Discrete and Continuous Dynamical Systems - Series S
Vol. 15, No. 12, December 2022, pp. 3667–3688
doi:10.3934/dcdss.2022039

SPLINE BASED HERMITE QUASI-INTERPOLATION FOR UNIVARIATE TIME SERIES

Antonella Falini and Francesca Mazzia*

Dipartimento di Informatica, Università degli Studi di Bari Aldo Moro, via Orabona, 4
Bari 70125, Italy

Cristiano Tamborrino

Dipartimento di Matematica, Università degli Studi di Bari Aldo Moro, via Orabona, 4
Bari 70125, Italy

Abstract. In this article the authors introduce a spline Hermite quasi-interpolation technique for the preprocessing operations of imputation and smoothing of univariate time series. The constructed model is then applied for the forecast and for the anomaly detection. In particular, for the latter case, algorithms based on the combination of quasi-interpolation, dynamic copulas and clustering have been proposed. Some numerical results are included showing the effectiveness of the presented techniques.

1. **Dedication.** This research paper is dedicated to the memory of Rosa Maria Mininni, Rosella for her friends. She started her research activity at the University of Bari sharing the office with Francesca Mazzia and, even if they worked on different subjects, they collaborate for the last 10 years on industrial projects. One of these projects is the PON Project “Change Detection in Remote Sensing” CUP H94F18000260006, that gives the financial support to the PhD position of Cristiano Tamborrino. Rosella was also the coordinator of the project “MAIA: Monitoraggio attivo dell’infrastruttura”, whose activity initiated, unfortunately, without her support. This research is based on some work we commenced to collaborate to, trying to merge our research interests. In particular, we recall the works related to saliency and change detection, made in collaboration with Antonella Falini [21, 23] and the work related to the inpainting techniques [2], all of them resulting from common projects with local industrial partners. The last section of the paper is devoted to the description of an algorithm based on dynamic copulas, a subject that she was following with extremely interest. For us it is like she is a coauthor of this paper. Thank you Rosella!

2020 *Mathematics Subject Classification.* Primary: 65D07, 65D10, 62M10; Secondary: 62M20.

Key words and phrases. Time series analysis, spline Hermite quasi-interpolation, imputation, smoothing, forecasting, anomaly detection.

The research of Antonella Falini is founded by PON Project AIM 1852414 CUP H95G18000120006 ATT1. The research of Francesca Mazzia is founded by the PON “Ricerca e Innovazione 2014-2020”, project “MAIA: Monitoraggio attivo dell’infrastruttura”, n. ARS01 00353. The research of Cristiano Tamborrino is funded by PON Project “Change Detection in Remote Sensing” CUP H94F18000260006.

* Corresponding author: Francesca Mazzia.

2. **Introduction.** A *time series* is a sequence of evenly spaced and ordered data collected at different time instants. Since they may be collected for several types of applications, there is a great variety of approaches for studying and analyzing time series [1, 25, 30]. Usually, the initial raw data have to be prepared for an effective exploratory analysis, hence, preprocessing techniques, such as normalization [26, 59], cleaning [38], smoothing and imputation [4, 48] are a necessary task. Time series models are useful in understanding the underlying structure that produces the observed data, to monitor and to have either feedback or feedforward control. They are used for many applications such as, economic forecasting [10], process and quality control [52], census analysis [11], network traffic anomaly detection for railway transportation critical infrastructure [3] and many more. In particular, in the present work we adopt a spline based Hermite quasi-interpolation operator to preprocess and to perform forecasting and anomaly detection on both, synthetic and real datasets.

In general, quasi-interpolation denotes an approach to construct efficient local approximants to a given set of data or to a given function, see [41] for a general introduction. The choice of quasi-interpolation is motivated by two main reasons. In fact, data collected from realistic scenarios are usually affected by errors, therefore, interpolation methods might be too stringent, beside suffering from the overfitting phenomenon. Moreover, since quasi-interpolation relies on a local construction, the computational cost is greatly reduced compared to global approaches such as interpolation.

When it comes to forecasting, the new unknown values should be predicted via historical data, hence, constructing a robust but at the same time versatile continuous model is of fundamental importance. Many time series models have been developed exactly for this task, including ARIMA [13], neural networks [29] and Fuzzy-neural autoregressive models [51], Garch model [28] and support vector machine based approaches [49]. Other approaches build the continuous model and hence construct a fitting curve by using smoothing algorithms, see e.g., [17, 36]. Following the same philosophy, we adopt the Hermite quasi-interpolation operator introduced in [42] to produce a smooth model which also gives a rather accurate prediction in the short-run. In particular, we present two valid techniques: one produces a C^{d-1} smooth spline of degree d; the second one gives a C^d smooth spline of degree $d+1$, with d degree of the chosen quasi-interpolant basis. Regarding anomaly detection, the task is to identify behaviors of the data that greatly differ from the standard trend, see [9] for a review. Anomalies are referred to particular interesting events or suspicious data records, like in case of floods, fires, or earthquakes. Many anomaly detection algorithms for time series are developed from outlier detection strategies, see for example [63] and references therein. Recently, neural networks, and hence, supervised learning, have been employed to improve the accuracy of the obtained results, [67]. In this paper we employ the quasi-interpolation operator to preprocess the data and hence, to produce a smooth model. Later on, an unsupervised learning approach, based on dynamic copulas is described. For modelling multiple dependencies, the use of copulas is a very powerful approach [18]. One of the main advantages, when using copulas in time series, is that a semiparametric approach can be chosen: while estimating the marginal distributions using non-parametric methods, the copula itself can then be estimated parametrically using the maximum likelihood estimation [34]. The nature of copulas can either be static or time-variant. *Time-varying* copulas, or equivalently, *dynamic* copulas, might be considered as the dynamic generalizations of a Pearson correlation or Kendall's tau and in practice, time-varying copulas are often assumed to follow the autoregressive moving average process (ARMA) [53].

The paper is organized as follows: on Section 3 we recall the main features of the adopted Hermite quasi-interpolant operator; on Section 4 we formalize the model notation and we explain the application of the proposed technique to the operations of imputation and smoothing for time series. On Section 5 follows an application to forecasting and in Section 6 to anomaly detection. Finally, on Section 7 we give some conclusions.

3. **B-spline Hermite quasi-interpolation.** Quasi-interpolation is a technique that allows to construct a local approximant by keeping low the computational cost and the needed degrees of freedom. Generally, the common way to express a univariate spline quasi-interpolant (QI) d-degree approximation reads as,

$$Q_d\, f(\cdot) = \sum_{j=-d}^{N-1} \lambda_j(f) B_{j,d}(\cdot), \tag{1}$$

where $B_{j,d}$ are d-degree B-splines assumed to be defined on an extended knot vector π:

$$\pi := \{\tau_{-d}, \dots, \tau_{N+d}\}, \quad \tau_j \leq \tau_{j+1},$$

and spanning the space,

$$\mathbb{S}_d^\pi := \langle B_{-d,d}, \dots, B_{N-1,d} \rangle.$$

The spline space $\mathbb{S}_d^\pi$ is constructed such that $\tau_0 := a$ and $\tau_N := b$ with $[a,b]$ interval where the function f is defined. The local linear functionals λ_j in (1) can be computed by using several methodologies, such as differential [16, 15], integral methods [58, 56] , and discrete approaches [41, 57]. In the present work we adopt the Hermite QI developed in [42], which is derived from an associated BS method, used in the context of Ordinary Differential Equations (ODEs). The BS methods are a specific class of linear multistep methods for which the construction of a Hermite spline interpolation scheme can be carried out locally by using the produced numerical solution and its numerical derivative at the mesh breakpoints. These sites are used as collocation points and define the knots of the constructed spline, see [44, 45] for further details. Therefore, the QI scheme needs only the knowledge of f and f' at the knots. It is proved that it is a projector in the space of C^{d-1} splines of degree d and that it has optimal approximation order $d+1$ on quasi-uniform meshes for $f \in C^{d+1}([a,b])$.

When dealing with numerical methods for ODEs, the unknown function values and its derivative values at some mesh points are provided as input data. In the current framework, however, f' values are not available. Hence, we rely on the variant QI scheme, introduced in [43] to produce numerical quadrature formulas, where the derivative information are approximated by using a symmetric finite difference scheme. In particular, the derivative values at the knots are computed as

$$\mathbf{f}' \approx \Gamma^\ell\, \mathbf{f},$$

where $\mathbf{f}' := (f'(t_0)\dots, f'(t_N))^\top$, $\mathbf{f} := (f(t_0), \dots, f(t_N))^\top$, and Γ^ℓ is a $(\ell+1)$-banded matrix, of dimension $(N+1)\times(N+1)$. The integer ℓ is chosen to be equal to $d+1$ when d is odd, and it is equal to $d+2$ when d is even. For the effective computation of the λ_j functionals, in general settings, we refer to [8]. Although the variant QI is not a projector on $\mathbb{S}_d^\pi$, it keeps the same approximation order of the original QI. Moreover, it exhibits super-optimal convergence properties when degree even is employed either on uniform or not-uniform meshes, for the treatment of singular integrals, see [20, 22].

