[as.character(lat)]] <-~metrics

  hw_valid_list[[as.character(lat)]] <- data.frame(
    date = validation_dates,
    Latitude = lat,
    actual = as.vector(valid_data),
    hw_forecast = hw_forecast,
    residuals = residuals_valid
  )
}

hw_forecast_df <- bind_rows(hw_forecast_list, .id = "Latitude")
residuals_analysis_df <- bind_rows(residuals_analysis_list, .id =
    "Latitude")
performance_metrics_df <- bind_rows(performance_metrics_list)
hw_valid_df <- bind_rows(hw_valid_list, .id = "Latitude")

#R script used to implement a neural network model with lagged
    inputs for temperature forecasting and validation across
    spatial~locations

create_lags <- function(data, lags) {
  lagged_data <- embed(data, lags + 1)
  lagged_data <- as.data.frame(lagged_data)
  colnames(lagged_data) <- c("Target", paste0("Lag_", 1:lags))
  return(lagged_data)
}

residuals_analysis_list <- list()
performance_metrics_list <- list()
forecast_valid_list <- list()

for (lat in unique_latitudes) {

  lat_data <- data %>% filter(Latitude == lat)
  lat_data <- lat_data %>% arrange(date)

  t2m_data <- lat_data$T2M

```

```
lags <- 14
lagged_data <- create_lags(t2m_data, lags)

n.valid <- 365
n.train <- nrow(lagged_data) -~n.valid

train_data <- lagged_data[1:n.train, ]
valid_data <- lagged_data[(n.train + 1):nrow(lagged_data), ]

nnet_model <- nnet(
  x = as.matrix(train_data[, -1]),
  y = train_data$Target,
  size = 5,
  linout = TRUE,
  maxit = 500
)

predictions <- predict(nnet_model, newdata = as.matrix(valid_data[, -1]))

residuals <- c(
  train_data$Target - predict(nnet_model, as.matrix(train_data[, -1])),
  valid_data$Target - predictions
)

rmse_valid <- sqrt(mean((valid_data$Target - predictions)^2, na.rm = TRUE))
mae_valid <- mean(abs(valid_data$Target - predictions), na.rm = TRUE)

ss_total_valid <- sum((valid_data$Target - mean(valid_data$Target, na.rm = TRUE))^2, na.rm = TRUE)
ss_res_valid <- sum((valid_data$Target - predictions)^2, na.rm = TRUE)
r2_valid <- 1 - (ss_res_valid / ss_total_valid)

metrics <- data.frame(
  Latitude = lat,
  RMSE_Valid = rmse_valid,
  MAE_Valid = mae_valid,
  R2_Valid = r2_valid,
  Train_Length = n.train,
  Valid_Length = n.valid
)

residuals_analysis_list[[as.character(lat)]] <- data.frame(
  date = lat_data$date[-(1:lags)],
  Latitude = lat,
  residuals = residuals
)

performance_metrics_list[[as.character(lat)]] <-~metrics

forecast_valid_list[[as.character(lat)]] <- data.frame(
```

```
    date = lat_data$date[(n.train + lags + 1):nrow(lat_data)],
    Latitude = lat,
    actual = valid_data$Target,
    forecast_valid = predictions,
    residuals = valid_data$Target - predictions
  )
}

residuals_analysis_df <- bind_rows(residuals_analysis_list, .id = "Latitude")
performance_metrics_df <- bind_rows(performance_metrics_list)
forecast_valid_df <- bind_rows(forecast_valid_list, .id = "Latitude")

#R script used to implement Neural Network Autoregression (NNAR) for temperature forecasting and model evaluation across spatial~locations

residuals_analysis_list <- list()
performance_metrics_list <- list()
forecast_valid_list <- list()

for (lat in unique_latitudes) {

  lat_data <- data %>% filter(Latitude == lat)

  lat_data <- lat_data %>% arrange(date)

  ts_data <- ts(lat_data$T2M, frequency = 365, start = c(1982, 1))

  n.valid <- 365
  n.train <- length(ts_data) -~n.valid

  train_data <- window(ts_data, end = c(1982 + (n.train - 1) / 365, 1))
  valid_data <- window(ts_data, start = c(1982 + n.train / 365, 1))

  nnetar_model <- nnetar(
    train_data,
    p = 8,
    P = 1,
    size = 5,
    repeats = 20
  )

  nnetar_forecast <- forecast(nnetar_model, h = n.valid)
  forecast_valid <- nnetar_forecast$mean

  residuals <- c(residuals(nnetar_model), valid_data - forecast_valid)

  rmse_train <- sqrt(mean(residuals(nnetar_model)^2, na.rm = TRUE))
  mae_train <- mean(abs(residuals(nnetar_model)), na.rm = TRUE)
```

