Computer Science and Information Systems 11(2):881–904 DOI: 10.2298/CSIS121215048G

Joint Propagation of Ontological and Epistemic Uncertainty across Risk Assessment and Fuzzy Time Series Models

Vasile Georgescu [1]

[1] Department of Statistics and Informatics, University of Craiova, 13 A.I.Cuza, Craiova, 200375, Romania
v_geo@yahoo.com

Abstract. This paper discusses hybrid probabilistic and fuzzy set approaches to propagating randomness and imprecision in risk assessment and fuzzy time series models. Stochastic and Computational Intelligence methods, such as Probability bounds analysis, Fuzzy α-levels analysis, Fuzzy random vectors, Wavelets decomposition and Wavelets Networks are combined to capture different kinds of uncertainty. Their most appropriate applications are probabilistic risk assessments carried out in terms of probability distributions with imprecise parameters and stochastic processes modeled in terms of fuzzy time series.

Keywords: risk assessment, fuzzy time series, probability bounds analysis, fuzzy random vectors, wavelets, Hukuhara difference.

1. Introduction: Challenges, Criticism and Rationales for a Novel Approach

Two kinds of uncertainty are contrasted in this paper (ontological, vs. epistemic uncertainty) and several techniques are addressed in order to capture and propagate both of them jointly across a specific model. The most relevant areas of applications carried out with such hybrid approaches are presented in what follows.

The first application area is addressed when attempting to extend the classical *probabilistic risk assessment* (PRA) modeling framework in such a way that allows probability distribution parameters to be imprecisely defined. Basically, PRA uses probability models to represent the likelihood of different risk levels in a population (i.e., randomness). In the standard probabilistic approach, inputs to the risk equation are described as *random variables* that can be defined mathematically by a probability distribution. The CDF for risk can be especially informative for illustrating the percentile corresponding to a particular risk level of concern. However, the presence of imprecision in a risk model adds another dimension to that of randomness and requires more comprehensive approaches to capture and represent the range within which the risk distribution might vary. Analytic methods for propagating the uncertainty (such as Probability bounds analysis, Fuzzy α-levels analysis) as well as stochastic simulation

techniques, can be combined or integrated to reach synergy. This may lead to hybrid simulation frameworks, such as Fuzzy Monte Carlo, in an attempt to find the output of a model that has both random variables (given by probability distributions) and fuzzy variables for the inputs.

The second application area is addressed when attempting to capture the inherent fuzzy and random nature of some stochastic processes, expressed in terms of fuzzy time series. Unfortunately, modeling, estimating and forecasting fuzzy time series faces the problem of non-invertibility of the standard Minkowski addition and multiplication by scalars in a fuzzy framework.

In contrast with the case of real numbers, for some set-defined quantities, such as intervals and fuzzy sets, the opposite of A is not the inverse of A in Minkovsky addition (unless $A=\{a\}$ is a singleton). This implies that, in general, additive simplification is not valid, i.e., $(A+C=B+C)\not\Rightarrow A=B$, or $(A+B)-B\neq A$.

To partially overcome this situation, the Hukuhara difference has been proposed instead of fuzzy subtraction, assuming that there exists a set C for which $C=A\,\sigma_{\mathrm{H}}\,B \Leftrightarrow A=B+C$.

Let $\mathcal{K}_C(\Re^p)$ be the class of the non-empty compact convex subsets of $\Re^p$. An important property of "σ_{H}" is that $A\,\sigma_{\mathrm{H}}\,A=\{0\}\quad \forall A\in\mathcal{K}_C(\Re^p)$ and $(A+B)\,\sigma_{\mathrm{H}}\,B=A\quad \forall A,B\in\mathcal{K}_C(\Re^p)$. The H-difference is unique, but it does not always exist. A necessary condition for $A\,\sigma_{\mathrm{H}}\,B$ to exist is that A contains a translate $\{c\}+B$ of B.

Unfortunately, even if the Hukuhara difference exists, some distortions may still appear when applying least squares estimation.

Several generalizations of Hukuhara difference have been proposed in an attempt to obtain a more tractable way to deal with fuzzy regression analysis, assuming that fuzzy estimates can be still obtained from the condition of minimizing the sum of square residuals, expressed as a difference between two fuzzy quantities: the response of a system and its model based estimation.

The standard assumption is to consider square-integrable random variables defined on a Hilbert space equipped with a suitable L_2-metric that allows the projection theorem to be still valid. However, the projection cannot be properly applied as usually onto a subspace, but rather onto cones (i.e., subject to some constraints), due to the lack of a general additive inverse in the space of fuzzy variables, which is only a semi-linear space. This may lead to distorted results such as obtaining fuzzy least squares estimates with negative spreads.

The first and most notable generalizations of Hukuhara difference proposed in the context of fuzzy linear regression was that of Diamond [2], or Diamond and Körner [4], as the least squares solution of equation $A+X=B$, i.e., $C=B\,\sigma_{\mathrm{H}}\,A$, if and only if $d(A+C,B)=\inf_{X\in\mathcal{Y}} d(A+X,B)$, in some L_2-type metric space $(\mathcal{Y},d)$.

When the usual Hukuhara difference $B\,\sigma_{\mathrm{H}}\,A$ exists, it coincides with the least squares solution defined above.

However, the extension of Hukuhara difference to an L_2-approximant, as proposed by Diamond, has some important limitations. The main criticism I can address to this approach is that it attempts to project the fuzzy data onto a projection cone as a whole, thus conserving a rigid structure and imposing constraints that are still too strong. Instead, I propose a suitable new method, based upon a partial decoupling principle. This exploits and extends an idea I introduced in 1998 (see [10]). It allows decomposing the monolithic fuzzy model into several more tractable crisp estimation sub-problems, starting from that one corresponding to modal values ($\alpha = 1$) in fuzzy data, and then proceeding in a decremental way for left and right α-level bounds, with α progressively decreasing towards 0. The estimates of modal values are not subject to any constraints, thus being obtained by applying the Hilbert space projection theorem directly onto the corresponding subspace. However, the estimates for the left and right α-level bounds can only be obtained by applying the projection theorem onto cones, in such a way to obtain least squares estimates without negative spreads. This leads to constrained quadratic programs, conveniently defined.

When applying to fuzzy time series estimation, the proposed approach is not only able to help decomposing, but also to make the process invertible, by recomposing a non-stationary fuzzy time series from its components, such as trend, cycle, seasonality and the simulated residuals, all of them properly defined as LR-fuzzy sets.

As an alternative to fuzzy estimation methods, computational intelligence techniques, based on wavelet decomposition and wavelet networks for nonlinear model fitting have been proposed to address fuzzy time series estimation and prediction.

2. Ontological versus Epistemic Uncertainty

Although the terms *imprecision* and *uncertainty* are often contrasted by fuzzy sets theorists, the latter is commonly used in a broad range of contexts with confusing connotations, varying from very specific to rather generic ones. According to EPA's guidance [6], an essential distinction is to be made between ontological and epistemic uncertainty.

Ontological uncertainty (also called randomness, variability, or aleatory / objective / irreducible uncertainty) arises from natural stochasticity, environmental variation across space or through time, genetic heterogeneity among individuals, and other sources of randomness.

Although randomness can often be better characterized by further specific study, it is not generally reducible by empirical effort. Randomness can be translated into risk (i.e., probability of some adverse consequence) by the application of an appropriate probabilistic model.

