

# Fuzzy Time Series Estimation and Prediction: Criticism, Suitable New Methods and Experimental Evidence

Vasile Georgescu

University of Craiova, 13 A.I.Cuza, Craiova, 200585, Romania

vgeo@central.ucv.ro

**Abstract:** This paper is devoted to exploring suitable methods for modelling, estimating and forecasting fuzzy time series, when facing the problem of non-invertibility of the standard Minkovsky addition and multiplication in a fuzzy framework. Some generalized versions of Hukuhara difference, which allow the fuzzy estimation problem to be handled in some  $L_2$ -type metric space, are first examined from a critical viewpoint. This leads us to propose a new estimation procedure, where the monolithic fuzzy model is broken in several more tractable crisp estimation sub-problems, based upon a partial decoupling principle. Our aim is to produce fuzzy estimations with non-negative spreads, capable not only 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. Computational Intelligence techniques such as wavelet decomposition and de-noising or nonlinear model fitting with wavelet networks are also addressed. Finally, the proposed methods are exemplified for a fuzzy time series with fuzzy daily temperatures (minimum, average and maximum values).

**Keywords:** Fuzzy time series estimation and prediction, Generalized Hukuhara difference, Projection cones vs. projection subspaces, Wavelet decomposition and de-noising, Nonlinear fitting with wavelet networks.

## 1. Introduction

The study of fuzzy time series has attracted great interest and is expected to expand rapidly. Fuzzy time series have an inherent fuzzy and random nature.

We 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  $\tilde{x}(\omega) = (\tilde{x}_1, \dots, \tilde{x}_n) \subseteq X$  of each elementary event  $\omega \in \Omega$ . We will restrict attention to the class  $F_C(\mathfrak{R})$  of normal convex fuzzy sets on  $\mathfrak{R}$ , whose  $\alpha$ -level sets are in the class  $K_C(\mathfrak{R})$  of nonempty compact real intervals.

A fuzzy random variable  $\tilde{X}$  is the fuzzy result of the uncertain mapping  $\tilde{X} : \Omega \rightarrow F_C(\mathfrak{R})$ , such that for each  $\alpha \in [0, 1]$  and  $\omega \in \Omega$ , the  $\alpha$ -level intervals  $X_\alpha(\omega) = [\inf(X(\omega))_\alpha, \sup(X(\omega))_\alpha]$ , generated by the mapping  $X_\alpha : \Omega \rightarrow K_C(\mathfrak{R})$ , are random sets. In other words,  $X_\alpha(\omega)$  are Borel-measurable w.r.t. the Borel  $\sigma$ -field generated by the topology associated with a suitable metric on  $K_C(\mathfrak{R})$ , usually the Hausdorff metric  $d_H$ .

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  $(\tilde{x}_t)_{t \in 1, 2, \dots, N}$  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.

For extending both the classical estimation theory and some computational intelligence techniques like wavelet analysis and wavelet networks to time series models with fuzzy data, appropriate assumptions should be stated and suitable methods should be developed.

Square-integrable random variables are assumed, defined on a Hilbert space equipped with a suitable  $L_2$ -metric that allows the projection theorem to be still valid. However, it 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. Using Hukuhara difference instead of fuzzy subtraction has been proposed to overcome the problem. Unfortunately, it does not always exist, and even if it exists,

some distortions may still appear when applying least square estimation.

A criticism of the existing fuzzy estimation methods in the literature is first addressed and suitable new methods are then proposed, based upon a partial decoupling principle. 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  $\alpha$ -level bounds, with  $\alpha$  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  $\alpha$ -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.

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. Parameter Estimation in Models with Fuzzy Data: A Critical View

The problem of parameter estimation in models with fuzzy data has been previously treated in literature from different points of view. Inversion of addition and multiplication has been typically addressed as a problem of interest. The reason is that, in fuzzy arithmetic, the standard Minkovsky addition and multiplication are not invertible.

### Using generalized Hukuhara difference instead of fuzzy subtraction: rationales and limitations

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 introduced as a set  $C$  for which  $C = A \sigma_H B \Leftrightarrow A = B + C$ .

This is important in regression analysis, where the estimates are obtained from the condition of minimizing the sum of squares residuals, expressed as a difference between the response of a system and its model based estimation.

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

Several generalizations of Hukuhara difference have been proposed. The first and most notable in the context of fuzzy linear regression was that of Diamond [1], or Diamond and Körner [3], as the least squares solution of equation  $A + X = B$ , i.e.,  $C = B \sigma_H A$ , if and only if  $d(A + C, B) = \inf_{X \in Y} d(A + X, B)$  in some  $L_2$ -type metric space  $(Y, d)$ . When the usual Hukuhara difference  $B \sigma_H A$  exists, it coincides with the least squares solution defined above.

The extension of Hukuhara difference to an  $L_2$ -approximant is discussed in what follows.

Let  $\mathcal{F}_C(\mathfrak{R}^p)$  denote the class of normal and convex upper semicontinuous fuzzy sets of  $\mathfrak{R}^p$  with bounded closure of the support, that is,

$$\mathcal{F}_C(\mathfrak{R}^p) = \left\{ U : \mathfrak{R}^p \rightarrow [0, 1] \mid U_\alpha \in \mathcal{K}_C(\mathfrak{R}^p) \text{ for all } \alpha \in [0, 1] \right\}$$

where  $U_\alpha$  is the  $\alpha$ -level set of  $U$  if  $\alpha \in (0, 1]$ , and  $U_0 = cl(\{x \in \mathfrak{R}^p \mid U(x) > 0\})$ .

Due to the lack of asymmetric element w.r.t. the addition,  $(\mathcal{F}_C(\mathfrak{R}^p), +, \cdot)$  is not a linear space. Let  $\mathbf{S}^{p-1}$  and  $\langle \cdot, \cdot \rangle$  be the unit sphere and the inner product in  $\mathfrak{R}^p$ . The

space  $F_C(\mathfrak{R}^p)$  can be embedded onto a cone of the class of the square integrable functions  $L(\mathbf{S}^{p-1} \times [0, 1])$  by means of the mapping  $s : F_C(\mathfrak{R}^p) \rightarrow L(\mathbf{S}^{p-1} \times [0, 1])$  which associates each  $U \in F_C(\mathfrak{R}^p)$  with its support functions  $s_U$  (where  $s_U : \mathbf{S}^{p-1} \times [0, 1] \rightarrow \mathfrak{R}$  is defined so that  $s_U(u, \alpha) = \sup_{w \in U_\alpha} \langle u, w \rangle$  for any  $u \in \mathbf{S}^{p-1}$  and  $\alpha \in [0, 1]$ ). The support function is semilinear, in the sense that  $s_{U+V} = s_U + s_V$  and  $s_{\lambda U} = \lambda \cdot s_U$  if  $\lambda \geq 0$ . Furthermore, if  $U \sigma_H V$  exists, then  $s_{U \sigma_H V} = s_U - s_V$  for all  $U, V \in K_C(\mathfrak{R}^p)$ .