```

  ss_total_train <- sum((train_data - mean(train_data, na.rm = TRUE))^2, na.rm = TRUE)
  ss_res_train <- sum(residuals(nnetar_model)^2, na.rm = TRUE)
  r2_train <- 1 - (ss_res_train / ss_total_train)

  rmse_valid <- sqrt(mean((valid_data - forecast_valid)^2, na.rm = TRUE))
  mae_valid <- mean(abs(valid_data - forecast_valid), na.rm = TRUE)

  ss_total_valid <- sum((valid_data - mean(valid_data, na.rm = TRUE))^2, na.rm = TRUE)
  ss_res_valid <- sum((valid_data - forecast_valid)^2, na.rm = TRUE)
  r2_valid <- 1 - (ss_res_valid / ss_total_valid)

  metrics <- data.frame(
    Latitude = lat,
    RMSE_Train = rmse_train,
    MAE_Train = mae_train,
    R2_Train = r2_train,
    RMSE_Valid = rmse_valid,
    MAE_Valid = mae_valid,
    R2_Valid = r2_valid,
    Train_Length = n.train,
    Valid_Length = n.valid
  )

  validation_dates <- lat_data$date[(n.train + 1):length(ts_data)]

  residuals_analysis_list[[as.character(lat)]] <- data.frame(
    date = lat_data$date,
    Latitude = lat,
    residuals = as.vector(residuals)
  )

  performance_metrics_list[[as.character(lat)]] <-~metrics

  forecast_valid_list[[as.character(lat)]] <- data.frame(
    date = validation_dates,
    Latitude = lat,
    actual = as.vector(valid_data),
    forecast_valid = as.vector(forecast_valid),
    residuals = as.vector(valid_data - forecast_valid)
  )
}

residuals_analysis_df <- bind_rows(residuals_analysis_list, .id = "Latitude")
performance_metrics_df <- bind_rows(performance_metrics_list)
forecast_valid_df <- bind_rows(forecast_valid_list, .id = "Latitude")
```

Listing A3. R script used to compute and fit spatiotemporal variograms for residual temperature data using the product-sum covariance model.

```r

# Load necessary libraries
library(strucchange)
library(dplyr)
library(tidyr)
library(lubridate)
library(forecast)
library(ggplot2)
library(spacetime)
library(covatest)
library(sp)
library(gstat)
library(RColorBrewer)
library(STRbook)
library(xts)
library(TSA)

# Load data
data <- read.csv("E:/e folder/Phd research work/longitude and 
    latitude RN/data T2m/NNEAR/final_result.csv")
data$date <- as.Date(data$date)

spat_part2 <- data %>% select(Longitude, Latitude) %>% distinct()
coordinates(spat_part2) <- ~Longitude + Latitude

# Set CRS to WGS84
proj4string(spat_part2) <- CRS("+proj=longlat +datum=WGS84")

temp_part2 <- unique(data$date)

STObj3 <- STFDF(
  sp = spat_part2,
  time = temp_part2,
  data = select(data, -date, -Longitude, -Latitude, -T2M)
)

proj4string(STObj3) <- CRS("+proj=longlat +datum=WGS84")

# Compute spatiotemporal variogram
st_variogram <- variogramST(
  residuals ~ 1,
  data = STObj3,
  tlags = 0:10,
  cutoff = 170,
  width = 170/15
)

plot(st_variogram, wireframe = TRUE,
     zlab = NULL,
     xlab = list("distance (km)", rot = 30),
     ylab = list("time lag (days)", rot = -35),
     scales = list(arrows = FALSE, z = list(distance = 5)))
```

```
st_variogram[1, 2:3] <-~0

# Spatial variogram
vario_sp <- filter(st_variogram, timelag == 0)

plot(vario_sp$dist, vario_sp$gamma,
     pch = 19, col = "blue",
     xlab = "Distance",
     ylab = "Semivariance")

manual_model <- vgm(psill = 3.8, model = "Exp", range = 12, nugget = 0.1)

fitted_values <- variogramLine(manual_model, dist_vector = vario_sp$dist)

lines(fitted_values$dist, fitted_values$gamma,
      col = "black", lwd = 2)

legend("bottomright",
       legend = c("Empirical Variogram", "Fitted Variogram Model"),
       col = c("blue", "black"),
       pch = c(19, NA),
       lty = c(NA, 1),
       lwd = c(NA, 2))

# Temporal variogram
vario_te <- filter(st_variogram, spacelag == 0)

vario_te$timelag_numeric <- as.numeric(gsub(" days", "", vario_te$timelag))

plot(vario_te$timelag_numeric, vario_te$gamma,
     pch = 19, col = "blue",
     xlab = "Time Lag (days)",
     ylab = "Semivariance")

temporal_model_EXP <- vgm(psill = 5.5, model = "Exp", range = 1, nugget = 0.9)

fitted_temporal_EXP <- variogramLine(
  temporal_model_EXP,
  dist_vector = vario_te$timelag_numeric
)

lines(fitted_temporal_EXP$dist,
      fitted_temporal_EXP$gamma,
      col = "black", lwd = 2)

legend("bottomright",
       legend = c("Empirical Variogram", "Fitted Variogram Model"),
       col = c("blue", "black"),
```