Epistemic uncertainty (also called subjective / reducible uncertainty) arises from incomplete knowledge about the world. Sources of epistemic uncertainty include measurement uncertainty, small sample sizes, detection limits and data censoring, ignorance about the details of the mechanisms and processes involved and other imperfections in scientific understanding. Epistemic uncertainty can in principle be reduced by focused empirical effort. It cannot be translated into probability, but it can

be used in hybrid approaches to generate bounds on probability distributions. Such bounds may be either crisp (e.g., probability bounds, where the interval arithmetic is used to propagate the imprecision regarding distribution parameters) or fuzzy (e.g., fuzzy randomness, where both the machineries of probability theory and fuzzy sets theory are combined).

3. Hybrid Representations of Ontological and Epistemic Uncertainty

3.1. Computing with Probability Bounds

Probability bounds analysis combines probability theory and interval arithmetic to produce probability boxes (p-boxes), structures that allow the comprehensive propagation of both randomness and epistemic uncertainty through calculations in a rigorous way.

If we have only partial information about the probability distribution, then we cannot compute the exact values $F(x)$ of the CDF. Instead, we can circumscribe $F(x)$ by a pair of functions $\underline{F}(x)$ and $\overline{F}(x)$, each one representing a CDF, which bounds the (unknown) actual CDF. Such a pair of CDFs is called a *probability bound*, or a *p-bound*, for short. For every x, the possible values of the probability $F(x)$ belongs to the interval $[\underline{F}(x), \overline{F}(x)]$.

In computations, it is often convenient to express a p-box in terms of its inverse functions ℓ and u defined on the interval of probability levels [0,1]. The function u is the inverse function of the upper bound on the distribution function and ℓ is the inverse function of the lower bound. These monotonic functions are bounds on the inverse of the unknown distribution function F

$$\ell(p) \geq F^{-1}(p) \geq u(p) \quad \textbf{(1)}$$

where p is probability level. Note that ℓ corresponds to $\overline{F}$ and u to $\underline{F}$.

It is simple to compute probability bounds for many cases in which the distribution family is specified, but only interval estimates can be given for the parameters. For instance, suppose that, from previous knowledge, it is assumed that a distribution is normal, but the precise values of the parameters that would define this distribution are uncertain. If there exist bounds on μ and σ (mean and standard deviation), bounds on the distribution can be obtained by computing the envelope of all normal distributions that have parameters within the specified intervals. These bounds are

$$\ell(p) = \max_{\theta} F_{\theta}^{-1}(p); \quad u(p) = \min_{\theta} F_{\theta}^{-1}(p) \quad \textbf{(2)}$$

where

$$\theta \in \{(\mu, \sigma) \mid \mu \in [\mu_\ell, \mu_u], \sigma \in [\sigma_\ell, \sigma_u]\} \quad (3)$$

and F is the CDF of a normal distribution with such parameters. In principle, making these calculations might be a difficult task since θ indexes an infinite set of distributions. However, in practice, finding the bounds requires computing the envelope over only four distributions: those corresponding to the parameter sets (μ_ℓ, σ_ℓ), (μ_ℓ, σ_u), (μ_u, σ_ℓ), and (μ_u, σ_u), as is shown in Figure 1. This simplicity is the result of how the family of distributions happens to be parameterized by μ and σ.

Nevertheless, it is just as easy to find probability bounds for cases with other commonly used distribution families such as lognormal, uniform, exponential, Cauchy, and many others.

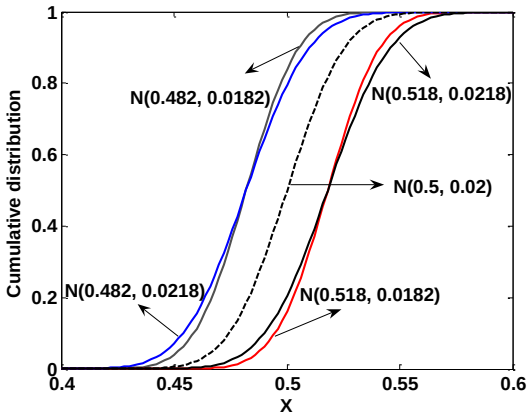

Fig. 1. Bounds on the CDF of a normal distribution with μ = [0.482, 0.518] and σ = [0.0182, 0.0218]. The dotted line is the CDF for the normal distribution with μ=0.5 and σ=0.02.

3.2. Fuzzy Random Variables and Processes

I consider an extension of the probability space $[\Omega; \mathcal{A}; P]$ by the dimension of fuzziness, i.e., by introducing a membership scale. This enables the consideration of imprecise observations as fuzzy realizations $\tilde{x}(\omega) = (\tilde{x}_1, \ldots, \tilde{x}_n) \subseteq X$ of each elementary event $\omega \in \Omega$. The attention will be restricted to the class $\mathcal{F}_C(\Re)$ of normal convex fuzzy sets on $\Re$, whose α-level sets are in the class $\mathcal{K}_C(\Re)$ of nonempty compact real intervals.

A *fuzzy* random variable $\tilde{X}$ is the fuzzy result of the uncertain mapping $\tilde{X}: \Omega \to \mathcal{F}_C(\Re)$, such that for each $\alpha \in [0,1]$ and $\omega \in \Omega$, the α-level intervals $X_\alpha(\omega) = [\inf(X(\omega))_\alpha, \sup(X(\omega))_\alpha]$, generated by the mapping $X_\alpha: \Omega \to \mathcal{K}_C(\Re)$,

are (compact convex) random sets. In other words, $X_\alpha(\omega)$ are Borel-measurable w.r.t. the Borel σ-field generated by the topology associated with a suitable metric on $\mathcal{K}_C(\Re)$, usually the Hausdorff metric d_H:

$$d_H(K,K') = \max\left\{\sup_{k\in K}\inf_{k'\in K'}|k-k'|, \sup_{k'\in K'}\inf_{k\in K}|k-k'|\right\} = \max\left\{\left|\inf K - \inf K'\right|, \left|\sup K - \sup K'\right|\right\} \quad (4)$$

The fuzzy probability distribution function $\tilde{F}(x)$ of $\tilde{X}$ is the set of probability distribution functions of all originals X_j of $\tilde{X}$ with the membership values $\mu(F(x))$. The quantification of fuzziness by fuzzy parameters leads to the description of the fuzzy probability distribution function $\tilde{F}(x)$ of $\tilde{X}$ as a function of the fuzzy bunch parameter $\tilde{s}$.

$$\tilde{F}(x) = F(\tilde{s}, x). \quad (5)$$

For the purposes of numerical evaluation, α-discretization is advantageously applied.

$$F(\tilde{s},x) = \{F_\alpha(x);\ \mu(F_\alpha(x)) \mid F_\alpha(x) = [\underline{F}_\alpha(x),\ \overline{F}_\alpha(x)],\ \mu(F_\alpha(x)) = \alpha,\ \forall\alpha\in[0,1]\}. \quad (6)$$

with $\underline{F}_\alpha(x) = \inf\{F(s,x) \mid s \in s_\alpha\}$ and $\overline{F}_\alpha(x) = \sup\{F(s,x) \mid s \in s_\alpha\}$.