We will restrict attention to LR-fuzzy intervals, defined as follows:

$$\begin{aligned} (m, c, \ell, r)_{LR}^\alpha &= [m_L(\alpha), m_R(\alpha)] = \\ &= [m_L - \ell \cdot L^{-1}(\alpha), m_R + r \cdot R^{-1}(\alpha)] = \\ &= [m - c - \Delta_L^\alpha, m + c + \Delta_R^\alpha,] \end{aligned}$$

where  $m = (m_L + m_R)/2$  is the modal point,  $m_L = m - c$ ,  $m_R = m + c$ ,  $c = (m_R - m_L)/2$  is the half-length of the flat region,  $\Delta_L^\alpha = \ell \cdot L^{-1}(\alpha)$ ,  $\Delta_R^\alpha = r \cdot R^{-1}(\alpha)$ , for all  $\alpha \in [0, 1]$ . Note that for LR-fuzzy numbers  $c = 0$ .

Taking into account that the unit sphere in  $\mathfrak{R}^1$  consists of only two points  $u \in \{-1, 1\}$ , the support function of  $M_{LR} = (m, c, \ell, r)$  is given by

$$\begin{aligned} s_M : \{-1, 1\} \times [0, 1] \rightarrow \mathfrak{R}, \quad s_M(u, \alpha) = \\ = \begin{cases} u \cdot m_L(\alpha) = -m + c + \ell \cdot L^{-1}(\alpha), & u = -1 \\ u \cdot m_R(\alpha) = m + c + r \cdot R^{-1}(\alpha), & u = 1 \end{cases} \end{aligned}$$

The support function is the way of defining an  $L_2$ -metric on the space of normal compact convex fuzzy sets  $F_C(\mathfrak{R}^p)$  using the  $L_2$ -metric on the Hilbert space of square-integrable functions  $L_2(\mathbf{S}^{p-1} \times [0, 1])$ . This space is equipped with the inner product

$$\begin{aligned} \langle M, N \rangle_\lambda &= \\ &= p \cdot \int_{[0,1]} \int_{\mathbf{S}^{p-1}} s_M(u, \alpha) s_N(u, \alpha) \mu(du) \lambda(d\alpha) \end{aligned}$$

where  $\mu$  is normalized Lebesgue measure on  $\mathbf{S}^{p-1}$  (i.e.  $\mu(\mathbf{S}^{p-1}) = 1$ ) and  $\lambda$  is normalized Lebesgue measure on  $[0, 1]$ .

The corresponding  $L_2$ -metric results of the form

$$\begin{aligned} \delta_2^{(\lambda)}(M, N) &= \|s_M - s_N\|_2 = \\ &= \left( p \cdot \int_{[0,1]} \int_{\mathbf{S}^{p-1}} W^2(u, \alpha) \mu(du) \lambda(d\alpha) \right)^{1/2} \end{aligned}$$

where  $W^2(u, \alpha) = |s_M(u, \alpha) - s_N(u, \alpha)|^2$ .

As mentioned above, the models  $Y = A + (-)X$  and  $Y + bX = A$  differ from each other. More importantly,  $B + (-)A$  is not compatible with the difference in the function space  $L_2(\mathbf{S}^{p-1} \times [0, 1])$ , that is, it does not have the support function  $s_B - s_A$ . But the Hukuhara difference defined as a solution for  $X$  in the equation  $A + X = B$ , if it exists, does coincide with the difference in  $L_2(\mathbf{S}^{p-1} \times [0, 1])$ . This property justifies the application of this difference instead of the fuzzy number  $B + (-)A$ .

For fuzzy numbers  $A, B$ , if the Hukuhara difference  $B \sigma_H A$  exists, it is given by

$$\begin{aligned} [B \sigma_H A]^\alpha &= \{ a \in \mathfrak{R}^d \mid [A]^\alpha + \{a\} \subseteq [B]^\alpha \}, \\ \alpha &\in [0, 1] \end{aligned}$$

In particular, the Hukuhara difference  $B \sigma_H A$  of two symmetric triangular fuzzy numbers  $A = (m_A, \ell_A)_{LL}^\alpha$  and  $B = (m_B, \ell_B)_{LL}^\alpha$  is well-defined if  $\ell_B \geq \ell_A$  and  $B \sigma_H A = (m_B - m_A, \ell_B - \ell_A)_{LL}^\alpha$ . Since an exact solution of the equation  $A + X = B$  is then impossible, one can find  $X$  such that  $A + X$  is the  $L_2$ -approximant to  $B$ . Write  $X = B \sigma_H A$ . For the space of symmetric triangular fuzzy numbers,  $B \sigma_H A = (m_B - m_A, 0)_{LL}^\alpha$  if  $\ell_B \geq \ell_A$ , which is actually a crisp number.

The Hukuhara difference can now be extended to the  $L_2$ -approximant  $B \sigma_H A$  by

$$\begin{aligned} B \sigma_H A = C \quad \text{iff} \quad \delta_2^{(\lambda)}(A + C, B) = \\ = \inf_X \delta_2^{(\lambda)}(A + X, B) \end{aligned}$$

Another generalized Hukuhara difference has been proposed recently in [18], as follows:

$$B \sigma_{\text{gH}} A = C \Leftrightarrow \begin{cases} (i) & B = A + C \\ (ii) & A = B + (-1)C \end{cases}$$

The definition given above allows us to express the generalized Hukuhara difference of compact convex fuzzy sets  $A, B \in F_C(\mathfrak{R}^p)$  by the use of the support functions. Consider  $A, B, C \in F_C(\mathfrak{R}^p)$  with  $C = B \sigma_{\text{gH}} A$ . Let  $s_A, s_B, s_C$  and  $s_{(-1)C}$  be the support functions of  $A, B, C$ , and  $(-1)C$ , respectively. In case (i), we have  $s_B = s_A + s_C$  and in case (ii), we have  $s_A = s_B + s_{(-1)C}$ . So,

$$s_C(u, \alpha) = \begin{cases} s_B(u, \alpha) - s_A(u, \alpha) & \text{in case (i)} \\ s_A(-u, \alpha) - s_B(-u, \alpha) & \text{in case (ii)} \end{cases}$$

i.e.

$$s_C(u, \alpha) = \begin{cases} s_B(u, \alpha) - s_A(u, \alpha) & \text{in case (i)} \\ s_{(-1)A}(u, \alpha) - s_{(-1)B}(u, \alpha) & \text{in case (ii)} \end{cases}$$

where for a fuzzy quantity  $C$ , the support functions are considered for each  $\alpha$ -cut and defined to characterize the compact  $\alpha$ -cuts  $C^\alpha$ :

$s_C : \mathbf{S}^{p-1} \times [0, 1] \rightarrow \mathfrak{R}$  defined so that

$$s_C(u, \alpha) = \sup_{v \in C^\alpha} \langle u, v \rangle, \quad \forall u \in \mathbf{S}^{p-1} \quad \text{and} \quad \alpha \in [0, 1].$$

Despite the obvious advantages of using the generalized Hukuhara difference instead of fuzzy subtraction, its application for the parametric estimation of a monolithic fuzzy model is still rigid and practically unsuitable in some real-world cases, as it will be motivated next. This will lead us to propose a new estimation procedure based upon a partial decoupling principle.

