www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 18(1), 2027, 354-371. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2027.v18i1.1516

Meteorological Time Series Interpolation: A Literature Review of Classical, Machine Learning and Deep Learning Approaches

Víctor Manuel Landassuri Moreno[1], Alejandro Moreno Martínez[1], Saturnino Job Morales Escobar[1], Saúl Lazcano Salas[1], Maricela Quintana López[1]

[1]Centro Universitario UAEM Valle de México Universidad Autónoma del Estado de México Boulevard Universitario S/N

Atizapán de Zaragoza, Estado de México, México

{vmlandassurim, almorenoma8218, sjmoralese, slazcanos, mquintanal}@uaemex.mx

Abstract. Meteorological time series are frequently affected by missing observations caused by sensor failures, communication problems, maintenance activities, or data acquisition errors. Subsequent analyses, including forecasting, climate monitoring, hydrological assessment, and environmental decision-making, may therefore be biased or weakened. In this review, interpolation methods for meteorological time series are reviewed, including classical, statistical, machine learning, and deep learning approaches. Emphasis is placed on the distinction between prediction and interpolation, since both tasks are often treated as equivalent. The literature is also organized according to reconstruction strategy, including recursive, bidirectional, probabilistic or generative, and hybrid or evolutionary approaches. Current challenges, evaluation practices, open gaps, and future trends for meteorological data reconstruction are discussed.

Keywords: Meteorological time series; interpolation; missing data; imputation; neural networks; deep learning; forecasting; climate data.

Article Info
Received Jul 09, 2026
Accepted Aug 11, 2026

1 Introduction

Meteorological time series are widely used to describe atmospheric behavior and support decisions in agriculture, water management, air quality monitoring, disaster prevention, and climate analysis. Complete, error-free records are the exception rather than the rule. Sensor malfunction, communication failures, maintenance windows, calibration drift, and plain human error routinely interrupt the acquisition process (Junninen et al., 2004; Che et al., 2018; Du et al., 2023).

Once observations start dropping out, the reliability of subsequent analyses is compromised: forecasting models lose accuracy, statistical summaries skew, temporal patterns blur, and decision-support systems become less trustworthy. This cascading effect has established the reconstruction of missing values — through interpolation or imputation — into the role of a near-mandatory preprocessing step in meteorological data analysis (Rubin, 1976; Lepot et al., 2017; Moritz and Bartz-Beielstein, 2017).

For decades, gaps in meteorological records were filled using classical interpolation techniques — linear interpolation, nearest neighbor, polynomial fitting, splines, inverse distance weighting, kriging. These methods are cheap to run and easy to justify, but they start to break down once the data turns nonlinear, nonstationary, noisy, or just messier than the models assume (Lepot et al., 2017; Li and Heap, 2014).

Machine learning and deep learning approaches have subsequently been introduced. Artificial neural networks, support vector regression, random forest imputation, recurrent networks, LSTM and GRU architectures, attention-based models, generative models, and transformers have all been applied in an effort to capture the kind of complex temporal dependencies that classical methods tend to miss (Smola and Schölkopf, 2004; Stekhoven and Bühlmann, 2012; Che et al., 2018; Cao et al., 2018; Du et al., 2023; Tashiro et al., 2021).

A handful of reviews already cover related ground, though none quite from the angle taken here. Lepot et al. (2017) give a broad account of interpolation in time series; Li and Heap (2014) focus on spatial interpolation in environmental sciences; and deep learning for time series imputation shows up in a few recent surveys (Fang and Wang, 2020; Wang et al., 2024; Kazijevs and Samad, 2023). Useful as these are, none of them bring together classical temporal methods, geostatistical reconstruction, machine learning, deep learning, neuroevolutionary models, and reconstruction strategies specifically for meteorological gap filling. This review addresses that gap, focusing specifically on the meteorological case rather than on interpolation or imputation as a general problem.

This review makes four main contributions: a) it clarifies the conceptual distinction between time series prediction and interpolation; b) it reviews classical, statistical, machine learning, and deep learning methods used for meteorological time series interpolation; c) it organizes a taxonomy of the methodological strategies reported in the literature; and d) it identifies current limitations, research gaps, and emerging trends in meteorological missing data reconstruction.

One research question guides the review: RQ1) how should prediction, interpolation, imputation, and gap filling be distinguished when missing values occur within meteorological time series?; RQ2) Which methodological families have been used, or could be transferred, to meteorological time series interpolation, and what assumptions are required?; RQ3) How can existing methods be organized according to reconstruction strategy rather than only by algorithmic architecture?; and RQ4) Under which data conditions, such as gap length, missingness pattern, spatial availability, multivariate dependence, and uncertainty requirements, are different method families likely to be more appropriate?

2 Prediction and Interpolation in Time Series

A time series may be represented as an ordered set of observations collected at successive time instants. In the univariate case, this sequence may be denoted as $X=x_1,x_2,...,x_t$ (equation 1), whereas a multivariate time series may be written as a matrix X ∈ $\mathbb{R}$^(T × D), where T denotes the number of time steps and D denotes the number of variables (Che et al., 2018; Du et al., 2023). When entries are unavailable, a binary mask M is commonly used to identify which observations are present and which ones must be estimated.

2.1 Time Series Prediction

Time series prediction, or forecasting, is usually defined as the estimation of future observations from past and present information (De Gooijer and Hyndman, 2006; Ben Taieb et al., 2012). Given a sequence of observations:

$$x_1,\ x_2,\ \ldots,\ x_t,\ . \tag{1}$$

the objective is to estimate one or more future values:

$$x_{t+1},\ x_{t+2},\ \ldots,\ x_{t+h},\ . \tag{2}$$

where h denotes the prediction horizon. In this task, future observations are not available to the model, and the temporal dynamics learned from historical data must be extrapolated.

Forecasting has been widely used in meteorology to estimate future temperature, precipitation, wind speed, humidity, solar radiation, atmospheric pressure, and air pollutant concentrations. Autoregressive models, artificial neural networks, recurrent neural networks, LSTM, GRU, and transformer-based models are among the methods commonly considered for this purpose (Hochreiter and Schmidhuber, 1997; Cho et al., 2014; Zhou et al., 2021).

2.2 Time Series Interpolation

Time series interpolation is concerned with the estimation of missing observations located within an already observed temporal sequence (Lepot et al., 2017; Moritz and Bartz-Beielstein, 2017). In this setting, missing values are surrounded, at least in principle, by known observations before and after the missing interval. For example:

$$x_1, x_2, x_3, ?, ?, x_6, x_7.. \quad (3)$$

Unlike forecasting, interpolation is not necessarily based on extrapolation beyond the available record. Instead, missing values are reconstructed from the temporal context around the gap. In meteorological applications, interpolation is commonly used to recover records lost because of sensor failures, communication interruptions, invalid measurements, or quality-control removal (Junninen et al., 2004; Lepot et al., 2017).

2.3 Main Differences Between Prediction and Interpolation

Although prediction and interpolation are related, different problems are addressed by them. Future values beyond the last available observation are estimated in prediction, whereas missing values inside the observed series are reconstructed in interpolation. This distinction is important because a method that performs well for one-step-ahead forecasting is not necessarily reliable when it is iterated through long missing intervals (Ben Taieb et al., 2012).

Table 1. Main differences between time series prediction and interpolation.