```
       pch = c(19, NA),
       lty = c(NA, 1),
       lwd = c(NA, 2))

# Product-sum spatiotemporal model
spatial_model <- vgm(psill = 3.8, model = "Exp", range = 12, nugget = 0.1)
temporal_model <- vgm(psill = 5.7, model = "Exp", range = 1, nugget = 0.9)

product_sum_model_updated <- vgmST(
  "productSum",
  space = spatial_model,
  time = temporal_model,
  k = 0.015
)

vgmf.metric <- fit.StVariogram(st_variogram, product_sum_model_updated)

plot(st_variogram, vgmf.metric,
     wireframe = TRUE,
     zlab = NULL,
     xlab = list("distance (km)", rot = 30),
     ylab = list("time lag (days)", rot = -35),
     scales = list(arrows = FALSE, z = list(distance = 5)))
```

Listing A4. R script used to compute the relative root mean square error (r-RMSE) for each spatial location and visualize its distribution using a box plot.

```

original_data <- read.csv("E:/e folder/Phd research work/longitude and latitude RN/data T2m/origional data of 6 day.csv")
forecast_results_df <- read.csv("E:/e folder/Phd research work/longitude and latitude RN/data T2m/NNEAR/forecast_6days_2024.csv")

error_data <- original_data %>%
  inner_join(forecast_results_df, by~= c("Longitude", "Latitude", "date"))
location_r_rmse <- error_data %>%
  group_by(Longitude, Latitude) %>%
  summarise(
    numerator = sum((T2M - final_forecast_T2M)^2, na.rm = TRUE),
    denominator = sum(T2M^2, na.rm = TRUE),
    r_RMSE = sqrt(numerator / denominator),
    .groups = "drop"
  )

boxplot(location_r_rmse$r_RMSE,
        main = "Box Plot of r-RMSE Values",
        ylab = "r-RMSE",
        col = "lightblue",
        border = "black",
        boxwex = 0.4)
```

```
text(x = 1.1, y = summary_stats["Min."],
     labels = paste("Min:", round(summary_stats["Min."], 4)), pos = 4)

text(x = 1.1, y = summary_stats["1st Qu."],
     labels = paste("Q1:", round(summary_stats["1st Qu."], 4)), pos = 4)

text(x = 1.1, y = summary_stats["Median"],
     labels = paste("Median:", round(summary_stats["Median"], 4)), pos = 4)

text(x = 1.1, y = summary_stats["3rd Qu."],
     labels = paste("Q3:", round(summary_stats["3rd Qu."], 4)), pos = 4)

text(x = 1.1, y = summary_stats["Max."],
     labels = paste("Max:", round(summary_stats["Max."], 4)), pos = 4)
```

Listing A5. R script used to perform interpolation of forecasted temperature and generate spatial temperature maps over the Apulia region.

```

library(raster)

forecast_results_df$date <- as.Date(
  as.numeric(forecast_results_df$date),
  origin = "1970-01-01"
)

setwd("C:/Users/Nouman Iqbal/Downloads/ITA_adm")
apulia <- st_read("./apulia_shapefile.shp")

apulia <- st_transform(apulia, crs = 4326)

bbox <- st_bbox(apulia)

grid_x <- seq(bbox["xmin"], bbox["xmax"], length.out = 500)
grid_y <- seq(bbox["ymin"], bbox["ymax"], length.out = 500)

prediction_grid <- expand.grid(Longitude = grid_x, Latitude = grid_y)

coordinates(prediction_grid) <- ~Longitude + Latitude
proj4string(prediction_grid) <- CRS("+proj=longlat +datum=WGS84")

unique_dates <- unique(forecast_results_df$date)

all_idw_results <- list()

for (d in unique_dates) {

  df_filtered <- forecast_results_df %>% filter(date == d)

```

```
  df_filtered$final_forecast_T2M <- as.numeric(
    as.character(df_filtered$final_forecast_T2M)
  )

  coordinates(df_filtered) <- ~Longitude + Latitude
  proj4string(df_filtered) <- CRS("+proj=longlat +datum=WGS84")

  idw_result <- idw(final_forecast_T2M ~ 1,
                    df_filtered,
                    prediction_grid,
                    idp = 2)

  idw_sf <- st_as_sf(
    as.data.frame(idw_result),
    coords = c("Longitude", "Latitude"),
    crs = st_crs(apulia)
  )

  idw_inside <- st_crop(idw_sf, apulia)
  idw_inside <- st_intersection(idw_inside, apulia)

  idw_inside <- idw_inside %>%
    mutate(
      Longitude = st_coordinates(.)[,1],
      Latitude = st_coordinates(.)[,2],
      date = d
    ) %>%
    as.data.frame()

  all_idw_results[[as.character(d)]] <- idw_inside
}

final_idw_data <- bind_rows(all_idw_results)

final_idw_data$date <- as.Date(final_idw_data$date, origin = "1970-01-01")

final_idw_data <- final_idw_data %>% distinct()