In all the experiments we chose $d = 2$, coincident auxiliary knots and coefficients λ_j in Eq.(1) provided in [42], Table 2.

4. **Model settings and preprocessing.** In the present work we deal with time series that can be defined as *stochastic processes*, i.e., a collection of random variables $\{x_t\}$, indexed over the time t. In particular, t is a discrete variable sampled at m time instants in a time domain $[T_1, T_2] \in \mathbb{R}$:

$$\{t_j, j = 1, \ldots, m\}.$$

We shall refer to the observed values of the stochastic process at every t_j as the set of *realizations*:

$$\{x_j, j = 1, \ldots, m\}.$$

The realizations are, in general, noisy observations of an underlying smooth function $f(t)$:

$$x_j = f(t_j) + \varepsilon_j \quad j = 1, \ldots, m,$$

where $\forall j, \varepsilon_j$ are independently distributed residuals with mean zero and constant variance.

We propose a uniform approach in estimating the function $f(t)$ and performing preprocessing operations such as *imputation* and *smoothing* by using the QI operator introduced in Section 3.

4.1. **Imputation.** The imputation process refers to estimate eventually missing values. Common techniques to deal with this issue include interpolation, moving average, decomposition and linear regression methods. For a practical overview and implementation in R-CRAN, we refer to [47].

We apply the QI operator to the data (t_j, x_j) with j index of known values. To approximate the first derivative, we use the mean value of backward and forward differences, when both are available. Otherwise, for extreme points, we use only either one of the two. Once the QI spline s is computed, the missing realization x_z is estimated as, $x_z := s(t_z)$, where t_z refers to the z^{th} time instant associated to the z^{th} missing realization. Note that the spline s is computed without taking into account the seasonality, i.e., periodicity of the data.

The algorithm has been tested on the "tsAirgap" time series dataset which consists of a complete series of 144 rows and an incomplete version with 13 missing values. The tsAirgap series represents the monthly totals of international airline passengers from 1949 to 1960[1], see [6]. The obtained results are displayed in Fig.1. We note that they are consistent with interpolation based strategies.

[1]The dataset can be downloaded at https://github.com/SteffenMoritz/imputeTS/tree/master/data

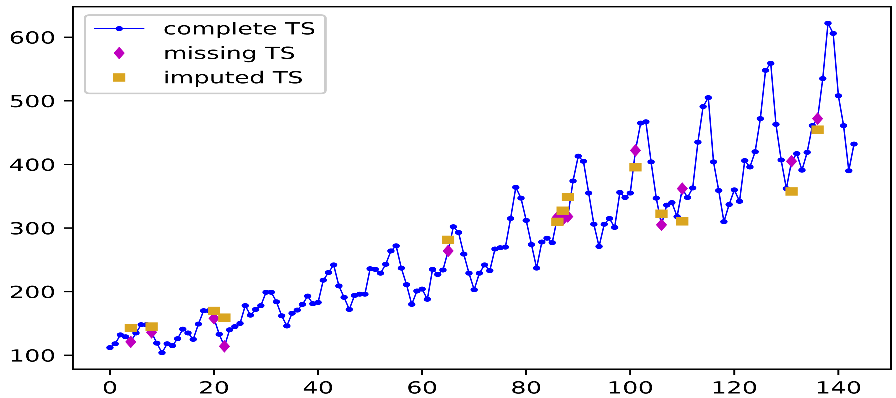


FIGURE 1. Results on "tsAirgap": the dots in blue are the known data, the magenta diamonds the missing values and the yellow squares are the imputed values.

4.2. **Smoothing.** One of the first steps that can be applied during preprocessing a time series is a *smoothing* technique. In general, smoothing out the irregular roughness can help to better identify patterns or trends.

A very simple approach to accomplish this task is to employ a *moving average smoother*,

$$y_j = \sum_{i=-k}^{k} a_i^{(j)} x_{j-i}, \tag{2}$$

where in general $a_i^{(j)} = \dfrac{1}{2k+1}$, $\sum_{i=-k}^{k} a_i^{(j)} = 1$, $\{x_t\}$ is the set of the realizations and the integer k is the window size.

When the data can be assumed to fluctuate around a steady mean value, i.e., there is no trend or consistent pattern of growth, then a *single exponential smoothing* can be a preferable choice,

$$\hat{x}_{j+1} = \alpha x_j + (1-\alpha)\hat{x}_j. \tag{3}$$

In Eq.(3) each new smoothed value $\hat{x}_{j+1}$ is computed as the weighted average of the current observation x_j and the previous smoothed observation $\hat{x}_j$, while α is the smoothing constant and its value is chosen according to which observation should be weighted more.

Although this technique is a smoothing method, it is principally used for short run forecasting as it is equivalent to the use of an *autoregressive integrated moving average* (ARIMA) model with no constant and with parameters:

- order $p = 0$,
- finite difference order equal to 1,
- moving average order $q = 1$.

When the data show a trend but no seasonality, then it is common to apply a *double exponential smoother*,

$$\begin{aligned}\hat{x}_{j+1} &= \alpha x_j + (1-\alpha)(\hat{x}_j - \hat{b}_j)\\ \hat{b}_{j+1} &= \beta(\hat{x}_{j+1} - \hat{x}_j) + (1-\beta)\hat{b}_j,\end{aligned} \tag{4}$$

where $0 < \alpha < 1$ and $0 < \beta < 1$. From Eq.(4) we notice that the new smoothed value $\hat{x}_{j+1}$ is computed via a single exponential smoother, while the smoothed *trend*. i.e., the variable b, is computed via a single exponential smoother on the first differences. This technique is equivalent to an ARIMA model as well, but with parameters $p = 0$, $q = 2$ and second order finite differences. Both models, (3)-(4) are strongly dependent on the choice of the smoothing factors α and β that can be obtained via non-linear optimization techniques, such as the Marquardt algorithm. Moreover, several choices can be made for the initialization values, i.e., $\hat{x}_1$ and $\hat{b}_1$. For more details we refer to [36, 60].

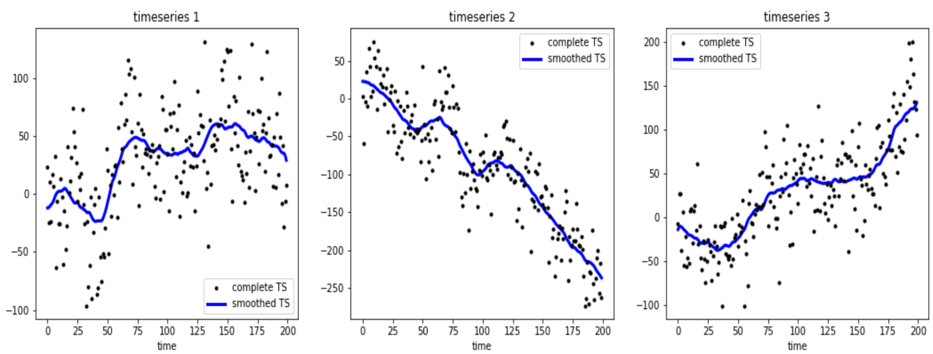


FIGURE 2. Three random walk patterns are generated. The quadratic continuous model s is computed by the smoothing technique QIH-LSQ$(15, 15, 1, 1)$.

Other common techniques include, *Kernel methods* [66], *lowess* [12] and smoothing splines methods. The latter ones can be generally expressed as,

$$\arg\min_s \sum_{j=1}^{m} (x_j - s(t_j))^2 + \gamma \int (s''(t))^2\, dt, \tag{5}$$

where primes denote differentiation. It can be shown that (5) has an explicit, finite-dimensional, unique minimizer which is a natural cubic spline with a knot at each t_j, see [27]. Note that Eq.(5) denotes a spline regression model with a penalization term which controls the uniformity, in terms of parameterization speed, of the achieved representation. Moreover, the degree of smoothness is controlled by the hyper-parameter $\gamma > 0$, see [31, 60] for additional details.