Aspect	**Prediction**	**Interpolation**
Objective	Estimate future values	Estimate missing internal values
Temporal position	After the last observed value	Inside an observed sequence
Available information	Past values only	Past and/or future surrounding values
Main challenge	Extrapolation	Reconstruction
Typical application	Forecasting future weather	Filling missing meteorological records
Evaluation	Forecasting error	Reconstruction error
Common metrics	RMSE, MAE, MAPE, NRMSE	RMSE, MAE, NRMSE

2.4 Prediction-Based Interpolation

Prediction models may also be adapted for interpolation. In this case, a model is trained to represent the temporal dynamics of the observed series and is then used to estimate missing values. Two common strategies are direct prediction and iterative prediction. These strategies correspond closely to single-step-ahead and multi-step-ahead forecasting settings: in single-step-ahead prediction, each estimate is obtained from observed input windows, whereas in multi-step-ahead prediction a model is recursively reused and its previous outputs may become inputs for subsequent estimates (Ben Taieb et al., 2012).

2.4.1 Direct Prediction

In direct prediction-based interpolation, each missing value is estimated independently from an input window of observed data. A fixed number of previous observations is supplied to the model, and the missing value is predicted directly:

$$\hat{x}_t = f(x_{t-p}, x_{t-p+1}, \dots, x_{t-1}) \quad (4)$$

where p denotes the input window size.

Error propagation is reduced by this strategy because predicted values are not repeatedly reused as inputs. However, the strategy may be limited when the missing interval is long and the required observed inputs are unavailable. In interpolation problems, direct strategies are most appropriate when the gap can be estimated from sufficient observed context without recursively feeding reconstructed values back into the model.

2.4.2 Iterative Prediction

In iterative prediction-based interpolation, the first missing value is predicted and then used as part of the input for estimating the next missing value. The same process is repeated until the complete gap has been reconstructed.

$$\hat{x}_{t+1} = f(x_{t-p+1}, \ldots, x_t) \tag{5}$$

$$\hat{x}_{t+2} = f(x_{t-p+2}, \ldots, \hat{x}_{t+1}) \tag{6}$$

Long missing intervals can be reconstructed with this strategy, but the procedure is more vulnerable to accumulated error because earlier predictions influence later predictions (Ben Taieb et al., 2012). In meteorological gaps, this effect is relevant because accumulated error may distort smooth transitions, extreme events, or daily and seasonal patterns.

3 Review Methodology

The searches were conducted iteratively during manuscript preparation, with the final update completed on 5 July 2026. The review was not designed as a PRISMA-style systematic review or meta-analysis, but as a structured methodological literature review focused on representative and verifiable contributions to meteorological time series interpolation. The consulted sources included ScienceDirect (Elsevier), IEEE Xplore, SpringerLink, ACM Digital Library, MDPI, NeurIPS Proceedings, AAAI Proceedings, Google Scholar, OpenAlex, and official publisher or DOI records. Searches combined terms related to meteorological time series interpolation, weather data imputation, missing data reconstruction, deep learning, neural networks, LSTM, transformers, evolutionary neural networks, EPNet, NEAT, and neuroevolution. Retrieved records were screened according to the criteria in Table 2, prioritizing studies with methodological relevance, verifiable bibliographic records, transferable insight for meteorological or environmental time series, and relevance to the taxonomy developed in this review. Foundational references and method-transfer papers were also retained through backward and forward citation checking when they supported definitions, baseline methods, or reconstruction strategies discussed in the manuscript.

Studies were selected based on their relevance to meteorological, climatological, hydrological, environmental, and air quality time series. The goal here was methodological synthesis, not a formal meta-analysis, so the coverage should be read as representative rather than exhaustive. Instead, reference works were selected according to their relevance, methodological contribution, data set characteristics, evaluation metrics, and relationship with missing data interpolation or imputation. Each selected reference was checked against at least one persistent scholarly record, such as Scopus, DataCite, an official publisher page, PMLR, NeurIPS, AAAI, Copernicus, arXiv, or another official proceeding repository. References were retained only when the title, authors, year, venue, and DOI or stable URL were consistent across the available records.

Table 2. Screening criteria used to select and organize the reviewed literature.

Criterion	Included	Excluded
Topical scope	Studies on interpolation, imputation, gap filling, missing data reconstruction, spatial interpolation, and time series modeling relevant to meteorological or environmental data.	Pure forecasting studies with no methodological relevance to internal gap reconstruction.
Methodological role	Foundational methods, representative applications, comparative studies, surveys, and recent deep learning or neuroevolutionary approaches that support the taxonomy.	Application reports that use a method without enough methodological detail or transferable insight.
Data setting	Meteorological, climatological, hydrological, air quality, environmental, geophysical, and general time series studies transferable to these settings.	Studies whose data assumptions are too distant from temporal or spatio-temporal reconstruction.
Validation record	Articles with DOI, official proceedings page, publisher page, or stable scholarly repository record.	Records whose title, authors, year, venue, DOI, or stable URL could not be verified.

An important methodological distinction relevant throughout this review concerns the mechanism underlying missingness. Following Rubin (1976), missing observations may be classified as missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). Meteorological gaps, however, rarely conform to the MCAR assumption implicit in the randomly masked benchmarks commonly used to evaluate machine learning and deep learning interpolators. Sensor failures, station outages, and quality-control removals tend to produce block-structured or systematic missingness that is often related to the underlying atmospheric conditions themselves — for example, extreme events that damage equipment or trigger automated flagging of anomalous readings. This structured missingness differs substantially from random masking (Schafer and Graham, 2002; Mitra et al., 2023), and is considered explicitly throughout the method families discussed in Sections 4–7.

4 Classical Interpolation Methods

Classical interpolation methods are widely used because they are simple, interpretable, and computationally inexpensive. They are also used as baselines against which more advanced approaches can be evaluated (Lepot et al., 2017; Moritz and Bartz-Beielstein, 2017). In meteorological applications, however, an apparently simple gap-filling problem may be complicated by variable behavior. For example, smooth curves may be suitable for temperature, whereas precipitation, wind speed, pollutant concentration, or pollutant peaks may show abrupt changes that are poorly represented by smooth interpolation curves (Junninen et al., 2004; Hofstra et al., 2008).

A further issue is the distinction between temporal interpolation at a single station and spatial or spatio-temporal reconstruction across a station network. Temporal interpolation exploits dependence along time, whereas spatial interpolation exploits dependence among locations. Meteorological data often require both types of information because missing values may coincide with local sensor failures, station downtime, or communication interruptions that affect a subset of the monitoring network (Haylock et al., 2008; Kondrashov and Ghil, 2006).

4.1 Linear Interpolation

Linear interpolation estimates a missing value by assuming a straight-line relation between two known neighboring observations. Given two known points (t_1, x_1) and (t_2, x_2), the value interpolated at time t is:

$$\hat{x}(t) = x_1 + (t - t_1)/(t_2 - t_1)(x_2 - x_1). \tag{7}$$

The method is efficient, easy to implement, and generally performs adequately — except when the meteorological variable behaves nonlinearly or shifts abruptly (Lepot et al., 2017). No extra parameters are introduced, which keeps it transparent and operationally simple. This method relies on an implicit assumption: change between two observed points is treated as monotonic and smooth. That assumption weakens fast for long gaps, or for variables with threshold behavior, strong diurnal cycles, or intermittent events like precipitation.

4.2 Nearest-Neighbor Interpolation

Nearest-neighbor interpolation assigns the value of the closest observed point to the missing one. Although simple to compute, discontinuities can appear, and the method struggles with variables that need smooth temporal transitions (Lepot et al., 2017). It tends to work better when the variable changes slowly relative to the sampling frequency, or when keeping the original observed value matters more than producing a smoothed estimate. Where derivatives, trends, or continuity matter downstream, it's a weaker choice.