### Rationales for avoiding the use of fuzzy multiplication

In order to estimate the parameter vector  $b$  in the case of a classical regression model  $y = Xb + e$ , we have to take an orthogonal projection of  $y$  onto the subspace  $\text{Im}(X)$

generated by the columns of  $X$ , i.e.  $\hat{y} = X \hat{b} = Py$ , with  $P = X(X'X)^{-1}X'$ .

When we deal with regression models defined with respect to fuzzy input-output data, the projection subspaces are somewhat ambiguous. The major difficulty is to obtain an explicit form for the product  $B \otimes X$ , where  $B$  is a fuzzy coefficient and  $X$  is a fuzzy-valued explanatory variable. The extension principle of Zadeh is the basic result from which all fuzzy operations are derived. As long as  $B$  is unknown, we cannot directly obtain an ex-ante specification of the membership function of the product  $B \otimes X$ . According to Nguyen's theorem, such a specification depends on the factor signs. Thus:

$$X = [x_L, x_R] > 0, \hat{B} = [\hat{b}_L, \hat{b}_R] > 0 \Rightarrow [\hat{Y}_L, \hat{Y}_R] = X \otimes \hat{B} = [x_L \cdot \hat{b}_L, x_R \cdot \hat{b}_R]$$

$$X = [x_L, x_R] < 0, \hat{B} = [\hat{b}_L, \hat{b}_R] > 0 \Rightarrow [\hat{Y}_L, \hat{Y}_R] = X \otimes \hat{B} = [x_L \cdot \hat{b}_R, x_R \cdot \hat{b}_L]$$

$$X = [x_L, x_R] > 0, \hat{B} = [\hat{b}_L, \hat{b}_R] < 0 \Rightarrow [\hat{Y}_L, \hat{Y}_R] = X \otimes \hat{B} = [x_R \cdot \hat{b}_L, x_L \cdot \hat{b}_R]$$

$$X = [x_L, x_R] < 0, \hat{B} = [\hat{b}_L, \hat{b}_R] < 0 \Rightarrow [\hat{Y}_L, \hat{Y}_R] = X \otimes \hat{B} = [x_R \cdot \hat{b}_R, x_L \cdot \hat{b}_L]$$

Some other cases could be taken into account if we consider opposite signs for the left and right interval bounds. The general definition is:

$$[\hat{Y}_L, \hat{Y}_R] = X \otimes \hat{B} = \begin{bmatrix} \min \{ x_L \cdot \hat{b}_L, x_L \cdot \hat{b}_R, x_R \cdot \hat{b}_L, x_R \cdot \hat{b}_R \}, \\ \max \{ x_L \cdot \hat{b}_L, x_L \cdot \hat{b}_R, x_R \cdot \hat{b}_L, x_R \cdot \hat{b}_R \} \end{bmatrix}$$

As opposite to the classical linear regression, where the parameters are obtained by projection onto a linear subspace, in the fuzzy linear model, however, there is no full linear structure and we also have to deal with the problem of negative spreads. Such a problem may occur because the multiplication of an LR-fuzzy number by a negative scalar reverses the spreads. This leads to imposing some constraints on the regression coefficients and results in a constrained optimization problem. Instead of a subspace of projection, we have cones of projections, according to the rules derived from the Nguyen's theorem.

With a positive explanatory variable  $X$  and a fuzzy slope-coefficient  $\hat{B} = [\hat{b}_L, \hat{b}_R] > 0$ , the projection cone for  $Y_L$  will be defined with respect to  $x_L$  and the projection cone for  $Y_R$  will be defined with respect to  $x_R$ .

By contrary, if  $\hat{B} = [\hat{b}_L, \hat{b}_R] < 0$ , the projection cone for  $Y_L$  will be defined with respect to  $x_R$  and the projection cone for  $Y_R$  will be defined with respect to  $x_L$ .

Now, let us consider one of the simplest models of fuzzy input-output data, with a non-fuzzy slope-coefficient  $Y = A + b \cdot X$  where  $A = (m_A, \ell_A, r_A)_{LR}$ ,  $b \in \mathfrak{R}$ ,  $X = (m_x, \ell_x, r_x)_{LR}$ . Figures 3 and 4 show how to choose the two projection cones (instead of projection subspaces):

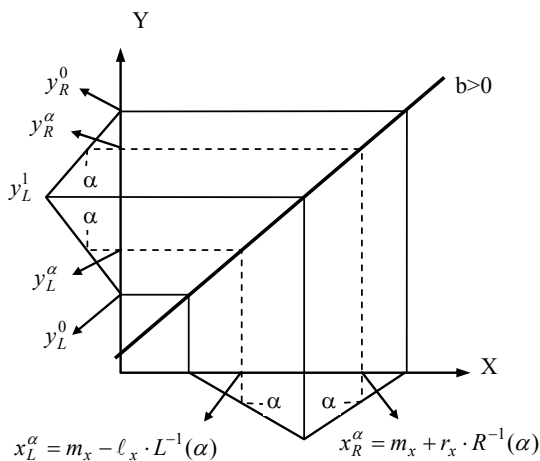
For  $b > 0$ :  $\hat{y} = (\hat{y}^L, \hat{y}^R) \in C_+$ , where

$$C_+ = \left\{ (m_x - \ell_x \cdot L^{-1}(\alpha), m_x + r_x \cdot R^{-1}(\alpha)) \mid m_x \in \mathfrak{R}; \ell_x, r_x \geq 0 \right\}$$

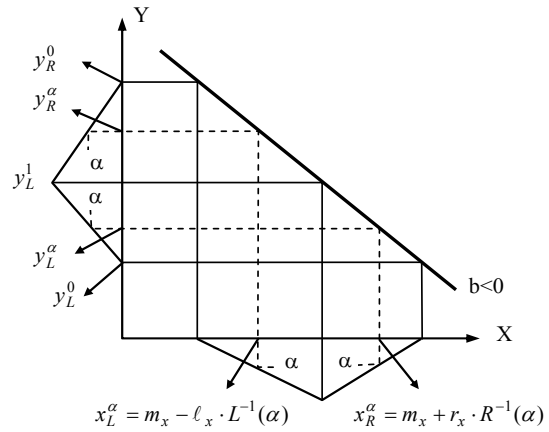
For  $b < 0$ :  $\hat{y} = (\hat{y}^L, \hat{y}^R) \in C_-$ , where

$$C_- = \left\{ (m_x + r_x \cdot R^{-1}(\alpha), m_x - \ell_x \cdot L^{-1}(\alpha)) \mid m_x \in \mathfrak{R}; \ell_x, r_x \geq 0 \right\}$$

In  $C_-$ , we must additionally consider  $L^{-1}(\alpha) = R^{-1}(\alpha)$  as a precondition for  $Y = A + b \cdot X$  to still be an LR-fuzzy number.