In the present paper we introduce two novel strategies based on the application of quasi-interpolation. Here we proceed with the description. In the first technique, denoted as QIH-LSQ$(k_x, k_{Dx}, d_x, d_{Dx})$, we quasi-interpolate the available realizations and we approximate the first derivative by using second order centered finite differences. Since the realizations are uniformly distributed, in our experiments, the derivatives approximation is carried out by using a constant step size $h = 0.1$,

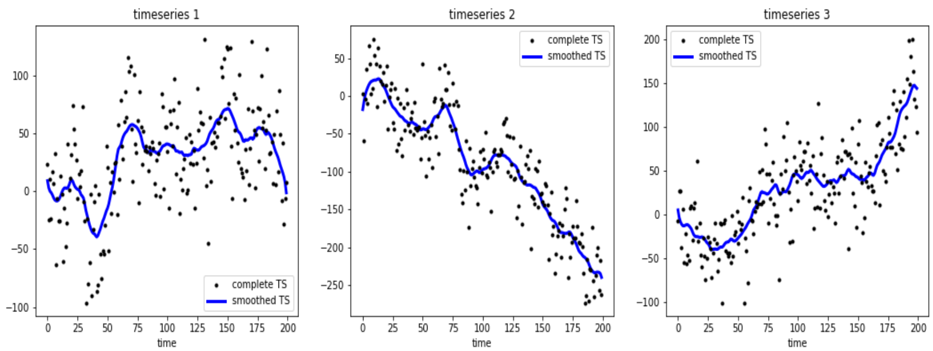


FIGURE 3. The same time series are now approximated with the quadratic continuous model obtained by using QIH-LSQ$(10, 10, 2, 2)$.

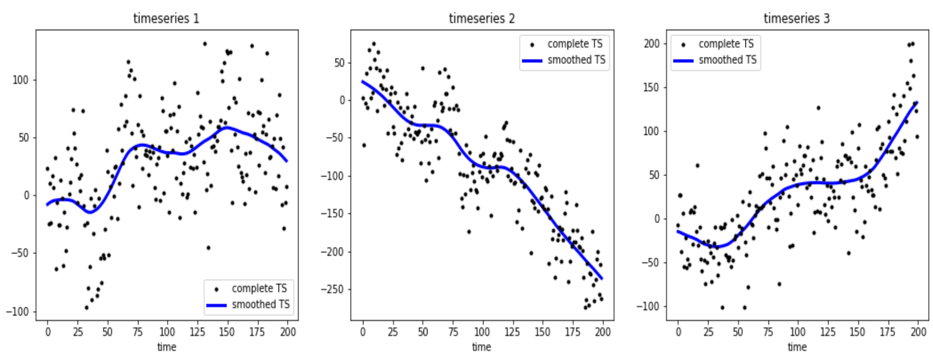


FIGURE 4. Given the original three random walk patterns, the use of QIH-I-LSQ$(15, 15, 1, 1)$ produces a cubic smooth continuous model s.

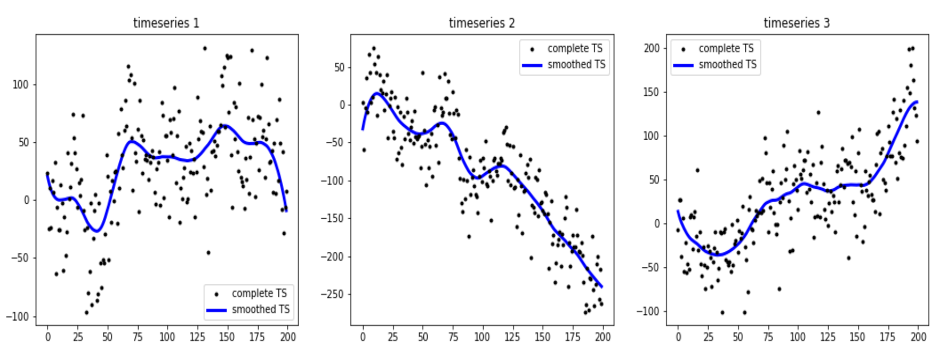


FIGURE 5. We see a cubic, smooth continuous model s generated with QIH-I-LSQ$(10, 10, 2, 2)$ approximating the random walk dataset.

empirically chosen. When the data are very noisy, the realizations can be denoised either by using a standard moving average technique or by computing the coefficients $a_i^{(j)}$ in Eq.(2) by adopting a centered least squares approximation with a polynomial of a given degree. A similar denoising approach is used on the first derivative after its approximation is achieved by using the denoised realizations. Resulting, hence in a double regularization effect.

This method depends on four parameters:

- k_x, k_{Dx}: windows sizes in formula (2) for the realizations and their derivatives, respectively;
- d_x, d_{Dx}: degrees of the least squares approximations.

The second technique, denoted as QIH-I-LSQ(k_x, k_{Dx}, d_x, d_{Dx}), is based on the application of the QI on both, the first and second derivatives which are approximated by a second order centered finite differences scheme. As first step, we compute the QI Hermite spline s' to construct a continuous approximation of the first derivative f'. Secondly,

$$s(t) := \int_{t_1}^{t} s'(z)\,dz - \left(\frac{1}{m} \sum_{j=1}^{m} \int_{t_1}^{t_j} s'(z)\,dz - x_j \right), \tag{6}$$

where the integration constant is chosen as:

$$-\left(\frac{1}{m} \sum_{j=1}^{m} \int_{t_1}^{t_j} s'(z)\,dz - x_j \right).$$

Lemma 4.1. *Given a set of realizations $\{x_t\}$, if the underneath smooth model $s(t)$ is constructed by using the introduced QI operator and formula (6), then*

$$\frac{1}{m} \sum_{i=1}^{m} s(t_i) = \frac{1}{m} \sum_{j=1}^{m} x_j.$$

Proof. By using expression (6) and explicitly writing down the arithmetic mean for $s(t)$, it results that:

$$\begin{aligned}
\frac{1}{m} \sum_{i=1}^{m} s(t_i) &= \frac{1}{m} \left(\sum_{i=1}^{m} \int_{t_1}^{t_i} s'(z)\,dz - \left(\frac{1}{m} \sum_{j=1}^{m} \int_{t_1}^{t_j} s'(z)\,dz - x_j \right) \right) \\
&= \frac{1}{m} \left(\sum_{i=1}^{m} \int_{t_1}^{t_i} s'(z)\,dz - \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{m} \int_{t_1}^{t_j} s'(z)\,dz + \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{m} x_j \right) \\
&= \frac{1}{m} \left(\sum_{i=1}^{m} \int_{t_1}^{t_i} s'(z)\,dz - \sum_{j=1}^{m} \int_{t_1}^{t_j} s'(z)\,dz + \sum_{j=1}^{m} x_j \right) \\
&= \frac{1}{m} \sum_{j=1}^{m} x_j.
\end{aligned}$$

□

The first technique QIH-LSQ produces a smooth spline $s(t)$ by only using the realizations and the approximated f' values. The second procedure QIH-I-LSQ uses the QIH on approximate values for f' to produce a smooth model s' and then uses an integral representation to derive the final smooth model s. The realizations and the derivatives can be denoised by adopting the same strategy used in QIH-LSQ. The parameters for the approximation of the second derivative are chosen equal to the ones of the first derivative for a better computational efficiency.

In the next test we apply the proposed two techniques to a random-walk dataset, generated with the function `sim_randomwalk` from the Python library `tsmoothie`[2]. In Figure 2 we applied QIH-LSQ with parameters $k_x = k_{Dx} = 15$ and $d_x = d_{Dx} = 1$, while in Figure 3, the QIH-LSQ is used with parameters $k_x = k_{Dx} = 10$ and $d_x = d_{Dx} = 2$. We see how raising the degree helps to better delineate the shape of the regularized continuous model. In Figure 4 we adopted QIH-I-LSQ with parameters $k_x = k_{Dx} = 15$ and linear degree, while in Figure 5 the QIH-I-LSQ is used with parameters $k_x = k_{Dx} = 10$ and second degree. We notice how the use of integration in Eq.(6) produces a smoother spline compared to simply using QIH-LSQ. Also, the values for k_x and k_{Dx} can be different; in this example they are set equal, as using different types of combination did not produce any significant change.

5. **Forecasting.**

Time series forecasting is a technique used to predict future values by using the information given by the continuous model based on the previous available data. The smoothing models described in Section 4.2 can be employed as well for this task.

We perform forecasting on two benchmarks datasets reporting the comparison with the simple and double exponential smoothing techniques, described in Section 4.2.