final_idw_data$var1.pred <- as.numeric(
  as.character(final_idw_data$var1.pred)
)

value_range <- range(final_idw_data$var1.pred, na.rm = TRUE)

temperature_palette <- c(
  "darkblue","blue","cyan","white","yellow","orange","red"
)

legend_breaks <- round(
  seq(value_range[1], value_range[2], length.out = 8), 2
)

ggplot(final_idw_data) +
  geom_tile(
```

```
    aes(x = Longitude, y = Latitude, fill = var1.pred)
  ) +
  geom_contour(
    aes(x = Longitude, y = Latitude, z = var1.pred),
    color = "black",
    bins = 8,
    breaks = seq(value_range[1], value_range[2], length.out = 10)
  ) +
  geom_sf(
    data = apulia,
    fill = NA,
    color = "black",
    linewidth = 1
  ) +
  scale_fill_gradientn(
    colors = temperature_palette,
    limits = value_range,
    breaks = legend_breaks,
    labels = sprintf("%.2f", legend_breaks),
    na.value = "gray90",
    guide = guide_colorbar(
      title = "Predicted Temperature (°C)",
      barwidth = 1.5,
      barheight = 10,
      title.position = "top",
      label.position = "right"
    )
  ) +
  coord_sf() +
  facet_wrap(~date, ncol = 3) +
  labs(title = "") +
  theme_minimal() +
  theme(legend.position = "right")
```

References

1. Pandolfi, A.; Pitzalis, M.; Nardi, M.; Lorenzetti, M.C. Evaluation of climatic risk for irrigation planning. In *Proceedings of the Symposium on Scheduling of Irrigation for Vegetable Crops Under Field Conditions*; International Society for Horticultural Science (ISHS): Leuven, Belgium, 1990.
2. Ronco, P.; Zennaro, F.; Torresan, S.; Critto, A.; Santini, M.; Trabucco, A.; Zollo, A.L.; Galluccio, G.; Marcomini, A. A risk assessment framework for irrigated agriculture under climate change. *Adv. Water Resour.* **2017**, *110*, 562–578. [CrossRef]
3. Chronopoulos, K.; Christoforides, E.; Kamoutsis, A.; Panagiotou, I. Development of a Spatial Methodology for Minimum Temperature Estimation for Early Frost Management in Agricultural Areas of Central Macedonia. *Environ. Earth Sci. Proc.* **2025**, *35*, 55. [CrossRef]
4. Kapur, B.; Steduto, P.; Todorovic, M. Prediction of climatic change for the next 100 years in the Apulia region, Southern Italy. *Ital. J. Agron.* **2007**, *2*, 365–371. [CrossRef]
5. Peli, M.; Rapuzzi, C.; Barontini, S.; Ranzi, R. Application of Benfratello's method to estimate the spatio-temporal variability of the irrigation deficit in a Mediterranean semiarid climate. *Hydrol. Res.* **2023**, *54*, 451–474. [CrossRef]
6. Jalilian, A.; Cappello, C.; Palma, M.; De Iaco, S. Generalized additive model with dynamic coefficients for spatiotemporal ozone predictions. *Environmetrics* **2025**, *37*, e70076.
7. Wild, M.; Behm, S.; Beck, C.; Cyrys, J.; Schneider, A.; Wolf, K.; Haupt, H. Mapping the Time-Varying Spatial Heterogeneity of Temperature Processes over the Urban Landscape of Augsburg, Germany. *Urban Clim.* **2022**, *43*, 101160. [CrossRef]
8. Castillo-Mateo, J.; Lafuente, M.; Asín, J.; Cebrián, A.C.; Gelfand, A.E.; Abaurrea, J. Spatial Modeling of Day-Within-Year Temperature Time Series: An Examination of Daily Maximum Temperatures in Aragón, Spain. *J. Agric. Biol. Environ. Stat.* **2022**, *27*, 487–505. [CrossRef]

9. Braga, A.; Laurini, M. Spatial Heterogeneity in Climate Change Effects across Brazilian Biomes. *Sci. Rep.* **2024**, *14*, 16414. [CrossRef] [PubMed]
10. Cressie, N.; Wikle, C.K. *Statistics for Spatio-Temporal Data*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
11. Galton, A. Time flies but space does not: Limits to the spatialisation of time. *J. Pragmat.* **2011**, *43*, 695–703. [CrossRef]
12. Kyriakidis, P.C.; Journel, A.G. Geostatistical Space–Time Models: A Review. *Math. Geol.* **1999**, *31*, 651–684. [CrossRef]
13. Li, J.; Heap, A.D. A Review of Comparative Studies of Spatial Interpolation Methods in Environmental Sciences: Performance and Impact Factors. *Ecol. Inform.* **2011**, *6*, 228–241. [CrossRef]
14. Cressie, N.; Huang, H.C. Classes of nonseparable, spatio-temporal stationary covariance functions. *J. Am. Stat. Assoc.* **1999**, *94*, 1330–1339. [CrossRef]