**Figure 1.** For  $b > 0$ , we have:  $y_L^\alpha = b \cdot x_L^\alpha$  and  $y_R^\alpha = b \cdot x_R^\alpha$ .



**Figure 2.** For  $b < 0$ , we have:  $y_L^\alpha = b \cdot x_R^\alpha$  and  $y_R^\alpha = b \cdot x_L^\alpha$ .

**Remark:** An additional condition is also active:  $\hat{b}_L = \hat{b}_R = \hat{b}$  (which appears only in the particular case of a non-fuzzy slope-coefficient).

Without prior information, an ex-ante specification of the model is not possible and we must take into account both projection cones  $C_+$  and  $C_-$  in order to check for admissible estimators, as solutions of the following two constrained quadratic programs:

$$QP_+^* : \inf_{\substack{b > 0, m_A \in \mathfrak{R} \\ \ell_A \geq 0, r_A \geq 0}} \delta_2^2(A + b \bar{X}_+, \hat{A} + \hat{b} \bar{X}_+);$$

where  $(\bar{X})_+ = (\bar{x}_L^\alpha, \bar{x}_R^\alpha)_{\alpha \in [0,1]}$ , i.e.

$$\inf_{\substack{b > 0, m_A \in \mathfrak{R} \\ \ell_A \geq 0, r_A \geq 0}} \frac{1}{2} \int_{[0,1]} \left[ (m_A - \hat{m}_A) - (\ell_A - \hat{\ell}_A) \cdot L^{-1}(\alpha) + (b - \hat{b})(\bar{m}_x - \bar{\ell}_x \cdot L^{-1}(\alpha))^2 + (m_A - \hat{m}_A) + (r_A - \hat{r}_A) \cdot L^{-1}(\alpha) + (b - \hat{b})(\bar{m}_x + \bar{r}_x \cdot R^{-1}(\alpha))^2 \right] \lambda(d\alpha)$$

$$QP_-^* : \inf_{\substack{b < 0, m_A \in \mathfrak{R} \\ \ell_A \geq 0, r_A \geq 0}} \delta_2^2(A + b \bar{X}_-, \hat{A} + \hat{b} \bar{X}_-);$$

where  $(\bar{X})_- = (\bar{x}_R^\alpha, \bar{x}_L^\alpha)_{\alpha \in [0,1]}$ , i.e.

$$\inf_{\substack{b < 0, m_A \in \mathbb{R} \\ \ell_A \geq 0, r_A \geq 0}} \frac{1}{2} \int_{[0,1]} \left[ \left( (m_A - \hat{m}_A) + (\ell_A - \hat{\ell}_A) \cdot L^{-1}(\alpha) + (b - \hat{b})(\bar{m}_x + \bar{r}_x \cdot R^{-1}(\alpha)) \right)^2 + \left( (m_A - \hat{m}_A) + (r_A - \hat{r}_A) \cdot R^{-1}(\alpha) + (b - \hat{b})(\bar{m}_x - \bar{\ell}_x \cdot L^{-1}(\alpha)) \right)^2 \right] \lambda(d\alpha)$$

For fuzzy models with multiple repressors, the number of projection cones to be tested becomes too large for being conveniently handled.

### Rationales for allowing coefficients with negative spreads

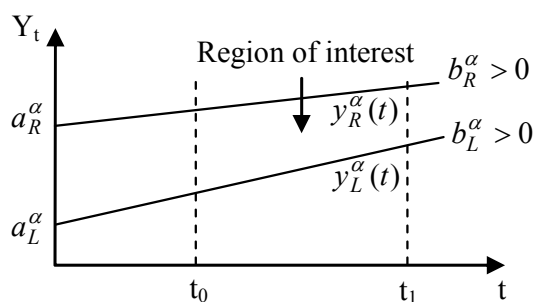
Let us consider the following fuzzy linear regression model:

$$Y = A + B \otimes X \Leftrightarrow$$

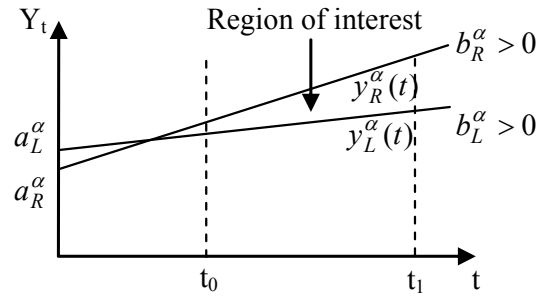
$$(y_L^\alpha, y_R^\alpha) = (a_L^\alpha, a_R^\alpha) + (b_L^\alpha, b_R^\alpha) \otimes (x_L^\alpha, x_R^\alpha)$$

We are clearly restricted to interpret the left and right slope-coefficients  $b_L^\alpha$  and  $b_R^\alpha$  as an ordered pair denoting the two bounds of an interval (i.e.  $b_L^\alpha \leq b_R^\alpha$ ). For a process with diminishing oscillations, similar to that one depicted in figure 3 by the two left and right  $\alpha$ -level bounds ( $y_L^\alpha(t)$  and  $y_R^\alpha(t)$ ), there is no solution within the fuzzy / interval arithmetic.

As long as we have reasons to consider such cases of practical interest, we need not check for the inequalities  $a_L^\alpha < a_R^\alpha$  and  $b_L^\alpha < b_R^\alpha$  and thus the pairs  $(a_L^\alpha, a_R^\alpha)$  and  $(b_L^\alpha, b_R^\alpha)$  will not be necessarily interpreted as  $\alpha$ -level intervals.



**Figure 3.** For the region of interest we have  $y_L^\alpha(t) < y_R^\alpha(t)$ ,  $a_L^\alpha < a_R^\alpha$ , but  $b_L^\alpha > b_R^\alpha$ , so  $[b_L^\alpha, b_R^\alpha]$  can't be an  $\alpha$ -level interval



**Figure 4.** For the region of interest we have  $y_L^\alpha(t) < y_R^\alpha(t)$ ,  $b_L^\alpha < b_R^\alpha$ , but  $a_L^\alpha > a_R^\alpha$ , so  $[a_L^\alpha, a_R^\alpha]$  can't be an  $\alpha$ -level interval

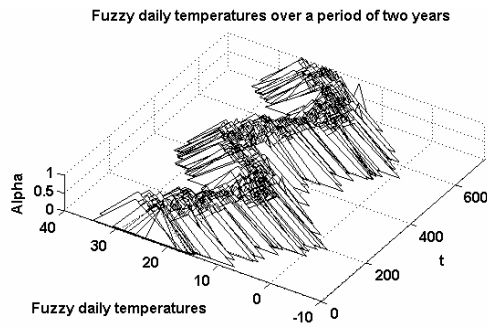
### 3. Estimating and Forecasting Fuzzy Time Series

Non-stationary time series may be analyzed by means of the fuzzy component model, which decomposes the fuzzy time series into a trend component, a cyclical or seasonal component and a fuzzy random residual component. The estimation technique proposed in this paper is based upon a partial decoupling principle, first proposed in [6]. 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  $\alpha$ -level bounds, with  $\alpha$  progressively decreasing towards 0. The estimates of modal values are not subjects 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  $\alpha$ -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.