With the aid of α-discretization a fuzzy random function may be formulated as a set of α-level sets of ordinary random functions

$$\tilde{X}(t) = \{X_\alpha(t);\ \mu(X_\alpha(t)) \mid X_\alpha(t) = [\underline{X}_\alpha(t),\ \overline{X}_\alpha(t)],\ \mu(X_\alpha(t)) = \alpha,\ \forall\alpha\in[0,1]\ \}. \quad (7)$$

A *fuzzy random process* $(\tilde{X}_t)_{t\in T}$ is defined as a family of fuzzy random variables $\tilde{X}_t$ over the space T of the time coordinate t.

A *fuzzy time series* $(\tilde{x}_t)_{t\in 1,2,\ldots,N}$ is a realization of a fuzzy random process $(\tilde{X}_t)_{t\in T}$ and consists of a temporally ordered sequence of fuzzy variables $\tilde{x}_t$, each one assigned to each discrete observation time.

4. Propagating Uncertainty in Risk Assessment Models

4.1. Propagating Randomness in Risk Models with no Parameter Uncertainty: One-dimensional Monte Carlo Analysis (1D MCA)

A Monte Carlo analysis that characterizes either uncertainty or variability in each input variable can be described as a one-dimensional Monte Carlo analysis (1D MCA). In its general form, the risk equation can be expressed as a function of multiple risk exposure variables (V_i): $\text{Risk} = f(V_1, \ldots, V_n)$.

Often the input distributions are assumed to be independent. The value of one variable has no relationship to the value of any other variable. In this case, a value for each variable (V_i) is selected at random from a specified PDF and the corresponding risk is calculated. This process is repeated many times (e.g., 10,000). Each iteration of a Monte Carlo simulation should represent a plausible combination of input values. A unique risk estimate is calculated for each set of random values. Repeatedly sampling (V_i) results in a frequency distribution of risk, which can be described by a PDF or a CDF. A sufficient number of iterations should be run to obtain numerical stability in percentiles of the output (e.g., risk distribution). The risk distributions derived from a PRA allow for inferences to be made about the likelihood or probability of risks occurring within a specified range of the input variables.

More complex Monte Carlo simulations can be developed that quantify some dependence between one or more input distributions by using conditional distributions or correlation coefficients.

4.2. Propagating Randomness and Epistemic Uncertainty Simultaneously: Two-dimensional Monte Carlo Analysis

A two-dimensional Monte Carlo Analysis (2D MCA) is a term used to describe a model that simulates both uncertainty and randomness in one or more input variables. Uncertainty in the parameter estimates can be represented in a PRA model as follows. Consider a random input variable whose parameter estimates are affected by uncertainty. Assume normal PDFs can be specified for both uncertain parameters: the mean and the standard deviation. Uncertainty in the mean is described by the normal PDF with parameters (μ_{mean}=5, σ_{mean}=0.5); similarly, uncertainty in the standard deviation is described by the normal PDF with parameters (μ_{SD} =1, σ_{SD} =0.5). A variable described in this way is called a second order random variable.

A two-dimensional Monte Carlo simulation is a nesting of two ordinary Monte Carlo simulations. Typically, the inner simulation represents natural variability of the underlying processes, while the outer simulation represents the analyst's uncertainty about the particular parameters that should be used to specify inputs to the inner

simulation. This structure means that each iterate in the outer simulation entails an entire Monte Carlo simulation, which can lead to a very large computational burden.

4.3. Probability Bounds Analysis Compared to Monte Carlo Simulation

When one or more point estimates defined in the risk model are uncertainty, a Monte Carlo analyst might employ a two-dimensional Monte Carlo simulation that includes an uncertainty (inner) loop for each uncertain point estimate.

In a probability bounds analysis, the same interval of possible values used in the Monte Carlo analyst's uncertainty loop replaces the point estimate; however the semi-analytic nature of the probability bounds analysis results in an exact representation of the stated uncertainty. As the number of times the inner loop is called in the Monte Carlo simulation approaches infinity, the result of the Monte Carlo analysis converges on the probability bounds result.

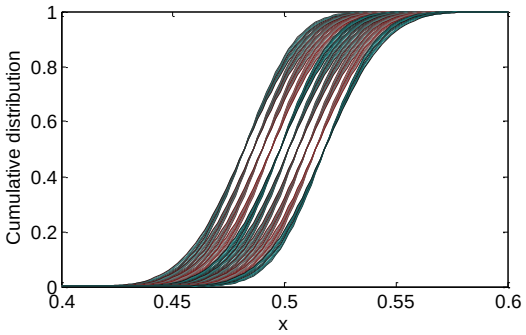

Fig. 2. Monte Carlo vs. Probability Bounds (see also Fig.1): uncertainty regarding the exact value of the parameters of a probability distribution, where μ = [0.482, 0.518] and σ = [0.0182, 0.0218].

4.4. Hybrid Approaches to Propagating Randomness and Fuzziness in Risk Assessment

The idea is to find the output of a model $g\left(X_1,\ldots,X_n,\tilde{\tilde{X}}_1,\ldots,\tilde{\tilde{X}}_m\right)$ that has both random variables $X_1,\ldots,X_n$, given by probabilistic distributions, and fuzzy variables $\tilde{\tilde{X}}_1,\ldots,\tilde{\tilde{X}}_m$, for the inputs. To estimate the output of this generalized model, most researchers attempt to eliminate or transform one type of uncertainty to another before performing a simulation (e.g. possibility to probability transformation). Guyonnet et al. (2003) first proposed a "hybrid approach" with both fuzzy and random types of uncertainty without transforming one type to another. They calculated the Inf and Sup values of the model g considering all the values that are located within the $\alpha-$cuts of the input fuzzy sets and suggested that minimization and maximization algorithm can

be used for finding Inf and Sup values of a general model. However, in their application, the model was a simple monotonic function, and the Inf and Sup values were identified directly without using minimization or maximization algorithms.

A more tractable way to propagating both randomness and fuzziness is based on a fuzzy generalization of the Monte Carlo (FMC) simulation framework, which integrates fuzzy arithmetic method with Monte Carlo simulation to find the output of a model with both fuzzy and probabilistic inputs.

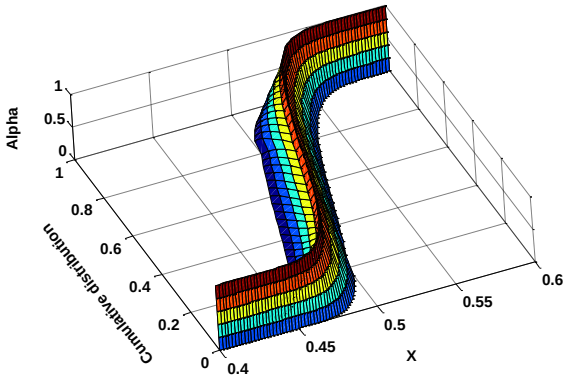

Fig. 3. 3D view of fuzzy CDF resulting from the output of FMC by aggregating α-CDF bounds.

Since in FMC, fuzzy arithmetic (in α-cut form) is performing for each sample set, the output of FMC is represented as a number of fuzzy sets with random variation. This randomness results from random sampling of random input parameters. The fuzzy CDF is used for finding the *fuzzy probability* of not exceeding a given threshold and a *fuzzy quantile* corresponding to a given probability.