15. De Cesare, L.; Myers, D.; Posa, D. Product-sum covariance for space-time modeling: An environmental application. *Environ. Off. J. Int. Environ. Soc.* **2001**, *12*, 11–23.
16. Gneiting, T. Nonseparable, stationary covariance functions for space–time data. *J. Am. Stat. Assoc.* **2002**, *97*, 590–600. [CrossRef]
17. Ma, C. Spatio-temporal covariance functions generated by mixtures. *Math. Geol.* **2002**, *34*, 965–975. [CrossRef]
18. Ma, C. Families of spatio-temporal stationary covariance models. *J. Stat. Plan. Inference* **2003**, *116*, 489–501. [CrossRef]
19. Kolovos, G.; Christakos, G.; Hristopulos, D.T.; Serre, M.L. Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. *Adv. Water Resour.* **2004**, *27*, 815–830. [CrossRef]
20. De Iaco, S.; Myers, D.E.; Posa, D. Nonseparable space-time covariance models: Some parametric families. *Math. Geol.* **2002**, *34*, 23–42. [CrossRef]
21. Gneiting, T.; Genton, M.G.; Guttorp, P. Geostatistical space-time models, stationarity, separability, and full symmetry. *Monogr. Stat. Appl. Probab.* **2006**, *107*, 151. [CrossRef]
22. Stein, M.L. Space–time covariance functions. *J. Am. Stat. Assoc.* **2005**, *100*, 310–321. [CrossRef]
23. Gregori, P.; Porcu, E.; Mateu, J.; Sasvári, Z. On potentially negative space time covariances obtained as sum of products of marginal ones. *Ann. Inst. Stat. Math.* **2008**, *60*, 865–882.
24. De Iaco, S.; Posa, D. Characteristics of some isotropic covariance models with negative values. *Spat. Stat.* **2025**, *68*, 100905. [CrossRef]
25. Alegria, A.; Emery, X. Versatile parametric classes of covariance functions that interlace anisotropies and hole effects. *Stat. Sin.* **2026**, *36*, 127–146.
26. De Iaco, S.; Palma, M.; Posa, D. On the use of difference-based covariance models. *Stat. Methods Appl.* **2026**, *35*, 181–205. [CrossRef]
27. De Iaco, S.; Posa, D.; Palma, M. Complex-Valued Random Fields for Vectorial Data: Estimating and Modeling Aspects. *Math. Geosci.* **2013**, *45*, 557–573. [CrossRef]
28. De Iaco, S. The cgeostat software for analyzing complex-valued random fields. *J. Stat. Softw.* **2017**, *79*, 1–32. [CrossRef]
29. De Iaco, S.; Palma, M.; Posa, D. Covariance functions and models for complex-valued random fields. *Stoch. Environ. Res. Risk Assess.* **2003**, *17*, 145–156. [CrossRef]
30. Rodrigues, A.; Diggle, P. A Class of Convolution-Based Models for Spatio-Temporal Processes with Non-Separable Covariance Structure. *Scand. J. Stat.* **2010**, *37*, 553–567. [CrossRef]
31. De Iaco, S.; Posa, D. Positive and negative non-separability for space-time covariance models. *J. Stat. Plan. Inf.* **2013**, *143*, 378–391. [CrossRef]
32. Carrera-Hernández, J.; Gaskin, S. Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. *J. Hydrol.* **2007**, *336*, 231–249. [CrossRef]
33. Hengl, T.; Heuvelink, G.B.; Perčec Tadić, M.; Pebesma, E.J. Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theor. Appl. Climatol.* **2012**, *107*, 265–277.
34. Kilibarda, M.; Hengl, T.; Heuvelink, G.B.; Gräler, B.; Pebesma, E.; Perčec Tadić, M.; Bajat, B. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *J. Geophys. Res. Atmos.* **2014**, *119*, 2294–2313. [CrossRef]
35. Ahmadi, F.; Nazeri Tahroudi, M.; Mirabbasi, R.; Khalili, K.; Jhajharia, D. Spatiotemporal trend and abrupt change analysis of temperature in Iran. *Meteorol. Appl.* **2018**, *25*, 314–321.
36. Cao, X.; Okhrin, O.; Odening, M.; Ritter, M. Modelling spatio-temporal variability of temperature. *Comput. Stat.* **2015**, *30*, 745–766. [CrossRef]
37. Bruno, F.; Guttorp, P.; Sampson, P.D.; Cocchi, D. A simple non-separable, non-stationary spatiotemporal model for ozone. *Environ. Ecol. Stat.* **2009**, *16*, 515–529.
38. Hoogland, T.; Heuvelink, G.B.; Knotters, M. Mapping water-table depths over time to assess desiccation of groundwater-dependent ecosystems in the Netherlands. *Wetlands* **2010**, *30*, 137–147. [CrossRef]
39. Massarelli, C.; Campanale, C.; Triozzi, M.; Uricchio, V.F. Dynamics of pesticides in surface water bodies by applying data mining to spatiotemporal big data. A case study for the Puglia Region. *Ecol. Inform.* **2023**, *78*, 102342. [CrossRef]