4.3 Polynomial Interpolation

Polynomial interpolation fits a polynomial function to estimate missing values, which lets it capture nonlinear behavior. However, high-degree polynomials tend to oscillate and overfit, especially with noisy meteorological data (Lepot et al., 2017). In practice, it works best as a local, low-degree approximation rather than a global reconstruction strategy. This oscillation risk increases whenthe data carries measurement noise, outliers, or irregular sampling intervals.

4.4 Spline Interpolation

Spline interpolation stitches together piecewise polynomial functions to produce smooth curves through known points. Cubic splines are the usual choice, since they give smooth transitions while avoiding some of the pitfalls of high-degree polynomials (Lepot et al., 2017; Moritz and Bartz-Beielstein, 2017). They suit meteorological variables with continuous dynamics well, but abrupt events tend to get oversmoothed, and values near discontinuities can end up unrealistic. For this reason, spline performance should not be judged only by global error metrics; the extent to which extremes and local variability are preserved is equally important. Their weak point is that they need assumptions about signal structure or state dynamics going in.

4.5 State-Space, EM, and Spectral Reconstruction

State-space models and expectation-maximization (EM) offer a statistical route into the missing-data problem. Missing values and model parameters can be estimated at the same time, which helps when missingness is not confined to isolated points (Shumway and Stoffer, 1982; Kalman, 1960). Climate and geophysical work have also leaned on covariance-based and spectral techniques — singular spectrum analysis and its multichannel extensions, paired with spatial correlations — to reconstruct incomplete records (Schneider, 2001; Kondrashov and Ghil, 2006; Hocke and Kämpfer, 2009; Musial et al., 2011).

These approaches ask more of the analyst than simple interpolation, but they pay off when oscillatory modes, seasonal structure, or correlated missing segments are involved. Their main limitation is that they require prior assumptions about signal structure or state dynamics. It's worth pairing their use with an explicit look at dominant components and, where available, uncertainty estimates.

4.6 Kriging and Spatio-Temporal Interpolation

Kriging estimates missing values from spatial or spatio-temporal correlation and is well suited to settings in which meteorological stations are spread across a region. It does, however, require assumptions about spatial dependence and can become computationally expensive at scale (Matheron, 1963; Cressie, 1990; Li and Heap, 2014). Comparative studies of daily climate interpolation highlight an important consideration: how well spatial methods perform depends on topography, station density, variable type, and whether auxiliary variables like elevation are folded in (Hofstra et al., 2008; Haylock et al., 2008).

Kriging and related geostatistical techniques are most effective when a missing value at one station can reasonably be inferred from its neighbors. Their performance declines when missingness spreads across the whole network, or when local phenomena do not correlate well with nearby stations. Recent environmental interpolation work also pushes back against treating reconstructed series as if they were fully observed — uncertainty should travel alongside the reconstructed values, not disappear (Hristopulos et al., 2008; Lepot et al., 2017).

5 Machine Learning Methods for Meteorological Data Interpolation

Machine learning methods learn relationships directly from data, without forcing a parametric form on them beforehand. That makes them well suited to meteorological records where nonlinear behavior, variable interactions, or noise get in the way (Smola and Schölkopf, 2004; Ahmed et al., 2010; Stekhoven and Bühlmann, 2012). Unlike purely local interpolation, these methods can take lagged values, neighboring stations, calendar variables, elevation, and other predictors as input features. That flexibility is useful, but it opens the door to problems around feature selection, data leakage, and overfitting.

A common evaluation setup removes observed values on purpose and compares the reconstructed estimates against what was hidden. The missing-data pattern chosen for this test matters more than it might seem - random gaps, block gaps, station-level outages, and seasonal gaps can each lead to quite different conclusions about how well a model actually performs (Rubin, 1976; Schäfer and Graham, 2002; White et al., 2011; Ben Taieb et al., 2012). Machine learning interpolators, in other words, should be tested under missingness scenarios that reflect the real monitoring problem rather than a convenient artificial one.

5.1 Support Vector Regression

Support Vector Regression (SVR) captures nonlinear relationships through kernel functions, and has found use in both prediction and missing-data estimation, particularly on datasets that aren't too large (Smola and Schölkopf, 2004). For time series interpolation, SVR takes lagged observations and auxiliary variables as predictors. Its performance depends on kernel choice, regularization, and the construction of the input window. It is often easier to train than a neural network on moderate-size datasets; however, temporal structure must be manually encoded into the features rather than learned automatically.

5.2 Random Forest and Ensemble Methods

Tree-based ensembles — Random Forest, Gradient Boosting — handle nonlinear relationships and variable interactions well, and hold up reasonably against noise while incorporating multiple meteorological predictors (Breiman, 2001). Random forest imputation is particularly notable because it captures nonlinear relationships without committing to any single parametric form (Stekhoven and Bühlmann, 2012). These methods also produce variable-importance estimates, useful for determining whether a reconstruction relies primarily on temporal lags, neighboring stations, or exogenous variables.

Other imputation strategies, including nearest-neighbor methods, low-rank reconstruction, and chained equations, originated in fields such as bioinformatics and survey statistics, but still offer useful baselines here, since they formalize the relationship between observed and missing entries (Troyanskaya et al., 2001; White et al., 2011). Their main limitation is that temporal order can be disregarded in noisy series unless lags and time-dependent covariates are explicitly incorporated.

5.3 Gaussian Process Regression

Gaussian Process Regression gives probabilistic predictions along with uncertainty estimates, which is a real advantage for meteorological interpolation — though the computational cost climbs quickly with dataset size (Rasmussen and Williams, 2006). Conceptually, this approach is closely related to kriging, but it is typically framed within a broader machine learning setting, where kernels can encode temporal smoothness, periodicity, spatial distance, or some mix of all three. The probabilistic output is especially valuable when the reconstructed values will feed into later decision-making or uncertainty-sensitive modeling.

Scalability is the main constraint. Large station networks or high-frequency records often call for sparse approximations or local models. Still, the core idea of explicitly modeling uncertainty, which deterministic neural interpolators tend to omit entirely, keeps Gaussian processes relevant despite the computational cost.

6 Deep Learning Methods

Deep learning has drawn considerable attention for time series interpolation, largely because it can capture nonlinear temporal dependencies and complex patterns in large datasets (Goodfellow et al., 2016; Che et al., 2018; Cao et al., 2018; Du et al., 2023). Its main advantage is representation learning: instead of relying only on manually chosen lagged variables, these models learn internal features straight from raw or lightly processed temporal windows. This capability comes at a cost, however, as high-capacity models can readily overfit artificial missing patterns and may perform poorly when applied to different stations, seasons, or climate regimes.

Therefore, deep learning is not a suitable tool for all meteorological interpolation problems. Recent surveys and benchmarks make clear that performance depends heavily on data type, missingness pattern, missing rate, variable behavior, and how the evaluation itself is designed (Fang and Wang, 2020; Kazijevs and Samad, 2023; Wang et al., 2024). Classical or statistical methods can still be the better choice when the dataset is small, gaps are short, or the variable is smooth, interpretable, and operationally stable. Deep models are most justified when multivariate dependencies are strong, high-frequency records are available, or temporal interactions are too complex for simpler methods.