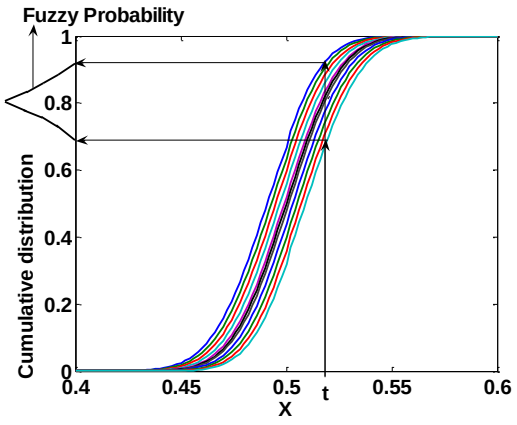

Fig. 4. The fuzzy probability of not exceeding a specific threshold $t \in X$.

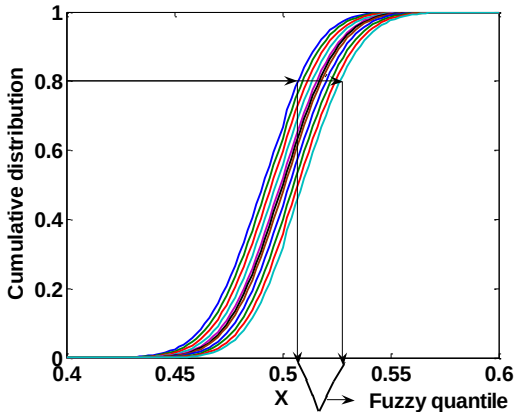

Fig. 5. A fuzzy quantile corresponding to a given probability.

5. Propagating Randomness and Fuzziness Jointly Across Fuzzy Time Series Models

The estimation technique proposed in this paper is based upon a partial decoupling principle, an early form of which I proposed in [10]. It allows decomposing the monolithic fuzzy model into several crisp models, starting from that one corresponding to modal values ($\alpha = 1$) in fuzzy data, and then proceeding in a decremental way for left and right α-level bounds, with α progressively decreasing towards 0. The estimates of modal values are not subject to any constraints, thus being obtained by projecting the 1-level data directly onto the corresponding subspace. However, the estimates for the left and right α-level bounds can only be obtained by projecting the α-level data onto cones, in such a way to obtain least squares estimates without negative spreads. This leads to constrained quadratic programs, conveniently defined.

5.1. Fuzzy Data: Minimum, Average and Maximum Daily Temperatures Registered at a Local Weather Station

The observed sequence consists of the minimum, average and maximum daily temperatures registered at a local weather station. The fuzzy time series is represented in figure 6 and the corresponding empirical fuzzy cumulative distribution function in figure 7.

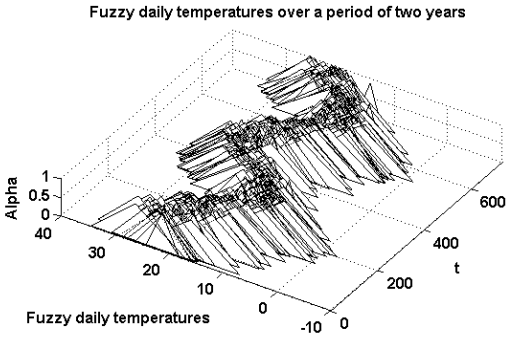

Fig. 6. Fuzzy daily temperatures over a period of two years (730 days).

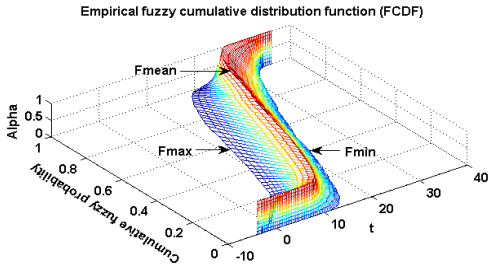

Fig. 7. Empirical Fuzzy Cumulative Distribution Function (FCDF), showing the fuzzy probability of not exceeding a given temperature, or reversely, the fuzzy quantile corresponding to a given probability.

5.2. The Fuzzy Time Series Model for Daily Temperatures, with Fuzzy Trend and Fuzzy Cyclical Component

In what follows, we will exemplify some suitable methods for modelling and forecasting non-stationary fuzzy time series, based on the fuzzy component model, which decomposes the fuzzy time series into a trend component, a cyclical component (or, alternatively, a seasonal component) and a fuzzy residual component:

$$\tilde{Y}(t) = \tilde{T}(t) + \tilde{C}(t) + \tilde{u}(t) \quad . \tag{8}$$

If we restrict to the class of triangular fuzzy numbers, which are a special case of LR fuzzy numbers, we can decompose the fuzzy model into 3 crisp models: one model for the modal (average) values $Y^C(t)$ and two models for the minimum values $Y^L(t)$ and maximum values $Y^R(t)$, respectively.

$$\begin{pmatrix} Y^L(t) \\ Y^C(t) \\ Y^R(t) \end{pmatrix} = \begin{pmatrix} T^L(t) \\ T^C(t) \\ T^R(t) \end{pmatrix} + \begin{pmatrix} C^L(t) \\ C^C(t) \\ C^R(t) \end{pmatrix} + \begin{pmatrix} u^L(t) \\ u^C(t) \\ u^R(t) \end{pmatrix}. \quad \textbf{(9)}$$

Due to the properties of minimum, mean and maximum, the following order relations hold (implying positive spreads for $Y(t)$, $T(t)$, $C(t)$ and $u(t)$):

$$Y^L(t) \le Y^C(t) \le Y^R(t); \qquad T^L(t) \le T^C(t) \le T^R(t); \quad \textbf{(10)}$$

$$C^L(t) \le C^C(t) \le C^R(t); \qquad u^L(t) \le u^C(t) \le u^R(t).$$

The model for the modal values is estimated without any constrains, i.e., by orthogonal projection of the observed values onto the appropriate subspace. The other two models for the minimum and maximum values are estimated subject to some non-negativity restrictions on spreads, corresponding to the time series components: trend ($\hat{T}^C - \hat{T}^L \ge 0$, $\hat{T}^R - \hat{T}^C \ge 0$), cyclical component ($\hat{C}^C - \hat{C}^L \ge 0$, $\hat{C}^R - \hat{C}^C \ge 0$) and residuals ($\hat{u}_1^C - \hat{u}_1^L \ge \xi > 0$, $\hat{u}_1^R - \hat{u}_1^C \ge \xi > 0$). This leads to constrained quadratic programs, i.e., to the projection of the observed values onto some cones.