40. Dumitrescu, A.; Birsan, M.V.; Manea, A. Spatio-temporal interpolation of sub-daily (6 h) precipitation over Romania for the period 1975–2010. *Int. J. Climatol.* **2016**, *36*, 1331–1343.

41. Kilibarda, M.; Tadić, M.P.; Hengl, T.; Luković, J.; Bajat, B. Global geographic and feature space coverage of temperature data in the context of spatio-temporal interpolation. *Spat. Stat.* **2015**, *14*, 22–38. [CrossRef]
42. Gething, P.W.; Atkinson, P.M.; Noor, A.M.; Gikandi, P.W.; Hay, S.I.; Nixon, M.S. A local space–time kriging approach applied to a national outpatient malaria data set. *Comput. Geosci.* **2007**, *33*, 1337–1350. [CrossRef] [PubMed]
43. Gething, P.W.; Noor, A.M.; Gikandi, P.W.; Hay, S.I.; Nixon, M.S.; Snow, R.W.; Atkinson, P.M. Developing geostatistical space–time models to predict outpatient treatment burdens from incomplete national data. *Geogr. Anal.* **2008**, *40*, 167–188. [CrossRef] [PubMed]
44. Giorgi, E.; Diggle, P.J.; Snow, R.W.; Noor, A.M. Geostatistical methods for disease mapping and visualisation using data from spatio-temporally referenced prevalence surveys. *Int. Stat. Rev.* **2018**, *86*, 571–597. [CrossRef] [PubMed]
45. Macharia, P.M.; Giorgi, E.; Noor, A.M.; Waqo, E.; Kiptui, R.; Okiro, E.A.; Snow, R.W. Spatio-temporal analysis of *Plasmodium falciparum* prevalence to understand the past and chart the future of malaria control in Kenya. *Malar. J.* **2018**, *17*, 340. [CrossRef] [PubMed]
46. Noor, A.M.; Gething, P.W.; Alegana, V.A.; Patil, A.P.; Hay, S.I.; Muchiri, E.; Juma, E.; Snow, R.W. The risks of malaria infection in Kenya in 2009. *BMC Infect. Dis.* **2009**, *9*, 180. [CrossRef] [PubMed]
47. Noor, A.M.; Kinyoki, D.K.; Mundia, C.W.; Kabaria, C.W.; Mutua, J.W.; Alegana, V.A.; Fall, I.S.; Snow, R.W. The changing risk of *Plasmodium falciparum* malaria infection in Africa: 2000–10: A spatial and temporal analysis of transmission intensity. *Lancet* **2014**, *383*, 1739–1747. [CrossRef] [PubMed]
48. Muto, S.; Sugasawa, S.; Suzuki, M. Hedonic real estate price estimation with the spatiotemporal geostatistical model. *J. Spat. Econom.* **2023**, *4*, 10. [CrossRef]
49. Huang, Y.; Yang, B.; Ren, X.; Lu, Y.; Lan, M.; Gong, X. Spatio-temporal cokriging crime predictions using social media data: A multi-type case study in San Jose, California. *Comput. Urban Sci.* **2025**, *5*, 72. [CrossRef]
50. De Iaco, S.; Hristopulos, D.; Lin, G. Statistical Deep Learning for Spatial and Spatiotemporal Data. *Math. Geosci.* **2022**, *54*, 459–465. [CrossRef]
51. Lan, S. Learning temporal evolution of spatial dependence with generalized spatiotemporal gaussian process models. *J. Mach. Learn. Res.* **2022**, *23*, 1–53. [CrossRef]
52. Nag, P.; Sun, Y.; Reich, B. Spatio-temporal DeepKriging for interpolation and probabilistic forecasting. *Spat. Stat.* **2023**, *57*, 100773. [CrossRef]
53. Huang, J.; Lu, C.; Huang, D.; Qin, Y.; Xin, F.; Sheng, H. A spatial interpolation method for meteorological data based on a hybrid kriging and machine learning approach. *Int. J. Climatol.* **2024**, *44*, 5371–5380. [CrossRef]
54. Mohammadpour, M.; Roshan, H.; Arashpour, M.; Masoumi, H. Machine learning assisted kriging to capture spatial variability in petrophysical property modelling. *Mar. Pet. Geol.* **2024**, *167*, 106967. [CrossRef]
55. Jeong, M.; Koo, H. Evaluating spatio-temporal kriging with machine learning considering the sources of spatio-temporal variation. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 224. [CrossRef]
56. Duerr, I.; Merrill, H.R.; Wang, C.; Bai, R.; Boyer, M.; Dukes, M.D.; Bliznyuk, N. Forecasting urban household water demand with statistical and machine learning methods using large space-time data: A Comparative study. *Environ. Model. Softw.* **2018**, *102*, 29–38. [CrossRef]
57. Requia, W.J.; Vicedo-Cabrera, A.M.; Amini, H.; Schwartz, J.D. Short-term air pollution exposure and mortality in Brazil: Investigating the susceptible population groups. *Environ. Pollut.* **2024**, *340*, 122797. [CrossRef] [PubMed]
58. Wu, C.-D.; Zeng, Y.-T.; Lung, S.-C.C. A hybrid kriging/land-use regression model to assess $PM_{2.5}$ spatial-temporal variability. *Sci. Total Environ.* **2018**, *645*, 1456–1464. [CrossRef] [PubMed]