6.1 Multilayer Perceptron

The Multilayer Perceptron (MLP) is about as simple as neural architectures get. For interpolation, the input is typically a fixed-size temporal window, and the output is the estimated missing value:

$$\hat{x}_t = f(x_{t-p}, \ldots, x_{t-1}). \tag{8}$$

Despite its simplicity, an MLP models nonlinear relationships effectively and, when properly configured, can serve as a competitive baseline. It has a long track record in atmospheric science and environmental prediction, where it has been used to capture nonlinear relationships among meteorological variables (Gardner and Dorling, 1998; Zhang et al., 1998; Chen et al., 2002). For interpolation, an MLP can be trained on past observations, future observations, or both, and can take in extra variables — humidity, pressure, solar radiation, neighboring-station readings. Its main limitation is the absence of a built-in memory mechanism, so temporal structure must be derived entirely from the construction of the input window.

6.2 Evolutionary and Neuroevolutionary Neural Models

Evolutionary algorithms optimize neural networks by searching over weights, architectures, input variables, delays, activation functions, or hyperparameters. Early work pairing genetic algorithms with neural networks showed that evolutionary search could shape connectivity patterns and overall network structure (Whitley et al., 1990). From there, evolutionary artificial neural networks grew into an established research area, useful precisely when manual architecture selection proves difficult (Yao, 1999).

EPNet is a representative example, in which neural networks evolve through evolutionary programming operations - node addition, node deletion, connection addition, connection deletion, partial training (Yao and Liu, 1997). This is conceptually useful for meteorological interpolation, since the ideal number of lags, hidden units, and connections can shift from one station or variable to the next. NEAT takes a different neuroevolutionary route, evolving neural topologies through historical markings and incremental complexification (Stanley and Miikkulainen, 2002). Although it was not originally developed for interpolation, its core idea of growing topology from simple structures is directly relevant to the design of compact neural interpolators.

More recent neuroevolution research ties these ideas to modern neural architecture search: network components get designed, architecture search combines with learning, and models get optimized in settings where gradient-based search alone is not sufficient (Floreano et al., 2008; Stanley et al., 2019; Real et al., 2019; Miikkulainen et al., 2019). For meteorological interpolation specifically, evolutionary neural models look promising for hybrid systems that need to select lag structures, combine spatial and temporal predictors, or adapt architecture to different missingness patterns.

6.3 Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are built for sequential data — internal states carry information forward so temporal dependencies can be modeled. A well-known limitation is that conventional RNNs suffer from vanishing or exploding gradients over long sequences (Hochreiter and Schmidhuber, 1997). For interpolation this is still useful, since recurrent models can draw on the sequence both before and after a gap. How well that works in practice comes down to how missing values are represented and how uncertainty is handled.

6.4 Long Short-Term Memory Networks

LSTM networks fix some of the RNN's shortcomings through memory cells and gates that control the flow of information (Hochreiter and Schmidhuber, 1997). They are widely used in time series forecasting and have also been adapted for missing-data reconstruction. In meteorological interpolation specifically, LSTMs can pick up daily cycles, persistence, delayed effects, and nonlinear temporal dependencies. Other imputation strategies, including nearest-neighbor methods, low-rank reconstruction, and chained equations, originated in fields such as bioinformatics and survey statisticsLong consecutive gaps remain problematic, however, as recursive LSTM reconstructions accumulate error in ways that can degrade multi-step forecasting (Ben Taieb et al., 2012).

6.5 Gated Recurrent Units

GRUs simplify the LSTM design, usually with fewer parameters and still competitive results in time series tasks (Cho et al., 2014). GRU-D takes this further, adapting the GRU architecture for missing multivariate time series through masks and time intervals tied to the missingness pattern itself (Che et al., 2018). This represents a meaningful design choice, as the pattern of missingness can itself carry information. A run of systematic gaps might point to sensor maintenance, station malfunction, or environmental conditions interfering with the measurement, rather than pure randomness.

6.6 Bidirectional Recurrent Models

Bidirectional recurrent networks process a sequence in both forward and backward directions. This property is particularly relevant for interpolation because missing values are located inside the observed sequence, and the model is allowed to exploit both past and future context. BRITS is a representative example because missing values are treated as variables inside a bidirectional recurrent dynamical system (Cao et al., 2018). Bidirectional models better match the interpolation setting than pure forecasting models because they do not force the reconstruction to depend only on the past.

6.7 Attention-Based and Transformer Models

Transformer-based models use attention mechanisms to capture long-range dependencies in sequential data. Recent time series architectures have reported promising results in forecasting and imputation tasks. Informer adapts transformer ideas to long-sequence forecasting, whereas SAITS applies self-attention directly to time series imputation (Vaswani et al., 2017; Zhou et al., 2021; Du et al., 2023). Their ability to model global dependencies makes them relevant for meteorological interpolation, especially in complex data sets.

Recent transformer variants for time series introduce architectural changes that may also benefit interpolation. Autoformer combines decomposition and auto-correlation to represent long-term temporal structure (Wu et al., 2021). PatchTST divides time series into patches, thereby reducing series length and supporting longer input contexts (Nie et al., 2023). TimesNet transforms one-dimensional temporal variation into two-dimensional representations for multiple time series tasks, including imputation (Wu et al., 2023). iTransformer inverts the role of temporal and variable dimensions, which is useful for multivariate series where cross-variable dependencies are important (Liu et al., 2024). These models may provide useful concepts for reconstructing internal time series gaps, but their mechanisms suggest that careful validation is needed when long context and multivariate dependencies are present.

6.8 Generative Models

Generative models, including Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and diffusion models, may be used to reconstruct missing observations by learning the underlying data distribution. Such models are particularly useful when uncertainty and multiple plausible reconstructions are important. GAIN adapts adversarial learning to missing data imputation by training a generator and discriminator with partial missingness information (Yoon et al., 2018). GP-VAE combines variational autoencoders with a Gaussian process structure to obtain smooth and uncertainty-aware time series imputations (Fortuin et al., 2020). NAOMI addresses long-range sequence imputation through a non-autoregressive multiresolution strategy that reduces error propagation (Liu et al., 2019). CSDI is a recent example of a diffusion-based method for probabilistic time series imputation and interpolation (Tashiro et al., 2021).

7 A Methodological Taxonomy for Meteorological Time Series Interpolation

Most studies classify interpolation approaches by the algorithm that is used. However, in meteorological time series it is also useful to classify methods according to the reconstruction strategy. This distinction follows from differences between local gap filling, recursive multi-step estimation, bidirectional sequence reconstruction, probabilistic generation, and hybrid optimization (Lepot et al., 2017; Ben Taieb et al., 2012; Cao et al., 2018; Du et al., 2023; Tashiro et al., 2021). The purpose of this taxonomy is not to introduce a new algorithm, but to organize the methodological choices that influence how missing values are reconstructed.

The four reconstruction-strategy categories introduced in this taxonomy are not intended to replace the algorithm-based organization used in Sections 4–6 (classical, statistical, machine learning, and deep learning methods); rather, they provide a complementary, strategy-oriented lens that groups methods by how they use temporal context to reconstruct missing values, independent of algorithmic family. Table 3c maps the representative methods discussed in Sections 4–6 onto these four reconstruction strategies. Classical local methods such as linear, polynomial, and spline interpolation are not included in this mapping because, although they use both preceding and following observations, the taxonomy was developed primarily to organize machine learning and deep learning reconstruction strategies; their bidirectional use of context is noted separately in Section 4.