The regressors for the linear fuzzy trend are defined by the matrix:

$$T = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \dots & \dots \\ 1 & N \end{pmatrix}; \quad N = 730. \quad \textbf{(11)}$$

The simplest way for representing $C(t)$ as a periodic function, with $C(t) = C(t-p)$, is to assume harmonic functions, such as the sine or cosine: $\sin(2\pi t/p)$ and $\cos(2\pi t/p)$, where p (=365 in our case) is called period, its inverse $f = 1/p$ is called frequency, and $\omega = 2\pi f = 2\pi/p$ is the angular frequency. Thus, the regressors for $C(t)$ are combined in the matrix:

$$C = \begin{pmatrix} \cos(1 \cdot 2\pi/365) & \sin(1 \cdot 2\pi/365) \\ \cos(2 \cdot 2\pi/365) & \sin(2 \cdot 2\pi/365) \\ \dots & \dots \\ \cos(\mathrm{N} \cdot 2\pi/365) & \sin(\mathrm{N} \cdot 2\pi/365) \end{pmatrix}. \quad \textbf{(12)}$$

5.3. Estimating the Fuzzy Trend (De-trending)

Step1: First, we estimate the trend corresponding to the average daily temperature, without any constraints, i.e., as a projection of Y^C onto the subspace Im(T) generated by the columns of matrix T:

$$\hat{T}^C = P \cdot Y^C, \quad \text{where } P = T \cdot \left(T' \cdot T\right)^{-1} \cdot T'. \tag{13}$$

Step 2: Second, we estimate the trend corresponding to the minimum daily temperature, as a solution of a constrained quadratic program:

$$\min_{b^L_{trend}} \left\{ \left(b^L_{trend}\right)' T' T\, b^L_{trend} - 2\left(Y^L\right)' T\, b^L_{trend} \right\}. \tag{14}$$

subject to:

$$\begin{cases} \hat{T}^L \le \hat{T}^C \\ \hat{u}^L_1 + \xi \le \hat{u}^C_1 \end{cases} \Leftrightarrow \begin{cases} T \cdot b^L_{trend} \le \hat{T}^C \\ -T \cdot b^L_{trend} \le Y^C - \hat{T}^C - Y^L - \xi \end{cases}. \tag{15}$$

$\hat{T}^L \le \hat{T}^C$ means that the left spread of the fuzzy trend must be non-negative, i.e., $\hat{T}^C - \hat{T}^L \ge 0$, with $\hat{T}^C$ already estimated at step 1.

$\hat{u}^L_1 + \xi \le \hat{u}^C_1$ means that the left spread of the intermediary fuzzy residuals (after de-trending) must be strictly positive, i.e., $\hat{u}^C_1 - \hat{u}^L_1 \ge \xi > 0$. The reason for this is that the intermediary fuzzy residuals will be further decomposed into a cyclical component and final residuals (those obtained after removing both trend and cyclical component). At this step, we recommend for ξ a value between 0.1 and 0.5.

Step 3: Third, we estimate the trend corresponding to the maximum daily temperature, as a solution of a constrained quadratic program:

$$\begin{cases} \hat{T}^L \le \hat{T}^C \\ \hat{u}^L_1 + \xi \le \hat{u}^C_1 \end{cases} \Leftrightarrow \min_{b^R_{trend}} \left\{ \left(b^R_{trend}\right)' T' T\, b^R_{trend} - 2\left(Y^R\right)' T\, b^R_{trend} \right\}. \tag{16}$$

subject to:

$$\begin{cases} -T \cdot b^R_{trend} \le -\hat{T}^C & \Leftrightarrow \ \hat{T}^R \ge \hat{T}^C \\ T \cdot b^R_{trend} \le -Y^C + \hat{T}^C + Y^R - \xi & \Leftrightarrow \ \hat{u}^R_1 - \hat{u}^C_1 \ge \xi > 0 \end{cases}. \tag{17}$$

$\hat{T}^R \ge \hat{T}^C$ means that the right spread of the fuzzy trend must be non-negative, i.e., $\hat{T}^R - \hat{T}^C \ge 0$

$\hat{u}^R_1 - \hat{u}^C_1 \ge \xi > 0$ means that the right spread of the intermediary fuzzy residuals (after de-trending) must be strictly positive.

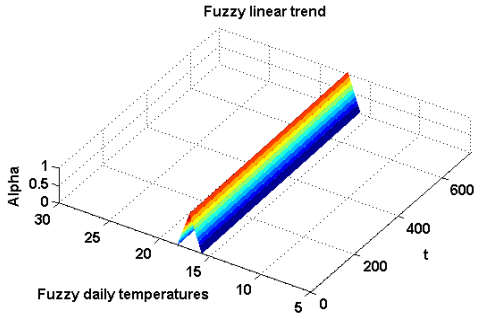

Fig. 8. Fuzzy linear trend.

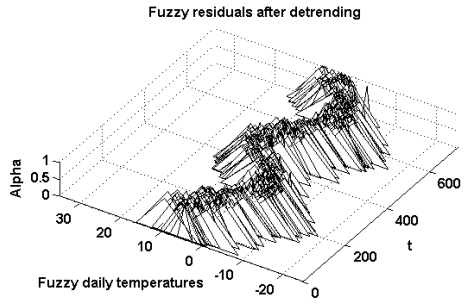

Fig. 9. Intermediary fuzzy residuals after detrending (the residuals are fuzzy sets in a proper sense – with non-negative spreads).

5.4. Estimating the Fuzzy Cyclical Component

The periodogram or sample spectrum shows the variation of the peak points of empirical daily temperature data. The maximum value of the periodogram is about 365 days.

Step1: First, we estimate the cyclical component corresponding to the average daily temperature, without any constraints, i.e., as a projection of $\hat{u}_1^C = Y^C - \hat{T}^C$ (the intermediary residuals after de-trending) on the subspace $\mathrm{Im}(C)$ generated by the columns of matrix C:

$$\hat{C}^C = P \cdot \hat{u}_1^C, \quad \text{where } P = C \cdot (C' \cdot C)^{-1} \cdot C'. \tag{18}$$

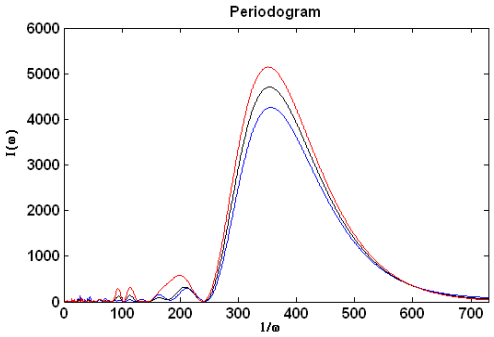

Fig. 10. Periodogram with a maximum corresponding to about 365 days (one year).

Step 2: Second, we estimate the cyclical component corresponding to the minimum daily temperature, as a solution of a constrained quadratic program:

$$\min_{b_{cycle}^{L}}\left\{\left(b_{cycle}^{L}\right)' C' C\, b_{cycle}^{L} - 2\left(\hat{u}_1^{L}\right)' C\, b_{cycle}^{L}\right\}. \tag{19}$$

subject to:

$$\begin{cases}\hat{C}^{L} \le \hat{C}^{C} \\ \hat{u}_2^{L} + \xi \le \hat{u}_2^{C}\end{cases} \Leftrightarrow \begin{cases} C \cdot b_{cycle}^{L} \le \hat{C}^{C} \\ -C \cdot b_{cycle}^{L} \le \hat{u}_1^{C} - \hat{C}^{C} - \hat{u}_1^{L} - \xi\end{cases}. \tag{20}$$

$\hat{C}^{L} \le \hat{C}^{C}$ means that the left spread of the fuzzy cyclical component must be non-negative, i.e., $\hat{C}^{C} - \hat{C}^{L} \ge 0$.

$\hat{u}_2^{C} - \hat{u}_2^{L} \ge \xi > 0$ means that the left spread of the final fuzzy residuals (after removing both trend and cyclical component) must be strictly positive. The reason for this is that the final fuzzy residuals will be further decomposed as a multivariate auto-regressive process (VAR). At this step, we recommend for ξ a value of about 0.1.