59. Hsu, C.-Y.; Zeng, Y.-T.; Chen, Y.-C.; Chen, M.-J.; Lung, S.-C.C.; Wu, C.-D. Kriging-based land-use regression models that use machine learning algorithms to estimate the monthly BTEX concentration. *Int. J. Environ. Res. Public Health* **2020**, *17*, 6956. [CrossRef] [PubMed]
60. Ding, X.; Zhao, Y.; Fan, Y.; Li, Y.; Ge, J. Machine learning-assisted mapping of city-scale air temperature: Using sparse meteorological data for urban climate modeling and adaptation. *Build. Environ.* **2023**, *234*, 110211. [CrossRef]
61. Zhang, G.P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **2003**, *50*, 159–175. [CrossRef]
62. Khashei, M.; Bijari, M. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Appl. Soft Comput.* **2011**, *11*, 2664–2675. [CrossRef]
63. Li, S.; Griffith, D.A.; Shu, H. Temperature prediction based on a space–time regression-kriging model. *J. Appl. Stat.* **2019**, *47*, 1168–1190. [CrossRef] [PubMed]
64. Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*, 2nd ed.; Holden-Day: San Francisco, CA, USA, 1976.
65. Mahajan, S.; Chen, L.J.; Tsai, T.C. Short-term PM2. 5 forecasting using exponential smoothing method: A comparative analysis. *Sensors* **2018**, *18*, 3223. [CrossRef] [PubMed]

66. Gibson, K. The Application of Machine Learning for Groundwater Level Prediction in the Steenkoppies Compartment of the Gauteng and North West Dolomite Aquifer, South Africa. Ph.D. Thesis, University of the Free State, Bloemfontein, South Africa, 2020.
67. Silva, E.S.; Hassani, H.; Heravi, S.; Huang, X. Forecasting tourism demand with denoised neural networks. *Ann. Tour. Res.* **2019**, *74*, 134–154. [CrossRef]
68. Hamzaçebi, C. Improving Artificial Neural Networks' Performance in Seasonal Time Series Forecasting. *Inf. Sci.* **2008**, *178*, 4550–4559. [CrossRef]
69. Pala, Z.; Atici, R. Forecasting sunspot time series using deep learning methods. *Sol. Phys.* **2019**, *294*, 50. [CrossRef]
70. Gardner, E.S. Exponential Smoothing: The State of the Art—Part II. *Int. J. Forecast.* **2006**, *22*, 637–666. [CrossRef]
71. Adji, M.W.; Hakim, D.K. Comparison of Double and Triple Exponential Smoothing Methods for Rainfall Prediction. *J. E-Komtek (Elektro-Komput.)* **2024**, *8*, 64–76. [CrossRef]
72. Zhang, G.; Patuwo, B.E.; Hu, M.Y. Forecasting with artificial neural networks: The state of the art. *Int. J. Forecast.* **1998**, *14*, 35–62.
73. Anders, U.; Korn, O. Model selection in neural networks. *Neural Netw.* **1999**, *12*, 309–323. [CrossRef] [PubMed]
74. Cappello, C.; De Iaco, S.; Posa, D. Testing the type of non-separability and some classes of space-time covariance function models. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 17–35. [CrossRef]
75. Cappello, C.; De Iaco, S.; Posa, D. covatest: An R Package for Selecting a Class of Space-Time Covariance Functions. *J. Stat. Softw.* **2020**, *94*, 1–14. [CrossRef]
76. Rouhani, S.; Hall, T.J. Space-time kriging of groundwater data. In Proceedings of the Geostatistics: Proceedings of the Third International Geostatistics Congress, Avignon, France, 5–9 September 1988; pp. 639–650.
77. Dimitrakopoulos, R.; Luo, X. Spatiotemporal modelling: Covariances and ordinary kriging systems. In *Proceedings of the Geostatistics for the Next Century: An International Forum in Honour of Michel David's Contribution to Geostatistics, Montreal*; Springer: Berlin/Heidelberg, Germany, 1993; pp. 88–93.
78. Rodríguez-Iturbe, I.; Mejía, J.M. The design of rainfall networks in time and space. *Water Resour. Res.* **1974**, *10*, 713–728. [CrossRef]
79. Varouchakis, E.A.; Hristopulos, D.T. Comparison of spatiotemporal variogram functions based on a sparse dataset of groundwater level variations. *Spat. Stat.* **2019**, *34*, 100245. [CrossRef]
80. Guo, L.; Lei, L.; Zeng, Z.-C.; Zou, P.; Liu, D.; Zhang, B. Evaluation of spatio-temporal variogram models for mapping XCO_2 using satellite observations: A case study in China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *8*, 376–385.
81. De Iaco, S.; Myers, D.E.; Posa, D. Space–time analysis using a general product–sum model. *Stat. Probab. Lett.* **2001**, *52*, 21–28. [CrossRef]