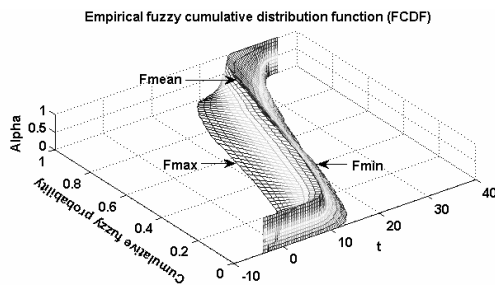
#### Fuzzy data: minimum, average and maximum daily temperatures registered at a local weather station

In what follows, we will exemplify some suitable methods for modelling and forecasting non-stationary fuzzy time series, based on the fuzzy component model. 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 5

and the corresponding empirical fuzzy cumulative distribution function in figure 6.



**Figure 5.** Fuzzy daily temperatures over a period of two years (730 days)



**Figure 6.** 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.

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

### The fuzzy time series model for daily temperatures, with fuzzy trend and fuzzy cyclical component

The fuzzy component model attached to the fuzzy time series is defined below.

$$Y(t) = T(t) + C(t) + u(t)$$

$$Y(t) = \begin{pmatrix} Y^L(t) \\ Y^C(t) \\ Y^R(t) \end{pmatrix}; T(t) = \begin{pmatrix} T^L(t) \\ T^C(t) \\ T^R(t) \end{pmatrix};$$

$$C(t) = \begin{pmatrix} C^L(t) \\ C^C(t) \\ C^R(t) \end{pmatrix}; u(t) = \begin{pmatrix} u^L(t) \\ u^C(t) \\ u^R(t) \end{pmatrix}$$

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  $u(t)$ ):

$$Y^L(t) \leq Y^C(t) \leq Y^R(t);$$

$$T^L(t) \leq T^C(t) \leq T^R(t);$$

$$C^L(t) \leq C^C(t) \leq C^R(t);$$

$$u^L(t) \leq u^C(t) \leq u^R(t)$$

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

$$T = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \dots & \dots \\ 1 & N \end{pmatrix}; N = 730$$

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:

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

### 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  $\text{Im}(T)$  generated by the columns of matrix  $T$ :

$$\hat{T}^C = P \cdot Y^C, \text{ where } P = T \cdot (T' \cdot T)^{-1} \cdot T'$$

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

$$\min_{b_{trend}^L} \left\{ (b_{trend}^L)' T' T b_{trend}^L - 2(Y^L)' T b_{trend}^L \right\}$$

$$\text{subject to: } \begin{cases} \hat{T}^L \leq \hat{T}^C \\ \hat{u}_1^L + \xi \leq \hat{u}_1^C \end{cases} \Leftrightarrow$$

$$\begin{cases} T \cdot b_{trend}^L \leq \hat{T}^C \\ -T \cdot b_{trend}^L \leq Y^C - \hat{T}^C - Y^L - \xi \end{cases}$$

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

$\hat{u}_1^L + \xi \leq \hat{u}_1^C$  means that the left spread of the intermediary fuzzy residuals (after detrending) must be strictly positive, i.e.,  $\hat{u}_1^C - \hat{u}_1^L \geq \xi > 0$ . 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  $\xi$  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:

$$\min_{b_{trend}^R} \left\{ (b_{trend}^R)' T' T b_{trend}^R - 2(Y^R)' T b_{trend}^R \right\}$$

subject to:

$$\begin{cases} -T \cdot b_{trend}^R \leq -\hat{T}^C & \Leftrightarrow \hat{T}^R \geq \hat{T}^C \\ T \cdot b_{trend}^R \leq -Y^C + \hat{T}^C + Y^R - \xi & \Leftrightarrow \hat{u}_1^R - \hat{u}_1^C \geq \xi > 0 \end{cases}$$

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

$\hat{u}_1^R - \hat{u}_1^C \geq \xi > 0$  means that the right spread of the intermediary fuzzy residuals (after detrending) must be strictly positive.

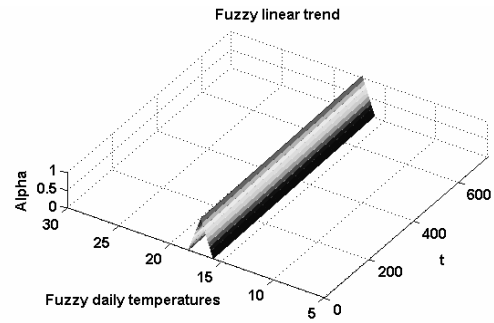


Figure 7. Fuzzy linear trend.

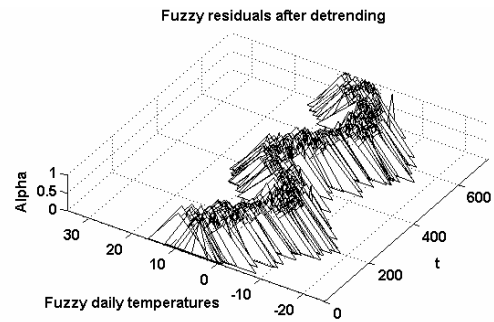


Figure 8. Intermediary fuzzy residuals after detrending (the residuals are fuzzy sets in a proper sense – with non-negative spreads).

### 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.

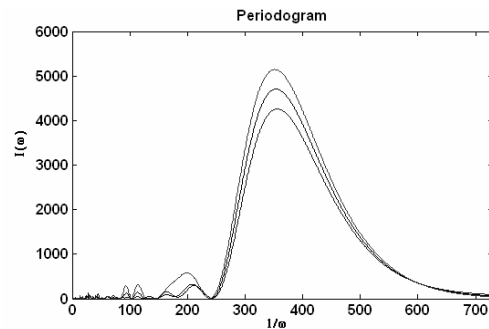


Figure 9. Periodogram with a maximum corresponding to about 365 days (one year)

**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  $\text{Im}(C)$  generated by the columns of matrix  $C$  :

$$\hat{C}^C = P \cdot \hat{u}_1^C; \text{ where } P = C \cdot (C' \cdot C)^{-1} \cdot C'$$

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

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

$$\text{subject to: } \begin{cases} \hat{C}^L \leq \hat{C}^C \\ \hat{u}_2^L + \xi \leq \hat{u}_2^C \end{cases} \Leftrightarrow$$

$$\begin{cases} C \cdot b_{cycle}^L \leq \hat{C}^C \\ -C \cdot b_{cycle}^L \leq \hat{u}_1^C - \hat{C}^C - \hat{u}_1^L - \xi \end{cases}$$