The first dataset contains sheep livestock population in Asia from 1961 to 2007. The time series consists of a total of 38 values, where the last 7 are considered unknown and hence, predicted. The forecasting task is achieved by using QIH-LSQ with parameters $k_x = 0, k_{Dx} = 4, d_x = 1, d_{Dx} = 0$ and QIH-I-LSQ$(0, 1, 0, 0)$. The results are shown in Figure 6, where also the comparisons with single exponential smoothing (SES) with $\alpha = 0.8$ and double exponential smoothing (DES) with parameters $\alpha = 0.8$ and $\beta = 0.8$ are reported. The choice for the setting values for the used parameters was empirically made. From the figure it is clear how the quasi-interpolation follows quite accurately the shape of the distribution of the data, while SES and DES seem affected by some delay. When it comes to forecast the data from year 2000 to 2007, QIH-LSQ and QIH-I-LSQ perform alike to DES; SES, at least for this example, achieved a rather poor accuracy. Just analyzing the predictive power of the proposed techniques, we see in Table 1, that QIH-LSQ and QIH-I-LSQ achieved the lowest normalized root mean square error (NRMSE). The considered example is a challenging one, especially for the first unknown values, where we can clearly observe a trend variation. First order methods such as SES or DES, and Hermite type methods, where the derivative, i.e., the direction, of the last observation is preserved, cannot accurately estimate the effective initial change. Hence, the observed behavior is expected, in practise.

Only analyzing the predictive accuracy, i.e., the preservation of the real values, of the model can have serious limitation under statistical point of view, see e.g., [54]. Therefore, we also included additional tests that can be significant in assessing the goodness of the produced model s. In particular, we performed the two-sample Kolmogorov-Smirnov test (KS-Test) and the Theil's statistics. The KS-Test [32]

[2] https://github.com/cerlymarco/tsmoothie

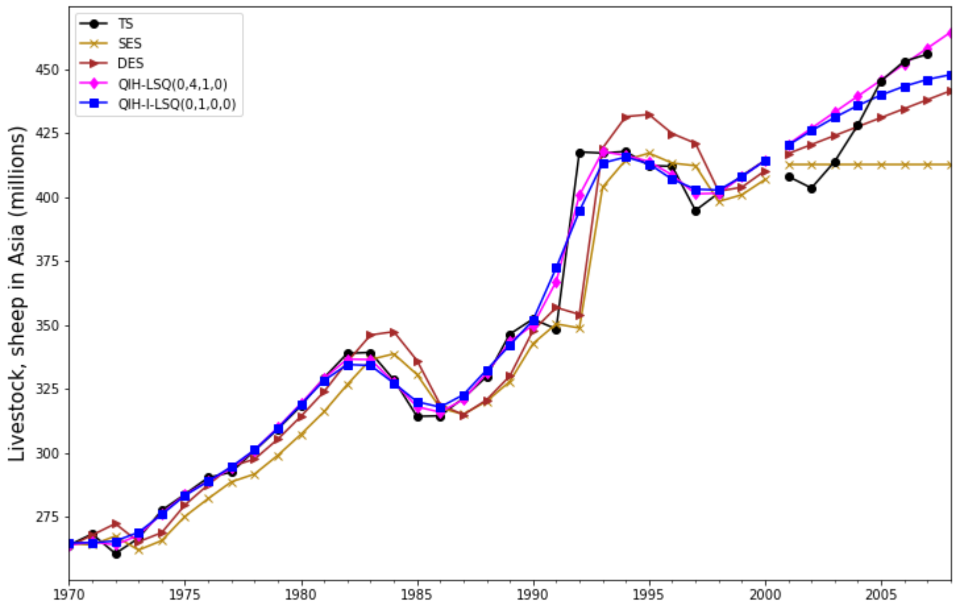


FIGURE 6. Smoothing and forecast of livestock, sheep in Asia. Comparison of SES, DES, QIH-LSQ$(0,4,1,0)$ and QIH-I-LSQ$(0,1,0,0)$.

allows to test the 'null hypothesis' (H_0), i.e., if the two considered distributions (the original one and the produced one) are equal or not. It consists of two output values: `statistic` and `p-value`. If the first output is greater than a critical value D (cf. e.g., [55]), then the two considered distributions are different. For $n = 7$ samples and significance 0.05, $D \approx 0.857$. The more the value for `statistic` is close to D the more likely H_0 cannot hold. Moreover, if the `p-value` is less than a certain threshold of significance, then, the H_0 can be 'rejected', hence, the two distributions are very likely different. Looking at the results of Table 1, intermediate columns, we see how the SES model has the highest `statistic` and, at the same time, a very small `p-value`. The H_0 would be rejected for `p-value`≤ 0.05. In the SES case, the `p-value` is very small but still above the threshold, hence, to be precise we would not be able to reject the null hypothesis. All the other models perform alike. The second chosen statistics is the U_2 Theil's uncertainty coefficient [65]. If $U_2 > 1$, then, the considered model is worse than employing a naive forecasting method. The coefficient U_2 measures the quality of the considered model. Observing the results collected in Table 1, last column, it is clear that SES and DES equivalently fail to produce a good fitting model, while the best quality fit, for this example, is provided by QIH-LSQ. We also computed the U_1 Theil's coefficient [64], but in this case, all the methods achieved comparable results.

The second dataset consists in Synthetic Aperture Radar (SAR) Persistent Scatterer data (PS), which were provided by Planetek-Italia Srl. The PS are time series in which each value measures the millimetric displacement of the ground every six days of an an area in Emilia Romagna within Modena and Bologna province territory. Generally, as told by the domain expert, this type of time series are not

	NRMSE	KS-Test		Theil's
		statistic	p-value	U_2
SES	1.20	0.71	0.053	1.47
DES	0.63	0.43	0.58	0.998
QIH-LSQ	0.59	0.43	0.42	0.69
QIH-I-LSQ	0.60	0.43	0.42	0.76

TABLE 1. Statistics ran on the results for the livestock sheep dataset.

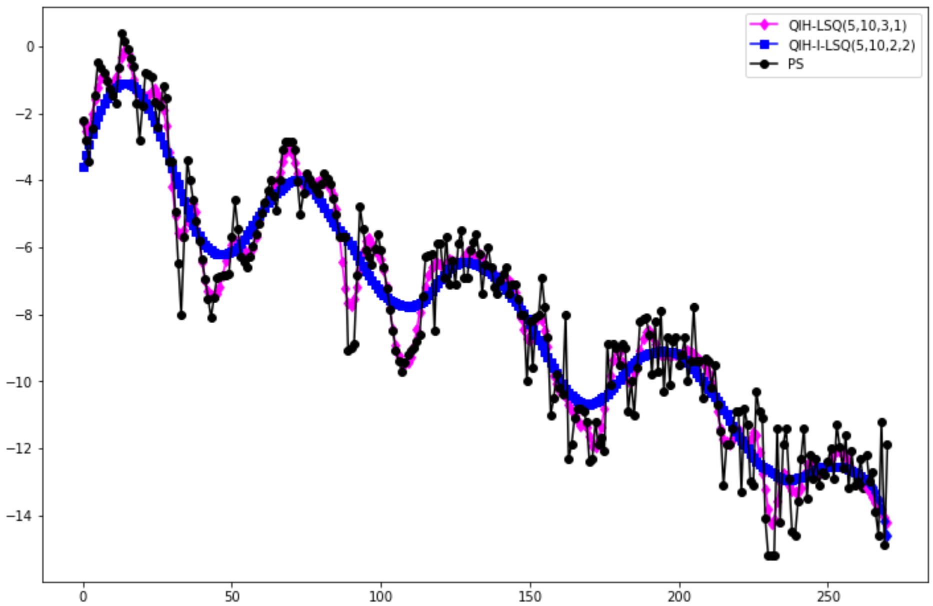


FIGURE 7. PS time series regularized with QIH-LSQ$(5, 10, 2, 1)$ and QIH-I-LSQ$(5, 10, 1, 1)$.

	NRMSE	KS-Test		Theil's
		statistic	p-value	U_2
SES	0.93	0.625	0.087	1.129
DES	0.99	0.625	0.087	0.84
QIH-LSQ	0.98	0.625	0.087	0.86
QIH-I-LSQ	1.10	0.5	0.282	0.749

TABLE 2. Statistics ran on the results for PS time series.

stationary and they are affected by noise, hence a regularization process is necessary before conducting any analysis of the series.

In Figure 7 we show one PS time series consisting of 271 time instants, which has been regularized by adopting the QIH-LSQ, with parameters $k_x = 5$, $k_{Dx} = 10$, $d_x = 2$ and $d_{D_x} = 1$ (magenta curve and diamond shape markers), or by using

QIH-I-LSQ with parameters $k_x = 5$, $k_{Dx} = 10$, $d_x = 1$ and $d_{D_x} = 1$ (blue curve and bullets), on the first 263 points employed as training set. Secondly, a forecasting analysis is performed on the last 8 points of the series. The results are shown in the Figure 8 where a zoom in of the last 8 points and the comparisons with SES ($\alpha = 0.8$) and DES ($\alpha = 0.8$, $\beta = 0.2$) can be seen.