82. Menegatto, V.A. Strictly positive definite kernels on the Hilbert sphere. *Appl. Anal.* **1994**, *55*, 91–101. [CrossRef]
83. Menegatto, V.A.; Oliveira, C.P.; Peron, A.P. Strictly positive definite kernels on subsets of the complex plane. *Comput. Math. Appl.* **2006**, *51*, 1233–1250. [CrossRef]
84. De Iaco, S.; Myers, D.; Posa, D. Strict positive definiteness of a product of covariance functions. *Commun. Stat.-Theory Methods* **2011**, *40*, 4400–4408. [CrossRef]
85. Gneiting, T. Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **2013**, *19*, 1327–1349. [CrossRef]
86. De Iaco, S.; Posa, D.; Cappello, C.; Maggio, S. Isotropy, symmetry, separability and strict positive definiteness for covariance functions: A critical review. *Spat. Stat.* **2019**, *29*, 89–108. [CrossRef]
87. Dumelle, M.; Ver Hoef, J.M.; Fuentes, C.; Gitelman, A. A linear mixed model formulation for spatio-temporal random processes with computational advances for the product, sum, and product–sum covariance functions. *Spat. Stat.* **2021**, *43*, 100510. [CrossRef]
88. De Cesare, L.; Myers, D.; Posa, D. Estimating and modeling space–time correlation structures. *Stat. Probab. Lett.* **2001**, *51*, 9–14. [CrossRef]
89. Fisher, R.A. On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond. Ser. A Contain. Pap. A Math. Phys. Character* **1922**, *222*, 309–368. [CrossRef]
90. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]
91. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [CrossRef]
92. Hodson, T.O. Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev. Discuss.* **2022**, *15*, 5481–5487. [CrossRef]
93. Roseto, R.; Dellino, P.; Capolongo, D. Spatial distribution and trend analysis of extreme rainfall time series in Apulia region (Italy). *Phys. Geogr. Quat. Dyn.* **2024**, *46*, 163–177.
94. IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013; 1535p.
95. Giannakopoulos, C.; Bindi, M.; Moriondo, M.; Le Sager, P.; Tin, T. *Climate Change Impacts in the Mediterranean Resulting from a 2 °C Global Temperature Rise*; WWF: Gland, Switzerland, 2005; 75p.

96. Fischer, G.; Tubiello, F.N.; van Velthuizen, H.; Wiberg, D.A. Climate change impacts on irrigation water requirements: Effects of mitigation, 1990–2080. *Technol. Forecast. Soc. Change* **2007**, *74*, 1083–1107. [CrossRef]
97. Cressie, N. The origins of kriging. *Math. Geol.* **1990**, *22*, 239–252. [CrossRef]
98. De Iaco, S.; Palma, M.; Posa, D. Choosing suitable linear coregionalization models for spatio-temporal data. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 1419–1434. [CrossRef]
99. Banerjee, S.; Gelfand, A.E.; Finley, A.O.; Sang, H. Gaussian predictive process models for large spatial data sets. *J. R. Stat. Soc. Ser. B* **2008**, *70*, 825–848. [CrossRef]
100. Cressie, N.; Johannesson, G. Fixed rank kriging for very large spatial data sets. *J. R. Stat. Soc. Ser. B* **2008**, *70*, 209–226. [CrossRef]
101. Katzfuss, M. A multi-resolution approximation for massive spatial datasets. *J. Am. Stat. Assoc.* **2017**, *112*, 201–214. [CrossRef]
102. Gneiting, T. Compactly supported correlation functions. *J. Multivar. Anal.* **2002**, **83**, 493–508. [CrossRef]
103. Furrer, R.; Genton, M.G.; Nychka, D.W. Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Stat.* **2006**, *15*, 502–523. [CrossRef]
104. Kaufman, C.G.; Schervish, M.J.; Nychka, D.W. Covariance tapering for likelihood-based estimation in large spatial datasets. *J. Am. Stat. Assoc.* **2008**, *103*, 1545–1555. [CrossRef]
105. De Iaco, S. New spatio-temporal complex covariance functions for vectorial data through positive mixtures. *Stoch. Environ. Res. Risk Assess.* **2022**, *36*, 2769–2787. [CrossRef]
106. De Iaco, S. Spatio-temporal generalized complex covariance models based on convolution. *Comput. Stat. Data Anal.* **2023**, *183*, 107709. [CrossRef]
107. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [CrossRef] [PubMed]
108. Wikle, C.K.; Zammit-Mangion, A. Statistical deep learning for spatial and spatio-temporal processes. *Annu. Rev. Stat. Appl.* **2023**, *10*, 247–270. [CrossRef]