Table 3 Reconstruction Method Classification by Algorithmic Family

Criterion Method / family	Algorithmic family (Section)	Excluded Reconstruction strategy (Section 7)
Direct prediction-based interpolation	Prediction-based (2.4.1)	Not recursive — single-step local estimation

Criterion Method / family	Algorithmic family (Section)	Excluded Reconstruction strategy (Section 7)
Iterative prediction-based interpolation	Prediction-based (2.4.1)	7.1 Recursive/Iterative
MLP (recursively applied)	Deep learning (6.1)	7.1 Recursive/Iterative
EPNet, NEAT, neuroevolutionary search	Evolutionary/neuroevolutionary (6.2)	7.4 Hybrid/Evolutionary
RNN (unidirectional)	Deep learning (6.3)	7.1 Recursive/Iterative
LSTM, GRU, GRU-D	Deep learning (6.4–6.5)	7.1 Recursive/Iterative (unless used bidirectionally)
BRITS	Deep learning (6.6)	7.2 Bidirectional
Kalman smoother / state-space (bidirectional)	Classical/statistical (4.5)	7.2 Bidirectional
Gaussian Process Regression	Machine learning (5.3)	7.3 Probabilistic/Generative
GAIN, GP-VAE, NAOMI, CSDI	Generative (6.8)	7.3 Probabilistic/Generative
Informer, SAITS, Autoformer, PatchTST, TimesNet, iTransformer	Attention/transformer (6.7)	7.2 Bidirectional (non-causal attention) or 7.3 depending on probabilistic output
Hybrid classical+ML, spatial+neural combinations	Hybrid (7.4)	7.4 Hybrid/Evolutionary

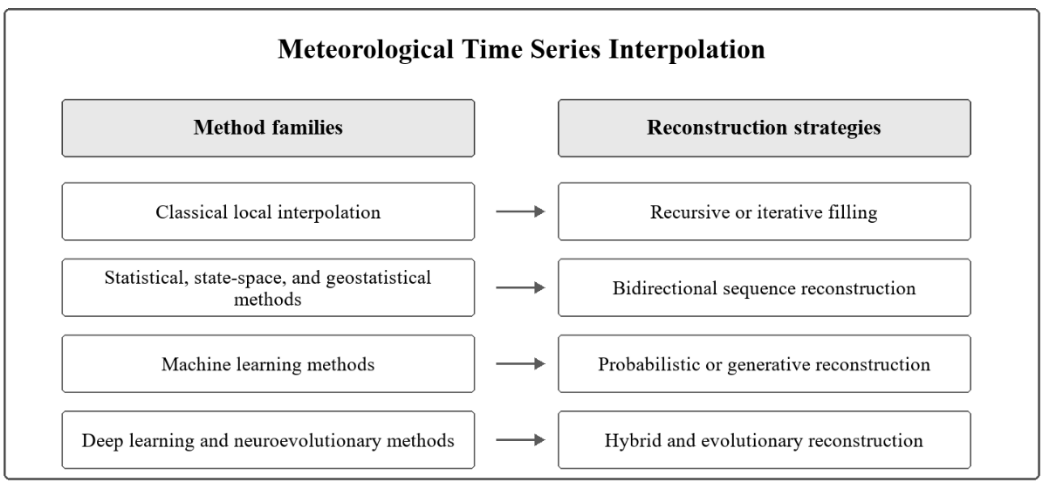

Fig. 1. *Methodological taxonomy found in the literature for meteorological time series interpolation.*

7.1 Recursive or Iterative Interpolation

Recursive or iterative interpolation methods reconstruct the missing interval one step at a time. This strategy is flexible because long consecutive gaps can be filled even when only part of the surrounding context is available. Nevertheless, error may accumulate, as occurs in recursive multi-step forecasting (Ben Taieb et al., 2012; Bontempi et al., 2013). In meteorological applications, accumulated error is particularly problematic when a reconstructed value is later used to infer another value inside the same gap. For this reason, recursive methods should be evaluated under block-missing scenarios rather than only under random missing values.

7.2 Bidirectional Interpolation

Bidirectional interpolation uses observations from both the past and the future. This strategy is highly relevant when the missing interval is surrounded by known values (Cao et al., 2018). It is methodologically closer to interpolation than to forecasting because information from both sides of the gap is used. Bidirectional recurrent models, smoothing models, and sequence-to-sequence architectures can therefore be interpreted as attempts to align the reconstruction with the temporal context before and after the missing interval.

7.3 Probabilistic and Generative Interpolation

Probabilistic and generative interpolation methods learn the distribution of the data and generate plausible missing values. Uncertainty-aware reconstructions may also be produced by these methods (Yoon et al., 2018; Fortuin et al., 2020; Tashiro et al., 2021). This is useful in meteorological data because several reconstructions may be plausible for long gaps or rapidly changing variables. Intervals or sampled trajectories can therefore be made visible to subsequent forecasting, climate monitoring, or decision-support models.

7.4 Hybrid and Evolutionary Interpolation

Hybrid approaches combine classical methods, statistical models, machine learning, deep learning, or optimization techniques in order to improve reconstruction accuracy. They are attractive when meteorological signals contain both smooth seasonal components and nonlinear short-term disturbances. Examples include spatial interpolation combined with neural networks, preliminary classical filling before deep learning, or evolutionary search over neural architectures and input windows (Yao and Liu, 1997; Stanley and Miikkulainen, 2002; Stanley et al., 2019).

Evolutionary neural approaches are relevant within this taxonomy because the structure of the interpolator can be optimized instead of being assumed in advance. Searches over lag selection, hidden nodes, connection density, or temporal and spatial predictors may be useful when the optimal representation differs across variables or stations.

8 Comparative Analysis of Existing Approaches

Representative studies supporting the methodological families discussed in this review are summarized in Table 3. The table is not intended to be exhaustive; rather, reference points are identified for organizing the comparison of meteorological time series interpolation methods.

Table 4. Representative studies for time series interpolation, imputation, forecasting strategy, and environmental reconstruction.

Year	Method	Data Type	Variable	Missing Pattern	Metric
1997	EPNet neuroevolution (Yao and Liu, 1997)	Neural network architecture search	Benchmark prediction tasks	Architecture and weight optimization	Prediction error
1998	MLP in atmospheric sciences (Gardner & Dorling, 1998)	Atmospheric data	Atmospheric variables	Prediction and modeling tasks	Error metrics
2001	Random Forest (Breiman, 2001)	Tabular and mixed data	Multiple variables	Nonlinear supervised learning	Prediction error
2002	NEAT (Stanley and Miikkulainen, 2002)	Neural topology evolution	Control benchmarks	Evolved topology and weights	Task fitness
2004	Statistical and neural imputation (Junninen et al., 2004)	Air quality time series	Pollutants	Missing monitoring records	Prediction error
2006	SSA gap filling (Kondrashov and Ghil, 2006)	Geophysical time series	Ocean, hydrology, atmosphere	Temporal and spatial gaps	Cross-validation
2008	Climate interpolation comparison (Hofstra et al., 2008)	Daily European climate data	Temperature, precipitation	Station interpolation	MAE, RMSE
2012	Multi-step forecasting strategies (Ben Taieb et al., 2012)	Multiple time series	NN5 competition series	Forecasting horizons	SMAPE
2014	Spatial interpolation review (Li and Heap, 2014)	Environmental spatial data	Environmental variables	Unsampled locations	Prediction error
2017	Time series interpolation review (Lepot et al., 2017)	Univariate time series	Environmental and engineering variables	Internal gaps	RMSE, uncertainty
2017	imputeTS algorithms (Moritz and Bartz-Beielstein, 2017)	Univariate time series	Sensor and benchmark series	Artificial and realistic gaps	RMSE, MAPE
2018	GRU-D (Che et al., 2018)	Multivariate time series	Health, geoscience, biology	Informative missingness	Classification and imputation error
2018	BRITS (Cao et al., 2018)	Multivariate time series	Air quality and health data	Multiple correlated missing values	MAE, RMSE
2018	GAIN (Yoon et al., 2018)	Multivariate tabular/time-indexed data	Multiple variables	Arbitrary missing values	RMSE

Year	Method	Data Type	Variable	Missing Pattern	Metric
2020	GP-VAE (Fortuin et al., 2020)	Multivariate time series	Healthcare, image-derived sequences	Missing multivariate values	Imputation error
2021	CSDI (Tashiro et al., 2021)	Healthcare and environmental time series	Multiple variables	Random and structured gaps	CRPS, MAE, RMSE
2021	Informer (Zhou et al., 2021)	Long-sequence time series	Energy, weather and traffic variables	Forecasting horizons	MSE, MAE
2023	TimesNet (Wu et al., 2023)	General time series	Multiple domains	Forecasting, imputation, classification	MSE, MAE
2023	SAITS (Du et al., 2023)	Multivariate time series	Healthcare, air quality and other real data	Artificial MCAR masking	MAE, MSE
2023	PatchTST (Nie et al., 2023)	Multivariate time series	Weather, traffic, energy	Long contexts	MSE, MAE

Table 4b. Methodological comparison of representative interpolation methods.