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

$\hat{u}_2^C - \hat{u}_2^L \geq \xi > 0$  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  $\xi$  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:

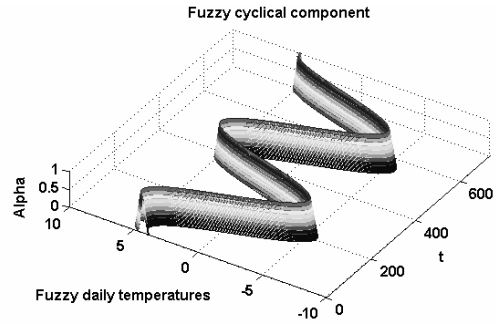
$$\min_{b_{cycle}^R} \left\{ (b_{cycle}^R)' C' C b_{cycle}^R - 2(\hat{u}_1^R)' C b_{cycle}^R \right\}$$

$$\text{subject to: } \begin{cases} \hat{C}^R \geq \hat{C}^C \\ \hat{u}_2^R - \hat{u}_2^C \geq \xi > 0 \end{cases} \Leftrightarrow$$

$$\begin{cases} -C \cdot b_{cycle}^R \leq -\hat{C}^C \\ C \cdot b_{cycle}^R \leq -\hat{u}_1^C + \hat{C}^C + \hat{u}_1^R - \xi \end{cases}$$

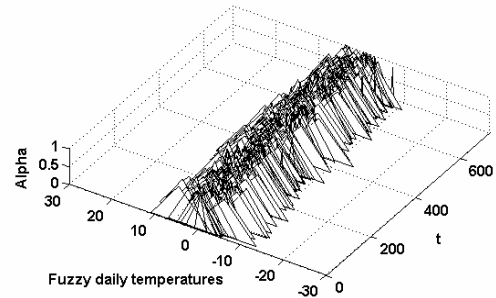
$\hat{C}^R \geq \hat{C}^C$  means that the right spread of the fuzzy cyclical component must be non-negative, i.e.,  $\hat{C}^R - \hat{C}^C \geq 0$

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



**Figure 10.** Fuzzy cyclical component.

Fuzzy residuals after removing both trend and cyclical component

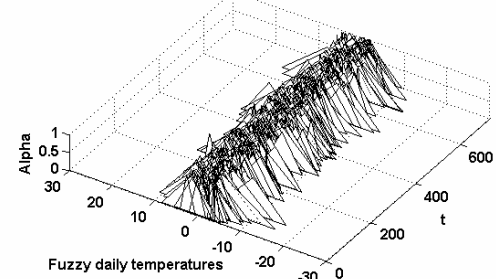


**Figure 11.** Fuzzy residuals after removing both trend and cyclical component (the residuals are fuzzy sets in a proper sense – with non-negative spreads).

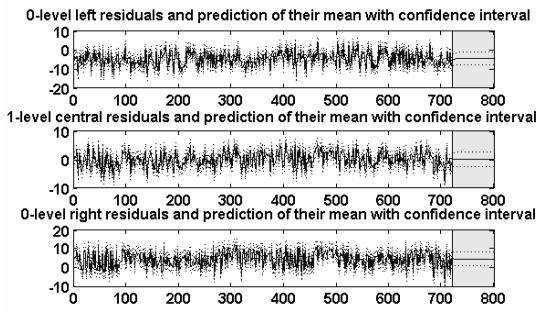
### Modeling and forecasting the fuzzy residuals as a VAR(4) process, after removing both trend and cyclical component

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.

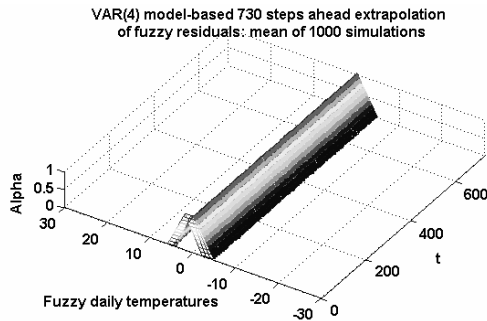
VAR(4) model-based 730 steps ahead extrapolation of fuzzy residuals: a single simulation



**Figure 12.** 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).

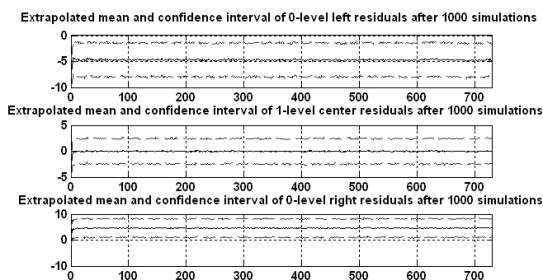


**Figure 13.** Fuzzy residuals (left, central and right) and prediction of their mean with confidence interval ( $\pm\sigma$ ).

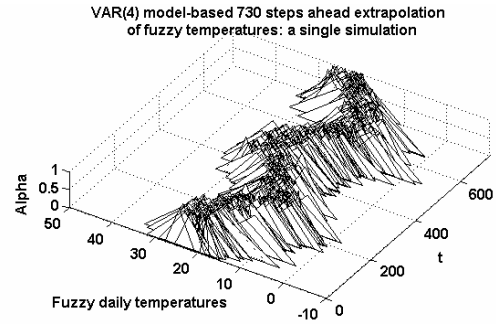


**Figure 14.** VAR(4) model-based 730 steps ahead extrapolation of fuzzy residuals: mean of 1000 simulations.

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.



**Figure 15.** Extrapolated mean and confidence interval of 0-level left / 1-level center / 0-level right fuzzy residuals, after 1000 simulations.



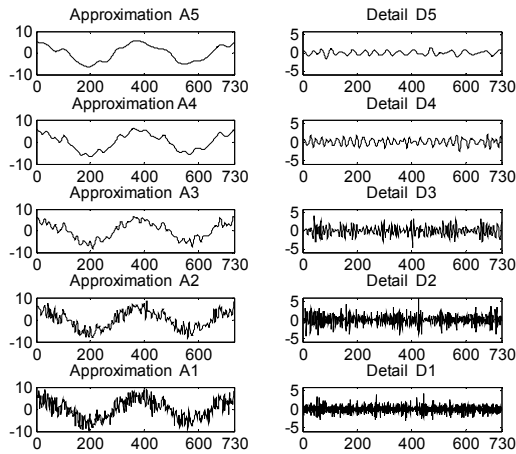
**Figure 16.** VAR(4) model-based 730 steps ahead extrapolation of fuzzy temperatures, additively recomposed from trend, cyclical component and simulated fuzzy residuals.