Step 3: Third, we estimate the cyclical component corresponding to the maximum daily temperature, as a solution of a constrained quadratic program:

$$\min_{b_{cycle}^{R}}\left\{\left(b_{cycle}^{R}\right)' C' C\, b_{cycle}^{R} - 2\left(\hat{u}_1^{R}\right)' C\, b_{cycle}^{R}\right\}. \tag{21}$$

subject to:

$$\begin{cases}\hat{C}^{R} \ge \hat{C}^{C} \\ \hat{u}_2^{R} - \hat{u}_2^{C} \ge \xi > 0\end{cases} \Leftrightarrow \begin{cases} -C \cdot b_{cycle}^{R} \le -\hat{C}^{C} \\ C \cdot b_{cycle}^{R} \le -\hat{u}_1^{C} + \hat{C}^{C} + \hat{u}_1^{R} - \xi\end{cases}. \tag{22}$$

$\hat{C}^R \geq \hat{C}^C$ means that the right spread of the fuzzy cyclical component must be non-negative, i.e., $\hat{C}^R - \hat{C}^C \geq 0$

$\hat{u}_2^R - \hat{u}_2^C \geq \xi > 0$ means that the right spread of the final fuzzy residuals (after removing both trend and cyclical component) must be strictly positive.

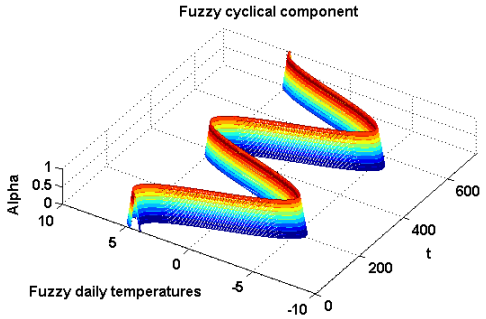

Fig. 11. Fuzzy cyclical component.

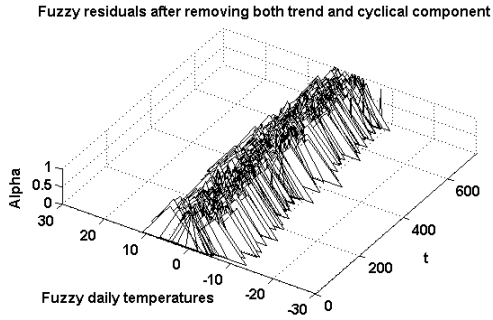

Fig. 12. Fuzzy residuals after removing both trend and cyclical component (the residuals are fuzzy sets in a proper sense – with non-negative spreads).

5.5. Modeling and Forecasting the Fuzzy Residuals as a VAR(4) Process, after Removing both Trend and Cyclical Component

The fuzzy residuals obtained after removing both trend and cyclical component can now be modeled as a multivariate auto-regressive process. A VAR(4) model has been chosen (among some other candidate models) based upon likelihood ratio tests and Akaike Information Criterion. This allows forecasting or simulating the residuals, starting from a sequence of the latest 10% observed historical temperatures.

Afterwards, based on the inversion property of the generalized Hukuhara difference, we can forecast the series of fuzzy daily temperatures by recomposing them from its components: trend component + cyclical component + the simulated residuals.

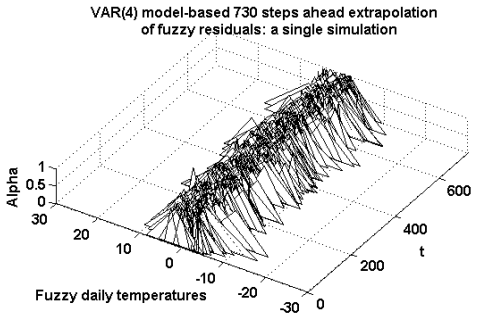

Fig. 13. VAR(4) model-based 730 steps ahead extrapolation of fuzzy residuals: a single simulation (the residuals are fuzzy sets in a proper sense – with non-negative spreads).

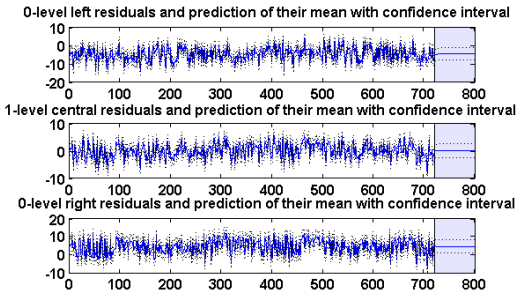

Fig. 14. Fuzzy residuals (left, central and right) and prediction of their mean with confidence interval ($\pm\sigma$).

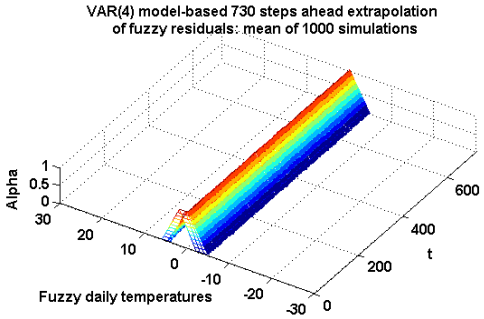

Fig. 15. VAR(4) model-based 730 steps ahead extrapolation of fuzzy residuals: mean of 1000 simulations.

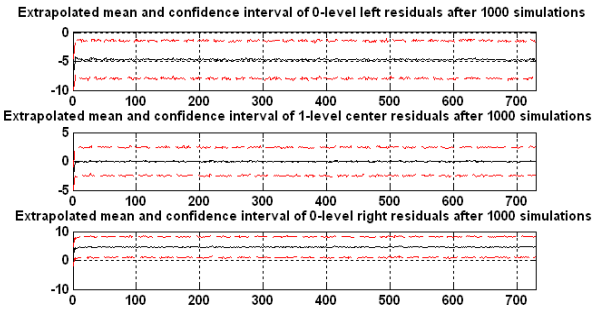

Fig. 16. Extrapolated mean and confidence interval of 0-level left / 1-level center / 0-level right fuzzy residuals, after 1000 simulations.

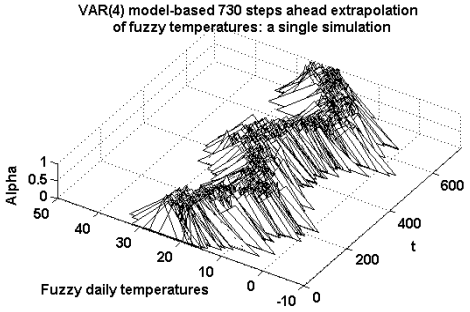

Fig. 17. VAR(4) model-based 730 steps ahead extrapolation of fuzzy temperatures, additively recomposed from trend, cyclical component and simulated fuzzy residuals.

5.6. Fuzzy Time Series Wavelet Decomposition and Nonlinear Model Fitting with Wavelet Networks

Another alternative to removing disturbances from a time series is de-noising data by wavelet decomposition.