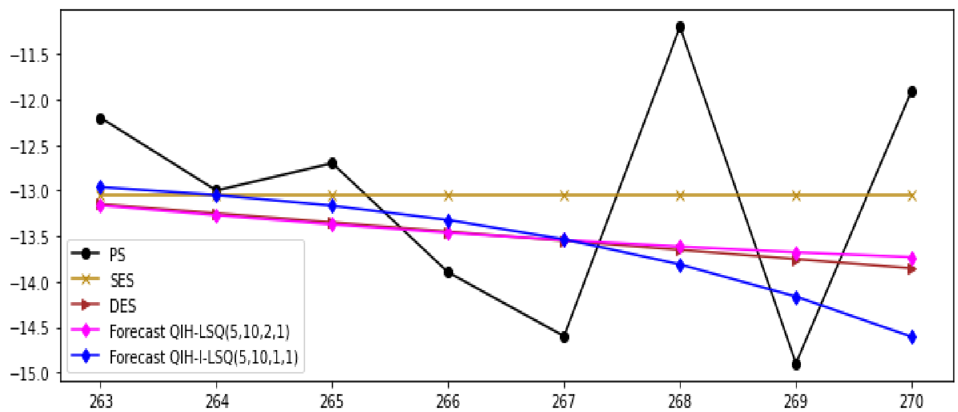


FIGURE 8. Zoom in of the forecasting task performed on the last 8 PS values. Results obtained with QIH-LSQ$(5, 10, 2, 1)$, QIH-I-LSQ$(5, 10, 1, 1)$ SES and DES.

From the picture it is clear that all the methods perform similarly when it comes to the forecasting task. Looking at the statistics reported in Table 2, we see how the NRMSE is quite high for all the considered approaches, assessing that somehow, none of them achieve a good predictive accuracy. However, the KS-Test highlights how SES, DES and QIH-LSQ provide a continuous model that is quite different from the original TS (the critical value $D \approx 0.75$, for $n = 8$ samples and significance equals to 0.05). While, QIH-I-LSQ, also according to U_2 Theil's coefficient, gives the best fit among the considered techniques.

We also tested the forecast algorithms on other time series present in the dataset of Persistent Scatterer obtaining results in line with the example shown here.

6. **Anomaly detection.** Anomaly detection, sometimes referred to as outliers or novelties detection, is the task to identify those observations that somehow deviate from what could be considered the standard behavior of the analyzed distribution. Unsupervised techniques do not need labeled test datasets and they assume that anomalies will occur as sporadic items. Supervised methods require a dataset that has been labeled as "normal" and "abnormal" and involves the need of a trained classifier. Finally, semi-supervised approaches construct a model representing normal behavior from a given normal training dataset, and then test the likelihood of a test instance to be generated by the utilized model. Since we propose an unsupervised anomaly detection method, we will briefly summarize three main algorithms that are largely used within this framework.

The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a non parametric density-based clustering algorithm, introduced for the first time in [19]. It groups together those points that have many nearby neighbors, i.e., high density areas; while it identifies as outliers, those points that lie in low-density regions.

The Local Outlier Factor (LOF) [7] is an unsupervised clustering algorithm which computes the local deviation of a given point from its neighbors. Outliers are identified as those points which have lower density than their neighbors. It shares with DBSCAN some concepts as “reachability distance” and “core distance”.

The Isolation Forest (IF) [39] ‘isolates’ observations by randomly selecting a feature and then randomly selecting a split value between the maximum and minimum values of the selected feature. It does not need any distance or density measure. Anomalies are identified as those random partitions which have a shorter path from the root node, see [40].

Here we present two classes of unsupervised anomaly detection algorithms, the first one is based on the application of the QI to construct a smooth model of the data and any of the clustering techniques described above, where as input to the clustering algorithm we give the differences between the original time series and the smoothed model $s(t)$ and the first derivative $s'(t)$, we call these algorithms QIH-I-DBSCAN, QIH-I-LOF and QIH-I-IF. To detect the anomalies it is important to smooth out only the outliers, so we used only the combination with the QIH-I-LSQ(1,1,1,1), that results the most effective for this purpose. To further improve the algorithm we also present a second procedure that employs the use of dynamic copulas in combination with the same classification algorithms. The algorithms derived by this procedure are called QIH-I-DC-DBSCAN, QIH-I-DC-LOF, QIH-I-DC-IF. In this case we use for the smoothing model only the procedure QIH-I-LSQ(0,0,0,0), that results to be the most effective.

To better understand the second procedure, we firstly recall some basic theoretical concepts about copulas. In probability theory and statistics, copulas are a useful tool to isolate the dependency structure in a multivariate distribution. In particular, we can construct any multivariate distribution by separately specifying the marginal distributions and the copula. See [18], [50] for a detailed description of copulas, here we recall only the definition and the main results. Let us consider a vector of random variables $(X_1, \dots, X_n)$, and let us suppose that the marginal cumulative distribution functions (CDFs) are continuous, i.e., $F_i(x) := \mathbf{P}[X_i \leq x]$ are continuous functions for $i = 1, \dots, n$. By applying the probability integral transform to each component, the new random vector $(U_1, U_2, \dots, U_n) := (F_1(X_1), F_2(X_2), \dots, F_n(X_n))$ has marginal CDFs that are uniformly distributed on the interval $[0, 1]$.

Definition 6.1. A n-dimensional copula, $C : [0; 1]^n \to [0; 1]$ is a CDF with uniform marginals. We write $C(u) = C(u_1; \dots ; u_n)$ for a generic copula.

Copulas establish a connection between multivariate distributions and their univariate margins, as described by the Sklar’s theorem [61]. In particular, we can construct any multivariate distribution F by separately specifying the marginal distributions $F_1, \dots, F_n$ and the copula C, by setting

$$F(x_1, \dots, x_n) := C(F_1(x_1), \dots, F_n(x_n)).$$

Moreover if the marginals are continuous functions, then the constructed copula C is unique. Thanks to Sklar’s theorem we can build very flexible classes of multivariate distribution models. There are many parametric copulas families available, which usually have parameters that control the strength of dependence There are three principal measures of dependence: *Pearson correlation coefficient*, *rank correlation measures*, namely Spearman’s rho and Kendall’s tau, that only depend on the unique copula of the joint distribution and *tail dependence coefficients*, that are a measure of dependence in the extremes of the distributions. In this work we use the Kendall's tau τ_t. It is defined as the difference between the probability of concordance and the probability of discordance, see [37]. Let (X, Y) be a vector of two continuous random variables, then the Kendall's tau for X and Y can be computed as,

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

The copulas are widely used to discover the dependence between two or more time series, especially for portfolio management, risk assessment, option pricing and coverage [5]. Usually the parameters of the copulas are chose statically, in other contexts see e.g., [53] these parameters are modelled in a dynamic way to discover the change in correlation over time. In this case the copulas are called *dynamic* copulas or *time varying* copulas. In our case, we use the dynamic approach, using $x_j - s(t_j)$ and $s'(t_j)$ for $j = 1, \ldots, m$, as time series for which we want to analyze the correlation over time, where s is the cubic spline obtained using the strategy QIH-I-LSQ(0,0,0,0) and s' is its derivative. In order to analyze the time-varying nonlinear correlation, we model the dynamic copula deriving an evolution equation for the Kendall's tau. We choose as best copula in this context the t-Student copula, that it is able to give a greater weight to the extreme values where the anomalies are located. The ARMA model (cf. [53]), with a lag of $q = 10$ orders is used to evaluate the correlation coefficient ρ_t of the Student copula according to the following evolution equation:

$$\rho_t = \Lambda\left(\omega + \beta\rho_{t-1} + \alpha\frac{1}{q}\sum_{j=1}^{q} T^{-1}(u_{t-j}, \nu)T^{-1}(v_{t-j}, \nu)\right), \tag{7}$$

where $\Lambda(x) := (1 - e^{-x})(1 + e^{-x})^{-1}$ is a logistic transformation and T^{-1} is the inverse function of the t-Student distribution with degrees of freedom ν. The logistic transformation is used to obtain ρ_t in $(-1, 1)$; ρ_{t-1} is the autoregressive term and the term multiplied by α represents the forcing variable.

For the estimation of the initial value ρ_0 of the t-Student copula, we adopt a non-parametric way to estimate the marginals and then we apply the maximum likelihood method (MLE). Specifically we use the empirical CDF $\hat{F}(x) := \dfrac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{X_i \leq x}$ with the transformed variables $u_j := \hat{F}(x_j - s(t_j))$ and $v_j := \hat{F}(s'(t_j))$ for $j = 1, \ldots, m$, as the marginal distributions.