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

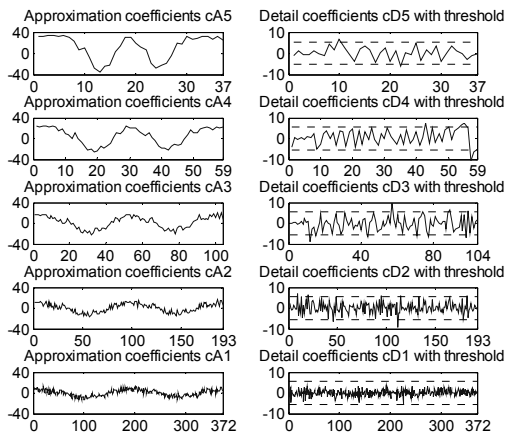
Another alternative to removing disturbances from a time series is de-noising data by wavelet decomposition.

The Discrete Wavelet Transform (DWT, [15]) 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 ([4]). 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 8. A level 5 decomposition with Sym8 wavelets and a fixed form soft thresholding is first performed (see figures 17 and 18).

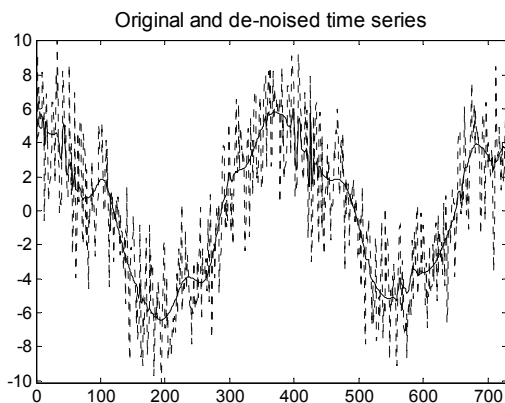


**Figure 17.** A level 5 decomposition of the average temperature time series using Sym8 wavelets: approximations and details. The successive approximations appear less and less noisy; however, they also lose progressively more high-frequency information.

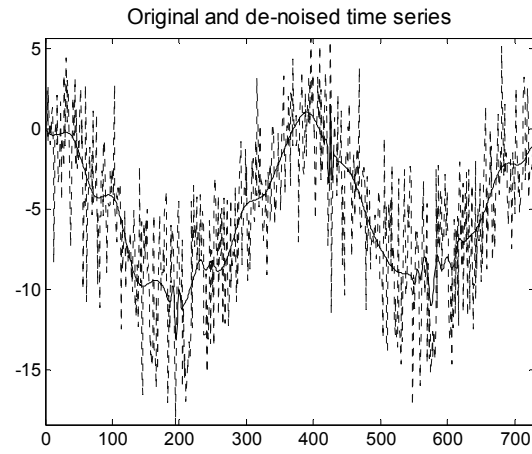


**Figure 18.** Approximation coefficients and detail coefficients with a global threshold.

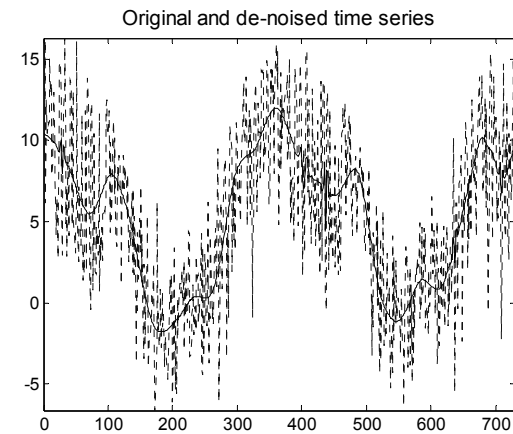
The initially de-trended average, minimum and maximum temperature time series are now de-noised in turn (see figures 19, 20 and 21).



**Figure 19.** De-trended vs. de-trended & de-noised average temperature time series



**Figure 20.** De-trended vs. de-trended & de-noised minimum temperature time series



**Figure 21.** 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.

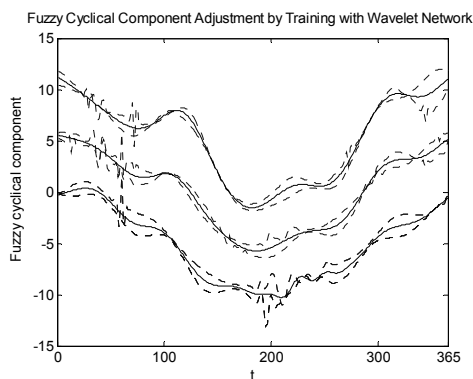
Wavelet networks attempt to combine the properties of the Wavelet decomposition previously described, along with the learning capabilities of feedforward neural networks. They employ wavelets instead of sigmoidal activations functions, are trained with a backpropagation-like algorithm and behave as universal 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  $L_2(\mathfrak{R}^n)$  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:

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

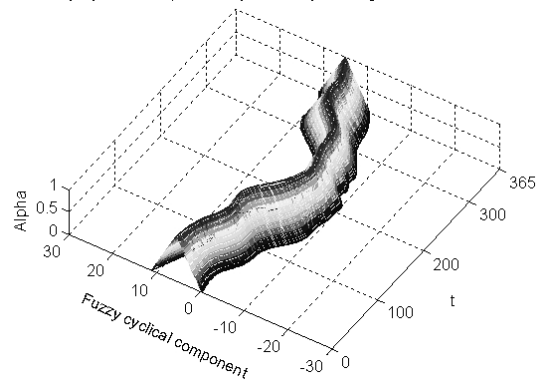
where  $\psi$  is a radial wavelet function,  $d_i \in \mathfrak{R}^n$  are dilatation parameters,  $t_i \in \mathfrak{R}^n$  are translation parameters,  $w_i \in \mathfrak{R}$  are linear weights,  $N$  is the number of wavelets,  $c \in \mathfrak{R}^n$  is the additional direct linear combination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

The input space is the set  $T = \{1, 2, \dots, 365\}$  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.,  $S_{\min}$ ,  $S_{\text{avg}}$  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  $S_{\min}$ ,  $S_{\text{avg}}$  and  $S_{\max}$ , i.e.  $\hat{S}_{\min}(t): T \rightarrow S_{\min}$ ,  $\hat{S}_{\text{avg}}(t): T \rightarrow S_{\text{avg}}$  and  $\hat{S}_{\max}(t): T \rightarrow S_{\max}$ , 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 22 and 23.