The Discrete Wavelet Transform (DWT) [17] uses scaled and shifted versions of a mother wavelet function, usually with compact support, to form either an orthonormal basis (Haar wavelet, Daubechies) or a bi-orthonormal basis (Symlets, Coiflets). Wavelets allow cutting up data into different frequency components (called approximations and details), and then studying each component with a resolution matched to its scale. They can help de-noise inherently noisy data through wavelet shrinkage and thresholding methods, developed by David Donoho [5]. The idea is to set to zero all wavelet coefficients corresponding to details in the data set that are less than a particular threshold. These coefficients are used in an inverse wavelet transformation to reconstruct the data set. An important advantage is that the de-noising is carried out without smoothing out the sharp structures and thus can help to increase the predictive performance.

We start with the de-trended fuzzy time series shown in figure 13. A level 5 decomposition with Sym8 wavelets and a fixed form soft thresholding is first performed (see figures 18 and 19).

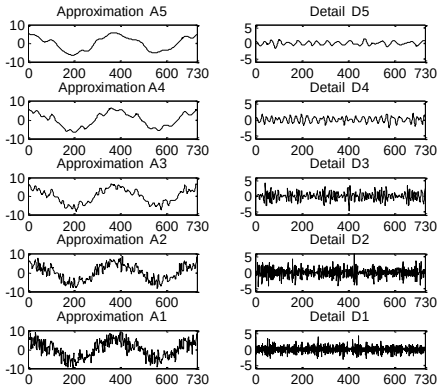

Fig. 18. A level 5 decomposition of the average temperature time series using Sym8 wavelets: approximations and details. The successive approximations appear less and less noisy; however, they also lose progressively more high-frequency information.

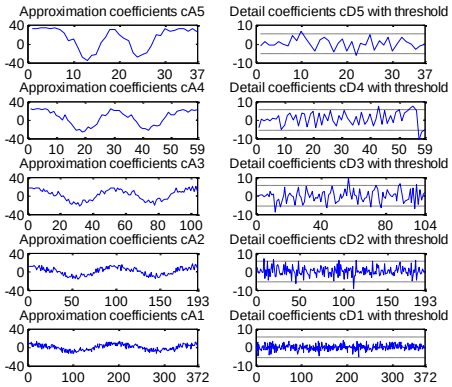

Fig. 19. Approximation coefficients and detail coefficients with a global threshold.

The initially de-trended average, minimum and maximum temperature time series are now de-noised in turn (see figures 20, 21 and 22).

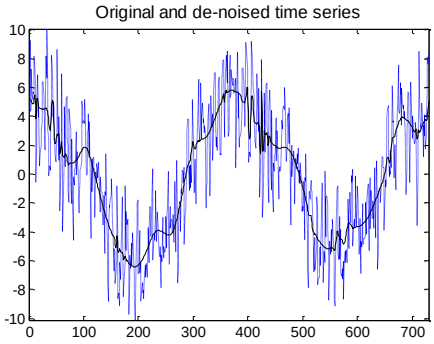

Fig. 20. De-trended vs. de-trended & de-noised average temperature time series.

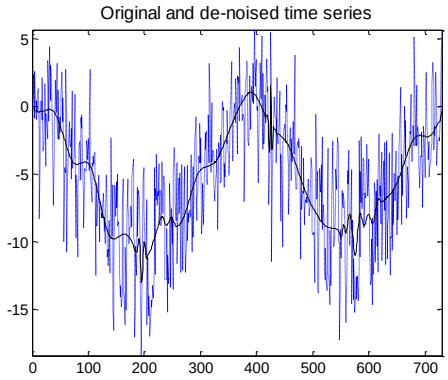

Fig. 21. De-trended vs. de-trended & de-noised minimum temperature time series.

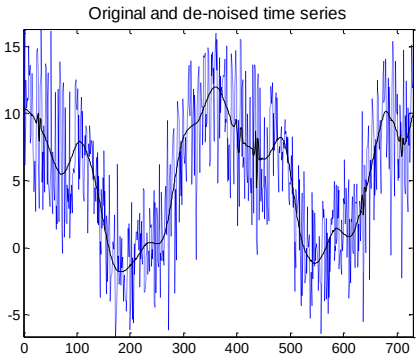

Fig. 22. De-trended vs. de-trended & de-noised maximum temperature time series.

However, the representations are not smooth enough, because of some weather turbulences that occur in certain time intervals. In order to produce smoothed

representations, the time series obtained after de-noising can be further fitted to some nonlinear approximation functions by using wavelet networks to learn them.

Wavelet networks attempt to combine the properties of the Wavelet decomposition previously described, along with the learning capabilities of feedforward neural networks. They employ wavelets instead of sigmoidal activations functions, are trained with a backpropagation-like algorithm and behave as universal approximators, being capable of estimating almost any computable function on a compact set arbitrarily closely. Their rigorous mathematical foundations and better localization and approximation properties allow hierarchical and multi-resolution learning as well as transparent design of the network. Wavelet networks can be easily generalized to the case of multidimensional nonlinear function approximation in order to approximate functions in $L_2(\Re^n)$ and their representation can be extended with radial wavelets that are better suited for approximation problems of large dimensions. This results in the following network structure:

$$g(x) = \sum_{i=1}^{N} w_i \psi(diag(d_i)(x - t_i)) + c' x + b \quad \textbf{(23)}$$

where Ψ is a radial wavelet function, $d_i \in \Re^n$ are dilatation parameters, $t_i \in \Re^n$ are translation parameters, $w_i \in \Re$ are linear weights, N is the number of wavelets, $c \in \Re^n$ is the additional direct linear combination parameters (direct connection parameters), and $b \in \Re$ is the bias parameter.

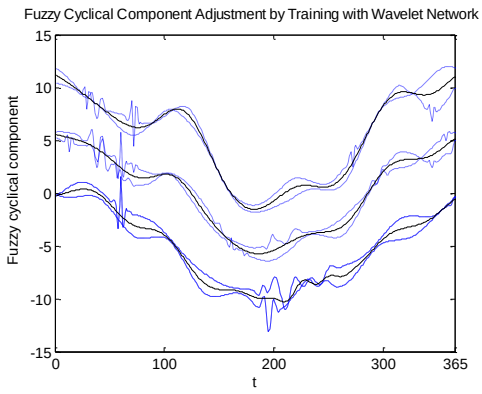

Fig. 23. 2D representation of the adjusted fuzzy cyclical component after training with wavelet network.

The input space is the set $T = \{1, 2, \ldots, 365\}$ of discrete time values. We train the wavelet network three times, for each set of minimum, average and maximum daily temperatures in turn. Each time, the output space is the set of de-trended and de-noised minimum, average and maximum daily temperatures, i.e., $S_{\min}$, S_{avg} and $S_{\max}$, respectively. 10 wavelets in the hidden layer and 10 iterations (epochs) are used for training. Finally, the adjusted fuzzy cyclical component is obtained by mapping T

onto $S_{\min}$, S_{avg} and $S_{\max}$, i.e. $\hat{S}_{\min}(t):T\to S_{\min}$, $\hat{S}_{avg}(t):T\to S_{avg}$ and $\hat{S}_{\max}(t):T\to S_{\max}$, respectively. 2D and 3D representations of the adjusted fuzzy cyclical component after training with wavelet network, starting from the de-trended and de-noised data, are shown in figures 23 and 24.