Method	Domain validation	Typical gap length	Missingness mechanism (as evaluated)	Spatial info used	Uncertainty estimation	Principal advantage	Main limitation
EPNet (Yao & Liu, 1997)	Transferred (benchmark architecture search)	N/A	N/A	No	No	Automatic node/connection addition-deletion	Not designed for interpolation specifically
MLP in atmospheric sciences (Gardner & Dorling, 1998)	Validated (atmospheric variables)	Short, local window	Not mechanism-specific	No	No	Captures nonlinear relationships among meteorological variables	No built-in memory; temporal structure must be hand-encoded
Random Forest (Breiman, 2001)	Transferred (tabular/mixed data)	Flexible (nonparametric)	Not mechanism-specific	Only if engineered as feature	Via variable importance only	Handles nonlinearity and variable interactions without a parametric form	Temporal order ignored unless lags are explicitly built in
NEAT (Stanley and Miikkulainen, 2002)	Transferred (control benchmarks)	N/A	N/A	No	No	Evolves compact topologies from simple structures	Not built for interpolation specifically
Statistical and neural imputation (Junninen et al., 2004)	Validated (air quality)	Short–moderate monitoring gaps	Real monitoring gaps (sensor/communication)	Not primary focus	Not emphasized	Combines statistical and neural methods for air-quality gap filling	Not detailed in manuscript — confirm
SSA gap filling (Kondrashov and Ghil, 2006)	Validated (geophysical: ocean, hydrology, atmosphere)	Temporal and spatial gaps	Structured (station/sensor-related)	Yes	Limited	Captures oscillatory/seasonal structure via spectral decomposition	Requires assumptions about signal structure
Climate interpolation comparison (Hofstra et al., 2008)	Validated (climate: temperature, precipitation)	Station-level spatial gaps	Spatial station gaps	Yes (core focus)	Compared across methods	Empirical comparison across six interpolation methods	Performance depends on topography, station density, variable type
Multi-step forecasting strategies (Ben Taieb et al., 2012)	Transferred (NN5 forecasting competition)	Multi-step horizons (forecasting)	N/A	No	No	Systematic direct-vs-iterative comparison informs error-accumulation discussion	Focused on forecasting, not internal gap reconstruction

Method	Domain validation	Typical gap length	Missingness mechanism (as evaluated)	Spatial info used	Uncertainty estimation	Principal advantage	Main limitation
Spatial interpolation review (Li and Heap, 2014)	Validated (environmental spatial data)	Spatial (unsampled locations)	N/A	Yes (core focus)	Method-dependent	Broad review of spatial interpolation applicable to environmental variables	Spatial focus, not temporal within-series gaps
Time series interpolation review (Lepot et al., 2017)	Validated (environmental & engineering variables)	General/internal gaps	General framework	Limited	Discussed as a criterion	Introduces performance criteria and uncertainty-assessment framework	General/introductory scope, not deep-learning-focused
imputeTS algorithms (Moritz and Bartz-Beielstein, 2017)	Validated (univariate sensor/benchmark series)	Artificial and realistic gaps	Artificial + realistic evaluated	No (univariate)	No	Accessible toolkit with multiple classical baselines	Univariate only; no spatial or multivariate dependence
GRU-D (Che et al., 2018)	**Transferred** (health, geoscience, biology)	Informative missingness	Explicitly modeled via masks + time intervals	No	No (deterministic)	Encodes the missingness pattern itself as an informative signal	Not validated on meteorological data in the cited work
BRITS (Cao et al., 2018)	**Partially validated** (air quality among tested domains)	Multiple correlated missing values	Correlated multivariate missingness	Via multivariate correlation	No	Bidirectional recurrent system uses both past and future context	Deterministic; no explicit uncertainty quantification
GAIN (Yoon et al., 2018)	**Transferred** (general tabular/multivariate benchmarks)	Arbitrary missing values	Arbitrary/general	No	No (adversarial point estimates)	Adversarial training captures complex data distributions	Not validated on meteorological data; GAN training can be unstable
GP-VAE (Fortuin et al., 2020)	**Transferred** (healthcare, image-derived sequences)	Multivariate missing values	General multivariate	No	Yes (GP-based)	Combines VAE + Gaussian process for smooth, uncertainty-aware imputation	Not validated on meteorological data; GP component is costly
CSDI (Tashiro et al., 2021)	**Partially validated** (healthcare and environmental series)	Random and structured gaps	Both evaluated	Not primary	Yes (diffusion-based, CRPS)	Probabilistic reconstruction suited to long gaps / fast-changing variables	Computationally intensive (diffusion sampling)
Informer (Zhou et al., 2021)	**Partially validated** (energy, weather, traffic)	Long forecasting horizons	N/A (forecasting)	No	No	Efficient attention mechanism for long-sequence forecasting	Designed for forecasting; interpolation use requires adaptation
TimesNet (Wu et al., 2023)	Transferred (general multi-domain, multi-task)	General	General (multi-task incl. imputation)	No	No	Transforms 1D variation into 2D representation; versatile across tasks	Meteorological-specific validation not detailed in manuscript
SAITS (Du et al., 2023)	**Partially validated** (air quality among tested domains)	Artificial MCAR masking	MCAR only	No	No	Self-attention imputation designed specifically for time series	Evaluated under MCAR masking, which may not reflect real meteorological missingness

Method	Domain validation	Typical gap length	Missingness mechanism (as evaluated)	Spatial info used	Uncertainty estimation	Principal advantage	Main limitation
PatchTST (Nie et al., 2023)	**Partially validated** (weather among tested domains)	Long contexts	N/A (forecasting-oriented)	No	No	Patch-based tokenization supports long input contexts efficiently	Primarily forecasting-oriented; interpolation use requires adaptation

While Table 3 identifies representative studies by year, data type, variable, and evaluation metric, it does not by itself indicate the practical trade-offs relevant to method selection. Table 3b extends this comparison along dimensions directly relevant to meteorological gap filling: whether the method has been validated on meteorological/environmental data or is included primarily for its methodological transferability, typical gap length addressed, the missingness mechanism considered in the original evaluation, whether spatial information is used, whether uncertainty is estimated, and the principal advantage and main limitation reported in Sections 4–6.

9 Practical Method Selection Guidelines

Method selection should be guided by the structure of the missingness problem rather than by model novelty alone. Practical baselines and caution points for common meteorological interpolation scenarios are summarized in Table 4.

Table 5. Practical guidance for selecting interpolation methods in meteorological time series.