The new ρ_t in Eq.(7) is then calculated with the MLE in order to estimate the parameters α, β and ω. Regarding the degrees of freedom ν, we choose to keep them fixed in order to simplify the derived model. The Kendall's tau τ_t time series can now be computed by using the correlation coefficient as $\tau_t = \dfrac{2}{\pi}\arcsin \rho_t$, see [62] for further details. It is worth noting that we use the Kendall's tau time series together with the difference between the original time series and the smoothed time series $s(t)$ as input in the selected algorithms to detect anomalies. By providing also the Kendall's tau time series as input, the anomaly detection improves compared to using only the original time series, as it is shown in the next examples. We ran the experiments by using the Python module `Scikit-learn` and the copulae libraries `Copulae`[3] and `pyvinecopulib`[4]. For the experiments we use one of the few publicly available dataset, Yahoo! Webscope S5 dataset[5]. It consists of four data classes, each containing either a set of synthetic or real web traffic values tagged with anomalies. Data class A2Benchmark, A3Benchmark, A4Benchmark represent synthetic time series of different length. Data class A1Benchmark is the most diverse one due to its real nature. It consists of time series representing metrics of various Yahoo! services.

To show the behavior of the algorithm we start by choosing some different time series from the classes. We first show an example with four time series: two time series are selected from the real-values dataset and two are taken from the synthetic one. Although the whole dataset contains both, the time series of the values and the vector with elements in $\{0, 1\}$ where 0 indicates the absence of anomaly and 1 the anomalous point, we conduct this example in a completely unsupervised fashion.

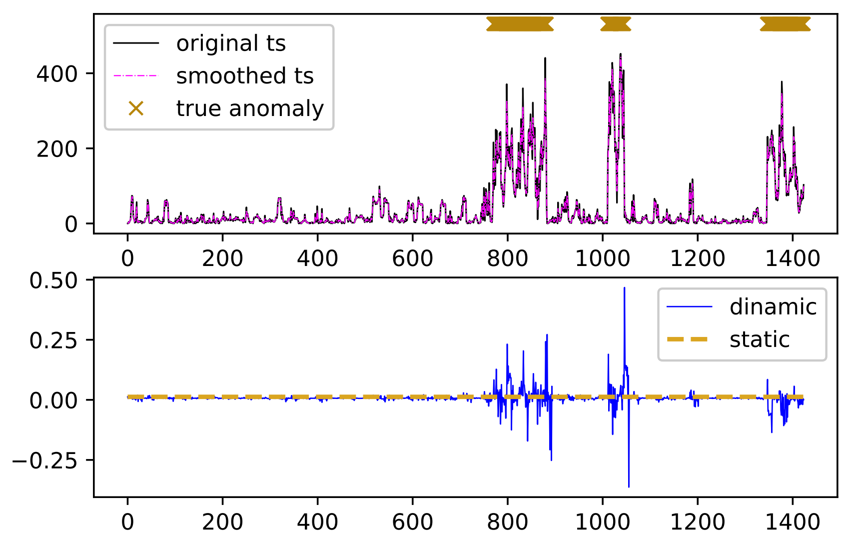


FIGURE 9. A1Benchmark-real19 time series and Kendall's tau time-varying copula.

The choice of the Kendall's tau together with the smooth function and its smooth derivative turns out to be very interesting. In fact, as it can be seen in the Figures 9, 10, 11, 12, it is clear how the peaks of the Kendall's tau evolution equation match with the anomalous points of the considered time series.

To show the different performance between QIH-I-* and QIH-I-DC-* in Figure 13 we plot the scatterplot for the time series A4Benchmark-TS10 of the input and output data of the two procedures, by showing in Figure 13 (a) the QIH-I-LOF and

[3] https://copulae.readthedocs.io/en/latest/

[4] https://vinecopulib.github.io/pyvinecopulib/

[5] The dataset can be downloaded at https://yahooresearch.tumblr.com/post/114590420346/a-benchmark-dataset-for-time-series-anomaly

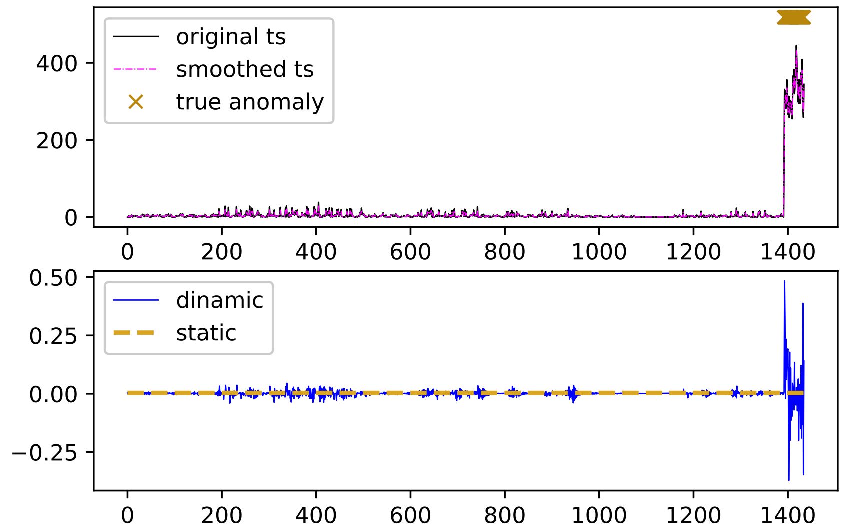


FIGURE 10. A1Benchmark-real25 time series and Kendall's tau time-varying copula.

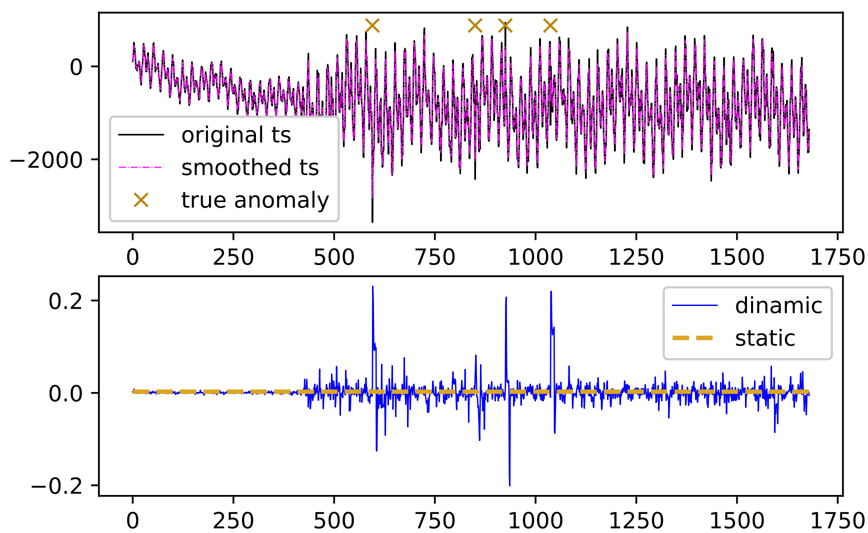


FIGURE 11. A4Benchmark-TS10 time series and Kendall's tau time-varying copula.

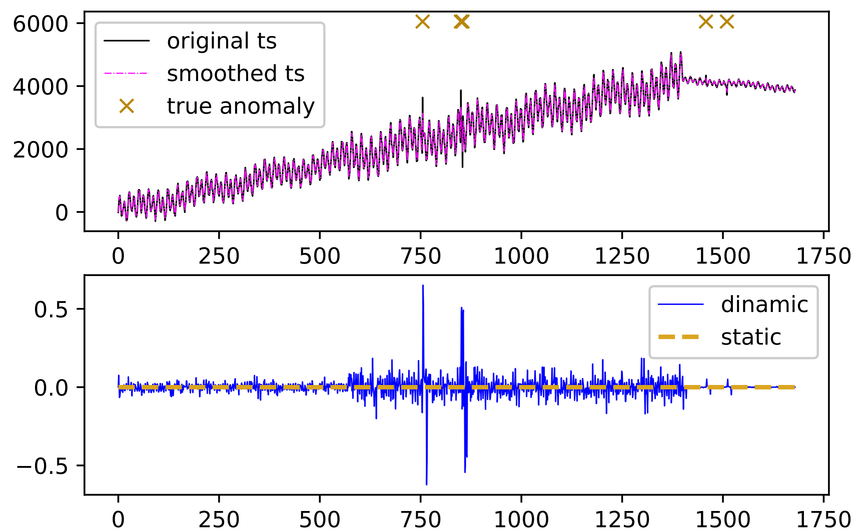
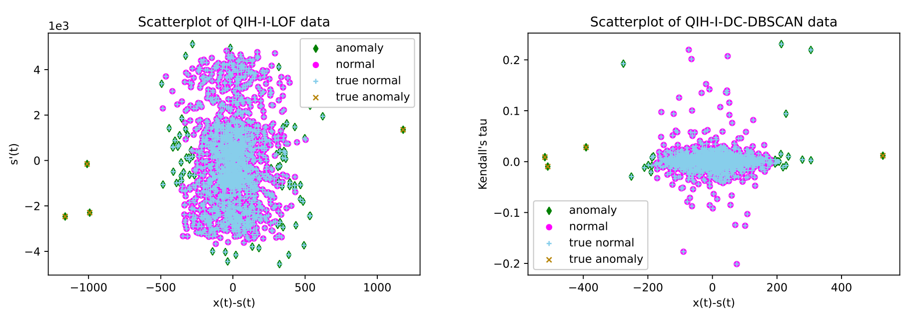


FIGURE 12. A4Benchmark-TS11 time series and Kendall's tau time-varying copula.