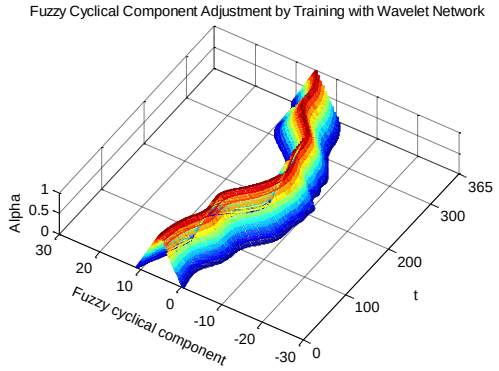

Fig. 24. 3D representation of the adjusted fuzzy cyclical component after training with wavelet network.

Finally, the fuzzy trend + fuzzy cyclical component can be re-compounded (figure 25).

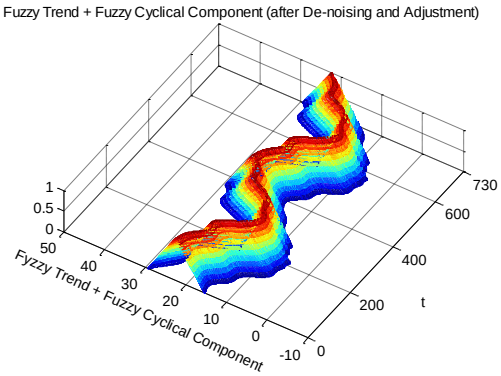

Fig. 25. Fuzzy trend + fuzzy cyclical component (after de-noising and adjustment).

6. Conclusion

Two kinds of uncertainty were contrasted in this paper (ontological, vs. epistemic uncertainty) and several techniques were addressed in order to capture and propagate both of them jointly across a specific model. The first application area is concerned

with extending the classical *probabilistic risk assessment* modeling framework in such a way that allows probability distribution parameters to be imprecisely defined. The presence of imprecision in a risk model adds another dimension to that of randomness and requires more comprehensive approaches to capture and represent the range within which the risk distribution might vary. Analytic methods for propagating the uncertainty (such as Probability bounds analysis, Fuzzy α-levels analysis) as well as stochastic simulation techniques (such as Monte Carlo Analysis), can be combined or integrated, in an attempt to find the output of a model that has both random and fuzzy variables for the inputs.

The second application area is addressed when attempting to capture the inherent fuzzy and random nature of some stochastic processes, expressed in terms of fuzzy time series. Suitable new methods for fuzzy time series estimation and prediction, using both the estimation theory and Computational Intelligence techniques were proposed.

We combined a generalized Hukuhara difference, which allows the fuzzy estimation problem to be handled in some L_2-type metric space, with a partial decoupling principle, which allows the monolithic fuzzy model to be broken in several more tractable crisp estimation sub-problems. This approach was proved to provide an efficient solution to the problem of non-invertibility of the standard Minkovsky addition and multiplication in a fuzzy feature space, while enabling to obtain fuzzy estimations in a proper sense (i.e., with non-negative spreads).

Alternatively, wavelet decomposition, a Computational Intelligence based technique, has been also used to de-noising fuzzy time series. Finally, starting from the de-trended and de-noised time series, wavelet networks have been employed as universal approximators to adjust the fuzzy cyclical component and thus to produce smoothed representations of the fuzzy time series components

References

1. Baudrit C., Guyonnet D., Dubois D.: Postprocessing the Hybrid Method for Addressing Uncertainty in Risk Assessments. Journal of Environmental Engineering 131(12), 1750-1754. (2005)
2. Diamond P.: Fuzzy least squares. Inform. Sci. 46, 141-157. (1988)
3. Diamond P., Kloeden P.: Metric Spaces of Fuzzy Sets: Theory and Applications. World Scientific, Singapore, (1994)
4. Diamond P., Körner R.: Extended fuzzy linear models and least squares estimates. Comput. Math. Appl. 33, 15-32. (1997)
5. Donoho D.: Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data. In Daubechies, I. (eds). Different Perspectives on Wavelets Amer. Math. Soc., Providence, R.I., 173-205. (1993)
6. EPA's guidance: Risk Assessment Guidance for Superfund (RAGS). Volume III - Part A: Process for Conducting Probabilistic Risk Assessment, (2001). [Online]. Available: www.epa.gov/oswer/riskassessment/rags3adt/index.htm
7. Ferson S., Tucker W.T.: Probability Bounds Analysis in Environmental Risk Assessments (Technical report). Applied Biomathematics, Setauket, New York (2003). [Online]. Available: www.ramas.com/pbawhite.pdf.

8. Ferson S., Ginzburg L., Kreinovich V., Nguyen H.T., Starks S.A.: Uncertainty in risk analysis: towards a general second-order approach combining interval, probabilistic, and fuzzy techniques. In Proceedings of the 2002 IEEE International Conference on Fuzzy Systems, 1342-1347. (2002)
9. Georgescu V.: Estimation of fuzzy regression models using quadratic programming. Economic Computation and Economic Cybernetics Studies and Research, 4, 105-124. (1997)
10. Georgescu V.: New estimation methods in fuzzy regression analysis, based on projection theorem and decoupling principle. Fuzzy Economic Review, III/1, 21-38. (1998)
11. Georgescu V.: On the Foundations of Granular Computing Paradigm, Fuzzy Economic Review, VIII/2, 73-105. (2003)
12. Georgescu V.: A Generalization of Symbolic Data Analysis Allowing the Processing of Fuzzy Granules, Lecture Notes in Artificial Intelligence, 3131, Springer, 215-226. (2004)
13. Georgescu V.: Fuzzy Time Series Estimation and Prediction: Criticism, Suitable New Methods and Experimental Evidence. Studies in Informatics and Control, Volume 19, Issue 3. (2010)
14. Guyonnet D., Bourgigne B., Dubois D., Fargier H., Côme B., Chilès J.: Hybrid approach for addressing uncertainty in risk assessments. Journal of environmental engineering 129(1), 68-78. (2003)
15. Hukuhara M.: Integration des applications measurables dont la valeur est un compact convexe, Funkcialaj Ekvacioj 10, 205–223. (1967)
16. Körner R.., Näther W.: Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates, Inform. Sci. 109, 95–118. (1998)
17. Mallat S. G., Peyré G.: A Wavelet Tour of Signal Processing: The Sparse Way, Academic Press, 3rd Edition (2009)
18. Möller B., Beer M.: Fuzzy Randomness – Uncertainty Computational Mechanics, Springer, Berlin and New York (2004)
19. Möller B., Reuter U.: Uncertainty Forecasting in Engineering, Springer-Verlag, Berlin Heidelberg. (2007)
20. Puri M.L., Ralescu D.A., Fuzzy random variables, J. Math. Anal. Appl. 114, 409–422. (1986)
21. Stefanini L.: A generalization of Hukuhara difference and division for interval and fuzzy arithmetic, Fuzzy Sets and Systems, Vol. 161, Issue 11, 1564-1584. (2010)
22. Terán P.: Probabilistic foundations for measurement modeling with fuzzy random variables, Fuzzy Sets and Systems 158(9), 973-986. (2007)

Vasile Georgescu is currently a professor of econometrics at University of Craiova, Romania and director of the Research Center for Computational Intelligence in Business and Economics. He has published more than 150 papers in fields such as computational intelligence, computational statistics, econometrics, cybernetics, pattern recognition and more than 20 books.

Received: December 15, 2012; Accepted: December 09, 2013