Scenario	Recommended starting point	Main cautions
Short gaps in smooth univariate variables	Linear interpolation, cubic splines, Kalman smoothing, or imputeTS baselines (Lepot et al., 2017; Moritz and Bartz-Beielstein, 2017; Kalman, 1960).	May oversmooth precipitation, wind gusts, pollutant peaks, or abrupt frontal changes.
Long consecutive gaps	State-space models, SSA/M-SSA, bidirectional recurrent models, generative models, or hybrid approaches (Kondrashov and Ghil, 2006; Cao et al., 2018; Tashiro et al., 2021).	Evaluate with block-missing scenarios; avoid relying only on random masking.
Multivariate station records	Random Forest, Gaussian Process Regression, BRITS, SAITS, or transformer variants (Breiman, 2001; Rasmussen and Williams, 2006; Cao et al., 2018; Du et al., 2023).	Prevent data leakage from future observations or neighboring stations during validation.
Spatial station networks	Kriging, spatio-temporal interpolation, Gaussian processes, or spatially informed machine learning (Matheron, 1963; Cressie, 1990; Li and Heap, 2014; Hofstra et al., 2008).	Performance depends on station density, topography, variable type, and spatial correlation.
Uncertainty-sensitive applications	Gaussian processes, state-space models, multiple imputation, diffusion models, or probabilistic generative models (Rasmussen and Williams, 2006; White et al., 2011; Tashiro et al., 2021).	Deterministic fills should not be treated as observed data without uncertainty reporting.
Small data or high interpretability requirements	Classical interpolation, state-space models, Random Forest, or Gaussian processes (Kalman, 1960; Breiman, 2001; Rasmussen and Williams, 2006).	High-capacity deep models may overfit and may be difficult to justify operationally in high-interpretability settings (Rudin, 2019).
Unknown lags, variables, or architecture	Hybrid/evolutionary search, EPNet, NEAT-inspired topology search, or feature-selection wrappers (Yao and Liu, 1997; Stanley and Miikkulainen, 2002; Stanley et al., 2019).	Computational cost and reproducibility must be controlled with transparent validation.

10 Open Challenges

Despite important advances, meteorological time series interpolation remains an open research problem. Relevant challenges include: a) long consecutive missing intervals, because recursive methods may accumulate error and artificial random masks may underestimate the difficulty of block gaps (Ben Taieb et al., 2012; Kazijevs and Samad, 2023); b) robustness to noise and anomalous observations, which is critical in environmental monitoring where missingness and outliers may occur together (Junninen et al., 2004; Musial et al., 2011); c) evaluation under realistic missing data mechanisms, including MCAR, MAR, MNAR, and structured missingness patterns (Rubin, 1976; Schäfer and Graham, 2002; Mitra et al., 2023); d) integration of spatial and temporal dependencies across station networks (Li and Heap, 2014; Hofstra et al., 2008; Kondrashov and Ghil, 2006); e) lack of standardized meteorological benchmark data sets that compare classical, statistical, machine learning, and deep learning methods under the same missingness scenarios (Fang and Wang, 2020; Wang et al., 2024); f) limited uncertainty quantification in deterministic neural interpolators (Rasmussen and Williams, 2006; Tashiro et al., 2021); g) interpretability of deep learning-based interpolators, especially when reconstructed values support operational decisions; and h) generalization across stations, regions, seasons, and climatic conditions.

11 Future Directions

Future research in meteorological time series interpolation may benefit from several emerging directions. Transformer-based models and long-sequence architectures may improve the modeling of long-range temporal dependencies (Zhou et al., 2021; Du et al., 2023; Wang et al., 2024). Graph neural networks can incorporate spatial relationships among meteorological stations. Physics-informed neural networks may integrate physical constraints into data-driven models so that reconstructions remain consistent with known atmospheric relationships. Generative models and diffusion-based approaches may provide uncertainty-aware reconstructions (Tashiro et al., 2021). Evolutionary optimization may also support architecture search, feature selection, and hybrid neural interpolators when model structure is not known a priori (Yao, 1999; Stanley and Miikkulainen, 2002). Finally, self-supervised learning and foundation models for time series may reduce the dependency on fully labeled data sets, although their usefulness will depend on realistic missingness patterns rather than only on random artificial masking.

12 Conclusions

Interpolation methods for meteorological time series were reviewed in this review, considering classical, machine learning, and deep learning approaches. A conceptual distinction between prediction and interpolation was presented to clarify the methodological differences between forecasting future observations and reconstructing missing internal values. Furthermore, a taxonomy based on reconstruction strategy was used to organize the literature from a methodological perspective.

This review was also situated within the context of existing survey studies. Although previous reviews have addressed interpolation for general time series, spatial interpolation, and deep learning-based imputation, the particular characteristics of meteorological data require an integrated perspective in which temporal dynamics, spatial monitoring networks, physical variability, missing-data mechanisms, and uncertainty are jointly considered. Accordingly, the taxonomy and practical guidelines presented in this review were developed to facilitate the selection of appropriate interpolation methods rather than to advocate a particular family of algorithms.

Meteorological interpolation continues to represent an active area of research owing to the complexity of atmospheric variables, the frequent occurrence of missing observations, and the importance of reliable data reconstruction for decision-making processes. Future developments are expected to be driven by hybrid models, bidirectional neural architectures, generative methods, uncertainty estimation techniques, evolutionary architecture search, and physically informed learning approaches. From a practical standpoint, it is recommended that newly proposed methods be assessed against robust classical and statistical baselines under representative missing-data scenarios, including extended gaps, station outages, structured patterns of missingness, and varying climatic conditions.

References

Ahmed, N. K., Atiya, A. F., El Gayar, N., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews, 29*(5–6), 594–621. https://doi.org/10.1080/07474938.2010.481556

Ben Taieb, S., Bontempi, G., Atiya, A. F., & Sorjamaa, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert Systems with Applications, 39*(8), 7067–7083. https://doi.org/10.1016/j.eswa.2012.01.039

Bontempi, G., Ben Taieb, S., & Le Borgne, Y.-A. (2013). Machine learning strategies for time series forecasting. In M.-A. Aufaure & E. Zimányi (Eds.), *Business intelligence: Second European Summer School, eBISS 2012, Brussels, Belgium, July 15–21, 2012, tutorial lectures* (pp. 62–77). Springer. https://doi.org/10.1007/978-3-642-36318-4_3

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324

Cao, W., Wang, D., Li, J., Zhou, H., Li, L., & Li, Y. (2018). BRITS: Bidirectional recurrent imputation for time series. *Advances in Neural Information Processing Systems, 31*.

Che, Z., Purushotham, S., Cho, K., Sontag, D., & Liu, Y. (2018). Recurrent neural networks for multivariate time series with missing values. *Scientific Reports, 8*, Article 6085. https://doi.org/10.1038/s41598-018-24271-9

Chen, Y., Kopp, G. A., & Surry, D. (2002). Interpolation of wind-induced pressure time series with an artificial neural network. *Journal of Wind Engineering and Industrial Aerodynamics, 90*(6), 589–615. https://doi.org/10.1016/S0167-6105(02)00155-1

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1724–1734). Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1179

Cressie, N. (1990). The origins of kriging. *Mathematical Geology, 22*(3), 239–252. https://doi.org/10.1007/BF00889887

De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting, 22*(3), 443–473. https://doi.org/10.1016/j.ijforecast.2006.01.001

Du, W., Côté, D., & Liu, Y. (2023). SAITS: Self-attention-based imputation for time series. *Expert Systems with Applications, 219*, Article 119619. https://doi.org/10.1016/j.eswa.2023.119619

Fang, C., & Wang, C. (2020). *Time series data imputation: A survey on deep learning approaches* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2011.11347

Floreano, D., Dürr, P., & Mattiussi, C. (2008). Neuroevolution: From architectures to learning. *Evolutionary Intelligence, 1*(1), 47–62. https://doi.org/10.1007/s12065-007-0002-4

Fortuin, V., Baranchuk, D., Raetsch, G., & Mandt, S. (2020). GP-VAE: Deep probabilistic time series imputation. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics* (Vol. 108, pp. 1651–1661). PMLR.

Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmospheric Environment, 32*(14–15), 2627–2636. https://doi.org/10.1016/S1352-2310(97)00447-0

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research: Atmospheres, 113*, Article D20119. https://doi.org/10.1029/2008JD010201

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hocke, K., & Kämpfer, N. (2009). Gap filling and noise reduction of unevenly sampled data by means of the Lomb-Scargle periodogram. *Atmospheric Chemistry and Physics, 9*, 4197–4206. https://doi.org/10.5194/acp-9-4197-2009

Hofstra, N., Haylock, M., New, M., Jones, P. D., & Frei, C. (2008). Comparison of six methods for the interpolation of daily, European climate data. *Journal of Geophysical Research: Atmospheres, 113*, Article D21110. https://doi.org/10.1029/2008JD010100

Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., & Kolehmainen, M. (2004). Methods for imputation of missing values in air quality data sets. *Atmospheric Environment, 38*(18), 2895–2907. https://doi.org/10.1016/j.atmosenv.2004.02.026

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, 82*(1), 35–45. https://doi.org/10.1115/1.3662552

Kazijevs, M., & Samad, M. D. (2023). Deep imputation of missing values in time series health data: A review with benchmarking. *Journal of Biomedical Informatics, 144*, Article 104440. https://doi.org/10.1016/j.jbi.2023.104440

Kondrashov, D., & Ghil, M. (2006). Spatio-temporal filling of missing points in geophysical data sets. *Nonlinear Processes in Geophysics, 13*, 151–159. https://doi.org/10.5194/npg-13-151-2006

Lepot, M., Aubin, J.-B., & Clemens, F. H. L. R. (2017). Interpolation in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment. *Water, 9*(10), Article 796. https://doi.org/10.3390/w9100796

Li, J., & Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software, 53*, 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008

Liu, Y., Yu, R., Zheng, S., Zhan, E., & Yue, Y. (2019). *NAOMI: Non-autoregressive multiresolution sequence imputation* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1901.10946

Liu, Y., Hu, T., Zhang, H., Wu, H., Wang, S., Ma, L., & Long, M. (2024). iTransformer: Inverted transformers are effective for time series forecasting. In *The Twelfth International Conference on Learning Representations*.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology, 58*(8), 1246–1266. https://doi.org/10.2113/gsecongeo.58.8.1246

Miikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., Raju, B., Shahrzad, H., Navruzyan, A., Duffy, N., & Hodjat, B. (2019). Evolving deep neural networks. In R. Kozma, C. Alippi, Y. Choe, & F. C. Morabito (Eds.), *Artificial intelligence in the age of neural networks and brain computing* (pp. 293–312). Academic Press. https://doi.org/10.1016/B978-0-12-815480-9.00015-3

Mitra, R., McGough, S. F., Chakraborti, T., Holmes, C., Copping, R., Hagenbuch, N., Biedermann, S., Noonan, J., Lehmann, B., Shenvi, A., Doan, X. V., Leslie, D., Bianconi, G., Sanchez-Garcia, R., Davies, A., Mackintosh, M., Andrinopoulou, E.-R., Basiri, A., Harbron, C., & MacArthur, B. D. (2023). Learning from data with structured missingness. *Nature Machine Intelligence, 5*, 13–23. https://doi.org/10.1038/s42256-022-00596-z

Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time series missing value imputation in R. *The R Journal, 9*(1), 207–218. https://doi.org/10.32614/RJ-2017-009

Musial, J. P., Verstraete, M. M., & Gobron, N. (2011). Technical note: Comparing the effectiveness of recent algorithms to fill and smooth incomplete and noisy time series. *Atmospheric Chemistry and Physics, 11*, 7905–7923. https://doi.org/10.5194/acp-11-7905-2011

Nie, Y., Nguyen, N. H., Sinthong, P., & Kalagnanam, J. (2023). A time series is worth 64 words: Long-term forecasting with transformers. In *The Eleventh International Conference on Learning Representations*.

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. MIT Press.

Real, E., Aggarwal, A., Huang, Y., & Le, Q. V. (2019). Regularized evolution for image classifier architecture search. *Proceedings of the AAAI Conference on Artificial Intelligence, 33*(1), 4780–4789. https://doi.org/10.1609/aaai.v33i01.33014780

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581–592. https://doi.org/10.1093/biomet/63.3.581

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*, 206–215. https://doi.org/10.1038/s42256-019-0048-x

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Shumway, R. H., & Stoffer, D. S. (1982). An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis, 3*(4), 253–264. https://doi.org/10.1111/j.1467-9892.1982.tb00349.x

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing, 14*(3), 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Stanley, K. O., & Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evolutionary Computation, 10*(2), 99–127. https://doi.org/10.1162/106365602320169811

Stanley, K. O., Clune, J., Lehman, J., & Miikkulainen, R. (2019). Designing neural networks through neuroevolution. *Nature Machine Intelligence, 1*, 24–35. https://doi.org/10.1038/s42256-018-0006-z

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics, 28*(1), 112–118. https://doi.org/10.1093/bioinformatics/btr597

Tashiro, Y., Song, J., Song, Y., & Ermon, S. (2021). CSDI: Conditional score-based diffusion models for probabilistic time series imputation. *Advances in Neural Information Processing Systems, 34*, 24804–24816.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., & Altman, R. B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics, 17*(6), 520–525. https://doi.org/10.1093/bioinformatics/17.6.520

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 5998–6008.

Wang, J., Du, W., Yang, Y., Qian, L., Cao, W., Zhang, K., Wang, W., Liang, Y., & Wen, Q. (2025). Deep learning for multivariate time series imputation: A survey. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence* (pp. 10696–10704). https://doi.org/10.24963/ijcai.2025/1187

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30*(4), 377–399. https://doi.org/10.1002/sim.4067

Whitley, D., Starkweather, T., & Bogart, C. (1990). Genetic algorithms and neural networks: Optimizing connections and connectivity. *Parallel Computing, 14*(3), 347–361. https://doi.org/10.1016/0167-8191(90)90086-O

Wu, H., Xu, J., Wang, J., & Long, M. (2021). *Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2106.13008

Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J., & Long, M. (2023). TimesNet: Temporal 2D-variation modeling for general time series analysis. In *The Eleventh International Conference on Learning Representations*.

Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE, 87*(9), 1423–1447. https://doi.org/10.1109/5.784219

Yao, X., & Liu, Y. (1997). A new evolutionary system for evolving artificial neural networks. *IEEE Transactions on Neural Networks, 8*(3), 694–713. https://doi.org/10.1109/72.572107

Yoon, J., Jordon, J., & van der Schaar, M. (2018). GAIN: Missing data imputation using generative adversarial nets. In *Proceedings of the 35th International Conference on Machine Learning* (Vol. 80, pp. 5689–5698). PMLR.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting, 14*(1), 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(12), 11106–11115. https://doi.org/10.1609/aaai.v35i12.17325

Žukovič, M., & Hristopulos, D. T. (2008). Environmental time series interpolation based on Spartan random processes. *Atmospheric Environment, 42*(33), 7669–7678. https://doi.org/10.1016/j.atmosenv.2008.05.062