(a) (b)

FIGURE 13. Scatterplot of $x(t) - s(t)$ vs $s'(t)$ for the QIH-I-LOF (a) and of $x(t) - s(t)$ vs Kendall's tau (b) for QIH-I-DC-DBSCAN, both for problem A4Benchmark-TS10. Computed normal points: magenta bullet. Computed anomalies: green diamond. True normal behavior: blue +. True anomalies: yellow x.

in Figure 13 (b) the QIH-I-DC-DBSCAN. It is evident how they use of Kendall's tau simplifies the task of identifying anomalies, as there is a minor dispersion within the normal points.

Finally we use the ground-truth present in the dataset, in order to compare the results obtained with our approach and with DBSCAN, LOF and IF applied on the original time series, in terms of RECALL, Overall Accuracy (OA) and ROC-AUC score. The first metric, intuitively, measures the ability to correctly identify all the anomalous samples. The OA metric computes the fraction of correct predictions over the total amount of samples. Finally, the ROC-AUC score is usually employed to assess the performance of a binary classifier by varying a discrimination threshold [24].

We stress that the experiments are conducted in a completely unsupervised fashion and that the aforementioned metrics are computed only at the end, when we can use the correct labels present in the ground-truth to evaluate the performance of our approach.

	RECALL	OA	ROC-AUC
QIH-I-DBSCAN	0.917	0.942	0.929
QIH-I-DC-DBSCAN	0.939	0.980	0.959
DBSCAN	0.984	0.067	0.523
QIH-I-LOF	0.973	0.955	0.964
QIH-I-DC-LOF	0.897	0.984	0.940
LOF	0.208	0.991	0.601
QIH-I-IF	0.940	0.791	0.865
QIH-I-DC-IF	0.958	0.882	0.920
IF	0.620	0.728	0.674

TABLE 3. Mean values of the recall, overall accuracy and ROC-AUC for the A4Benchmark for the compared algorithms.

Regarding DBSCAN, LOF and IF finding the correct input parameters it is not a trivial task, therefore we decided to test these algorithms using the available routines with the default parameters and giving as input the original time series. For our approach, regarding the DBSCAN, we have used a special routine to set the `eps` parameter. This is the most important parameter as it represents the threshold distance which discriminates outliers from normal samples. As it can be seen in Table 3, the default value for `eps`, which is equal to 0.5, used applying DBSCAN to the original time series is too small and as a result most of the points are considered as anomalies. The information given by the absolute error between the original time series and the smoothed one is used to set up the `eps` value. In particular, for the algorithm QIH-I-DBSCAN we used `eps` $= mean(|x - s(t)|)$. For the algorithm QIH-I-DC-DBSCAN we empirically decided to divide the time interval in 5 equals subintervals $I_j, j = 1, \ldots, 5$ and we consider as `eps` $= \min \left\{ mean(|x - s(t)|), mean(|x - s(t)|)|_{I_j}, j = 1, \ldots, 5 \right\} /5$. This technique allows to select a good value for `eps` even when the time series changes behavior as shown in figures 11–12.

Table 3 shows the potentiality of the algorithms that combines Hermite quasi-interpolation smoothing and dynamic copulas with clusterization algorithms, that for the chosen benchmark dataset, outperforms the using of any clustering algorithms employed on the original dataset. We observe that all the proposed algorithms could have a better behavior choosing in the correct way the parameters, but it is not our intention to make an effective comparison, that we know is usually very difficult to achieve, but we only want to show the potentiality of the proposed algorithms. In particular, we see that for some classes of problems the DBSCAN, LOF and IF algorithms are also very effective, but they have poor behavior for other examples, that are, however, handled very well by the combined algorithms.

7. **Conclusions.** Times series usually refers to stochastic processes of data collected in real scenarios, and hence, affected by noise or by error measurements. In this paper we present a valid tool based on Hermite quasi-interpolation which has a twofold use. On the one hand, we can successfully preprocess the considered time series by performing the operations of imputation and smoothing. On the other hand, we can provide a smooth model of the time series that can be applied to forecast and to anomaly detection tasks. In all the considered examples we obtained promising results compared to standard techniques. This is a first step for a deeper analysis of the application of the proposed QI scheme for the study of time series in different contexts.

Acknowledgments. We thank Planetek Italia s.r.l. for PS data. The authors are members of the INdAM Research group GNCS. This work was also developed within the project "Modelli e metodi per l'analisi di dati complessi e voluminosi" of the University of Bari.

REFERENCES

[1] W. Aigner, S. Miksch, H. Schumann and C. Tominski, *Visualization of Time-Oriented Data*, Springer Science & Business Media, 2011.

[2] A. Andrisani, R. M. Mininni, F. Mazzia, G. Settanni, A. Iurino, S. Tangaro, A. Tateo and R. Bellotti, Applications of PDEs inpainting to magnetic particle imaging and corneal topography, *Opuscula Mathematica*, **39** (2019), 453–482.

[3] T. Andrysiak, L. Saganowski and W. Mazurczyk, Network anomaly detection for railway critical infrastructure based on autoregressive fractional integrated moving average, *EURASIP Journal on Wireless Communications and Networking,* **245** (2016), 1–14.

[4] R. Armina, A. M. Zain, N. A. Ali and R. Sallehuddin, A review on missing value estimation using imputation algorithm, In *Journal of Physics: Conference Series,* **892** (2017).

[5] N. Benlagha and L. Noureddine, A time-varying copula approach for modelling dependency: New evidence from commodity and S & P500 markets, *Journal of Multinational Financial Management*, **892**, (2016).

[6] G. E. Box, G. M. Jenkins, G. C. Reinsel and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, John Wiley & Sons, 2016.

[7] M. M. Breunig, H. P. Kriegel, R. T. Ng and J. Sander, LOF: identifying density-based local outliers, In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, **29** (2000), 93–104.

[8] F. Calabrò, A. Falini, M. L. Sampoli and A. Sestini, Efficient quadrature rules based on spline quasi-interpolation for application to IGA-BEMs, *Journal of Computational and Applied Mathematics*, **338** (2018), 153–167.

[9] V. Chandola, A. Banerjee and V. Kumar, Anomaly detection: A survey, *ACM Computing Surveys (CSUR)*, **41** (2009), 1–58.

[10] M. P. Clements, P. H. Franses and N. R. Swanson, Forecasting economic and financial time-series with non-linear models, *International Journal of Forecasting*, **20** (2004), 169–183.

[11] W. P. Cleveland and G. C. Tiao, Decomposition of seasonal time series: A model for the census X-11 program, *Journal of the American Statistical Association*, **71** (1976), 581–587.

[12] W. S. Cleveland, Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, **74** (1979), 829–836.

[13] J. Contreras, R. Espinola, F. J. Nogales and A. J. Conejo, ARIMA models to predict next-day electricity prices, *IEEE Transactions on Power Systems*, **18** (2003), 1014–1020.

[14] C. de Boor, *Splines as Linear Combinations of B-Splines*, Lorentz, G.G., et al. (eds.) Approximation Theory II, pp. 1–47. Academic Press, San Diego, 1976.

[15] C. de Boor, *A Practical Guide to Splines*, revised edn. Springer, Berlin, 2001.

[16] C. de Boor and M. G. Fix, Spline approximation by quasi-interpolants, *J. Approx. Theory*, **8** (1973), 19–54.

[17] A. M. De Livera, R. J. Hyndman and R. D. Snyder, Forecasting time series with complex seasonal patterns using exponential smoothing, *Journal of the American Statistical Association*, **106** (2011), 1513–1527.

[18] F. Durante and C. Sempi, *Principles of Copula Theory,* 1^{st} edition, Chapman and Hall/CRC, New York, 2015.

[19] M. Ester, H. P. Kriegel, J. Sander and X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, *In Kdd,* **96** (1996), 226–231.

[20] A. Falini, C. Giannelli, T. Kanduč, M. L. Sampoli and A. Sestini, An adaptive IgA-BEM with hierarchical B-splines based on quasi-interpolation quadrature schemes, *Internat. J. Numer. Methods Engrg.*, **117** (2019), 1038–1058.

[21] A. Falini, G. Castellano, C. Tamborrino, F. Mazzia, R. M. Mininni, A. Appice and D. Malerba, Saliency detection for hyperspectral images via sparse-non negative-matrix-factorization and novel distance measures, In *2020 IEEE Conference on Evolving and Adaptive Intelligent Systems*, (EAIS) (2020), 1–8.

[22] A. Falini and T. Kanduč, A study on spline quasi-interpolation based quadrature rules for the isogeometric Galerkin BEM, In *Advanced Methods for Geometric Modeling and Numerical Simulation*, Springer, Cham., (2019), 99–125.

[23] A. Falini, C. Tamborrino, G. Castellano, F. Mazzia, R. M. Mininni, A. Appice and D. Malerba, Novel reconstruction errors for saliency detection in hyperspectral images, In *International Conference on Machine Learning, Optimization, and Data Science*, Springer, Cham. (2020), 113–124.

[24] T. Fawcett, An introduction to ROC analysis, *Pattern Recognition Letters*, **27** (2006), 861–874.

[25] T. C. Fu, A review on time series data mining, *Engineering Applications of Artificial Intelligence*, **24** (2011), 164–181.

[26] M. Gavrilov, D. Anguelov, P. Indyk and R. Motwani, Mining the stock market (extended abstract) which measure is best?, In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2000, 487–496.

[27] P. J. Green and B. W. Silverman, *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, Monographs on Statistics and Applied Probability, 58. Chapman & Hall, London, 1994.

[28] H. S. Guirguis and G. A. Felder, Further advances in forecasting day-ahead electricity prices using time series models, *KIEE International Transactions on Power Engineering*, **4** (2004), 159–166.

[29] J. J. Guo and P. B. Luh, Selecting input factors for clusters of Gaussian radial basis function networks to improve market clearing price prediction, *IEEE Transactions on Power Systems*, **18** (2003), 665–672.

[30] W. Härdle, H. Lütkepohl and R. Chen, A review of nonparametric time series analysis, *International Statistical Review*, **65** (1997), 49–72.

[31] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics. Springer, New York, 2009.

[32] J. L. Hodges, Jr., The significance probability of the Smirnov two-sample test, *Ark. Mat.*, **3** (1958), 469–486.

[33] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*. OTexts, 2018.

[34] H. Joe, *Dependence Modeling with Copulas*, Monographs on Statistics and Applied Probability, 134. CRC Press, Boca Raton, FL, 2015.

[35] R. H. Jones, Maximum likelihood fitting of ARMA models to time series with missing observations, *Technometrics*, **22** (1980), 389–395.

[36] P. S. Kalekar, Time series forecasting using holt-winters exponential smoothing, *Kanwal Rekhi School of Information Technology*, **4329008** (2004), 1–13.

[37] M. G. Kendall, *Rank Correlation Methods,* Griffin, 1948.

[38] W. Kim, B.-J. Choi, E.-K. Hong, S.-K. Kim and D. Lee, A taxonomy of dirty data, *Data Min. Knowl. Discov.*, **7** (2003), 81–99.

[39] F. T. Liu, K. M. Ting and Z. H. Zhou, Isolation forest, In *2008 Eighth IEEE International Conference on Data Mining*, (2008), 413–422.

[40] F. T. Liu, K. M. Ting and Z.-H. Zhou, Isolation-based anomaly detection, *ACM Transactions on Knowledge Discovery from Data*, **6** (2012), 1–39.

[41] T. Lyche and L. L. Schumaker, Local spline approximation, *J. Approx. Theory*, **15** (1975), 294–325.

[42] F. Mazzia and A. Sestini, The BS class of Hermite spline quasi-interpolants on nonuniform knot distributions, *BIT Numerical Mathematics*, **49** (2009), 611–628.

[43] F. Mazzia and A. Sestini, Quadrature formulas descending from BS Hermite spline quasi-interpolation, *J. Comput. Appl. Math.*, **236** (2012), 4105–4118.

[44] F. Mazzia, A. Sestini and D. Trigiante, B-spline multistep methods and their continuous extensions, *SIAM J. Numer. Anal.*, **44** (2006), 1954–1973.

[45] F. Mazzia, A. Sestini and D. Trigiante, BS linear multistep methods on non-uniform meshes, *JNAIAM J. Numer. Anal. Ind. Appl. Math.*, **1** (2006), 131–144.

[46] F. Mazzia, A. Sestini and D. Trigiante, The continous extension of the B-spline linear multistep methods for BVPs on non-uniform meshes, *Appl. Numer. Meth.*, **59** (2009), 723–738.

[47] S. Moritz and T. Bartz-Beielstein, ImputeTS: Time series missing value imputation in R, *R. J.*, **9** (2017), 207–218.

[48] F. Muharemi, D. Logofătu and F. Leon, Review on general techniques and packages for data imputation in R on a real world dataset, In *International Conference on Computational Collective Intelligence*, 2018, 386–395, Springer, Cham.

[49] K. R. Müller, A. J. Smola, G. Rätsch, B. Schölkopf, J. Kohlmorgen and V. Vapnik, Predicting time series with support vector machines, In *International Conference on Artificial Neural Networks*, 1997. Springer, Berlin, Heidelberg.

[50] R. B. Nelsen, *An Introduction to Copulas,* 2^{nd} edition, Springer Publishing Company, 2006.

[51] T. Niimura, H.-S. Ko and K. Ozawa, A day-ahead electricity price prediction based on a fuzzy-neuro autoregressive model in a deregulated electricity market, In *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02*, **2** (2002), 1362–1366.

[52] P. Omenzetter and J. M. W. Brownjohn, Application of time series analysis for bridge monitoring, *Smart Materials and Structures*, **15** (2006), 129.

[53] A. J. Pattom, Modelling asymmetric exchange rate dependence, *Internat. Econom. Rev.,* **10** (2006), 527–556.

[54] B. Ramosaj, L. Amro and M. Pauly, A cautionary tale on using imputation methods for inference in matched-pairs design, *Bioinformatics*, **36** (2020), 3099–3106.

[55] F. J. Rohlf and R. R. Sokal, *Statistical Tables*, Macmillan, 1995.

[56] P. Sablonnière, Positive spline operators and orthogonal splines, *J. Approx. Theory*, **52** (1988), 28–42.

[57] P. Sablonnière, Univariate spline quasi-interpolants and applications to numerical analysis, *Rend. Semin. Mat. Univ. (Torino)*, **63** (2005), 211–222.

[58] P. Sablonnière and D. Sbibih, Integral spline operators exact on polynomials, *Approx. Theory Appl.*, **10** (1994), 56–73.

[59] X. Shao, Self-normalization for time series: A review of recent developments, *J. Amer. Statist. Assoc.*, **110** (2015), 1797–1817.

[60] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and its Applications*, Fourth edition, Springer, Cham, 2017.

[61] M. Sklar, *Fonctions de Répartition à n Dimensions et Leurs Marges*, Publ. Inst. Statist. Univ. Paris, **8** (1959), 229–231.

[62] M. K. P. So and C. Y. T. Yeung, Vine-copula GARCH model with dynamic conditional dependence, *Comput. Statist. Data Anal.,* **76**, (2014), 655–671.

[63] M. Teng, Anomaly detection on time series, In *2010 IEEE International Conference on Progress in Informatics and Computing*, **1** (2010), 603–608.

[64] H. Theil, *Economic Forecasts and Policy*, North-Holland Pub. Co., 1961.

[65] H. Theil, *Applied Economic Forecasting*, North-Holland Pub. Co., 1971.

[66] G. S. Watson, Smooth regression analysis, *Sankhyā: The Indian Journal of Statistics, Series A*, **26** (1964), 359–372.

[67] C. Zhang, D. Song, Y. Chen, X. Feng, C. Lumezanu, W. Cheng, J. Ni, B. Zong, H. Chen and N. V. Chawla, A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data, In *Proceedings of the AAAI Conference on Artificial Intelligence,* **33** (2019), 1409–1416.

Received August 2021; revised December 2021; early access February 2022.

E-mail address: antonella.falini@uniba.it
E-mail address: francesca.mazzia@uniba.it
E-mail address: cristiano.tamborrino@uniba.it