**Figure 22.** 2D representation of the adjusted fuzzy cyclical component after training with wavelet network

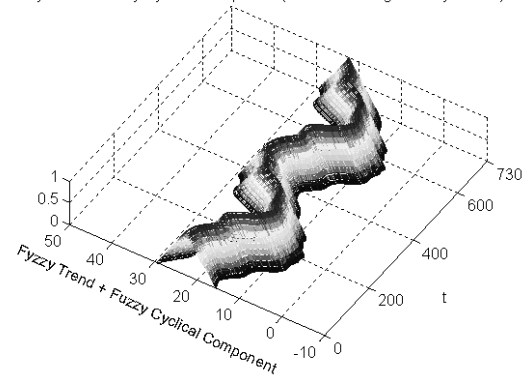
Fuzzy Cyclical Component Adjustment by Training with Wavelet Network



**Figure 23.** 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 24).

Fuzzy Trend + Fuzzy Cyclical Component (after De-noising and Adjustment)



**Figure 24.** Fuzzy trend + fuzzy cyclical component (after de-noising and adjustment)

## 5. A Fuzzy Generalization of Buys Ballot Decomposition Model for Trend and Seasonality

The fuzzy component model attached to the fuzzy time series of monthly temperatures is defined below, where  $S(t)$  is the seasonal component.

$$Y(t) = T(t) + S(t) + u(t)$$

$$Y(t) = \begin{pmatrix} Y^L(t) \\ Y^C(t) \\ Y^R(t) \end{pmatrix}; T(t) = \begin{pmatrix} T^L(t) \\ T^C(t) \\ T^R(t) \end{pmatrix};$$

$$S(t) = \begin{pmatrix} S^L(t) \\ S^C(t) \\ S^R(t) \end{pmatrix}; u(t) = \begin{pmatrix} u^L(t) \\ u^C(t) \\ u^R(t) \end{pmatrix}$$

$$Y^L(t) \leq Y^C(t) \leq Y^R(t);$$

$$T^L(t) \leq T^C(t) \leq T^R(t);$$

$$S^L(t) \leq S^C(t) \leq S^R(t);$$

$$u^L(t) \leq Y^C(t) \leq Y^R(t)$$

$$T = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \dots & \dots \\ 1 & N \end{pmatrix}; \quad S = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

$N = 24$ , period = 12.

The method is similar to that used in the case of daily temperatures, except for the seasonal component, for which we use the matrix  $S$  defined above, instead of matrix  $C$ .

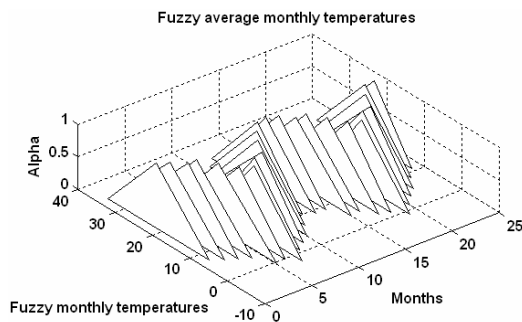


Figure 25. Fuzzy monthly temperatures

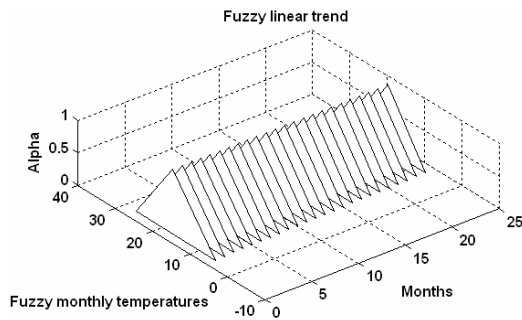


Figure 26. Fuzzy linear trend

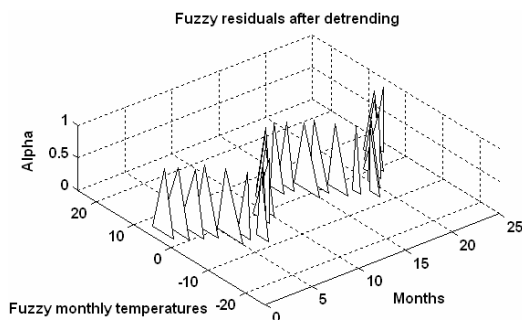


Figure 27. Fuzzy residuals after de-trending (the residuals are fuzzy sets in a proper sense – with non-negative spreads).

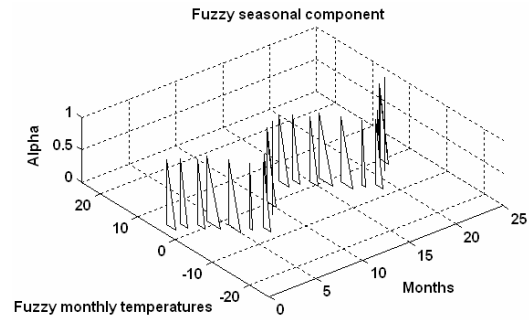


Figure 28. Fuzzy seasonal component.

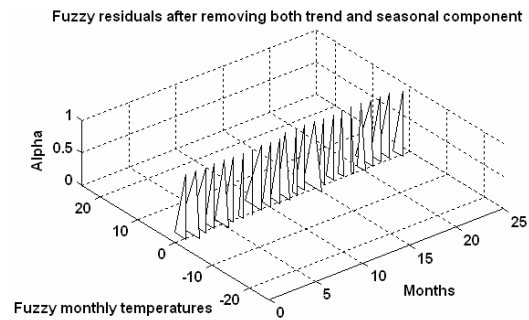


Figure 29. Fuzzy residuals after removing both trend and seasonal component (the residuals are fuzzy sets in a proper sense – with non-negative spreads)

## 6. Conclusion

This paper proposed suitable new methods for fuzzy time series estimation and prediction, using both the estimation theory and Computational Intelligence techniques.

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 (first proposed in [6]), 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. DIAMOND, P., **Fuzzy Least Squares**, Inform. Sci. 46, 1988, p. 141–157.
2. DIAMOND, P., P. KLOEDEN, **Metric Spaces of Fuzzy Sets: Theory and Applications**, World Scientific, Singapore, 1994.
3. DIAMOND, P., R. KÖRNER, **Extended Fuzzy Linear Models and Least Squares Estimates**, Comput. Math. Appl. 33, 1997, pp. 15–32.
4. DONOHO, D., **Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data**. In Different Perspectives on Wavelets, I. Daubechies (eds). Amer. Math. Soc., Providence, R.I., 1993, pp. 173–205.
5. EL ABIDINE SKHIRI, M. Z., CHTOUROU, M., **Synthesis of Denoising Wavelet Neural Networks**, Studies in Informatics and Control, 17/4, 2008
6. GEORGESCU, V., **Estimation of Fuzzy Regression Models using Quadratic Programming**, Economic Computation and Economic Cybernetics Studies and Research, 4, 1997, pp. 105-124.
7. GEORGESCU, V., **New Estimation Methods in Fuzzy Regression Analysis, Based on Projection Theorem and Decoupling Principle**, Fuzzy Economic Review, III/1, 1998, pp. 21-38.
8. GEORGESCU, V., **Fuzzy Control Applied to Economic Stabilization Policies**, Studies in Informatics and Control, 10/1, 2001, pp. 37-60
9. GEORGESCU, V., **Multivariate Fuzzy-termed Data Analysis: Issues and Methods**, Fuzzy Economic Review, VI/1, 2001, pp. 19-48.
10. GEORGESCU, V., **On the Foundations of Granular Computing Paradigm**, Fuzzy Economic Review, VIII/2, 2003, pp. 73-105.
11. GEORGESCU, V., **A Generalization of Symbolic Data Analysis Allowing the Processing of Fuzzy Granules**, Lecture Notes in Artificial Intelligence, 3131, Springer 2004, pp. 215-226
12. GEORGESCU, V., **Granular vs. Point-Wise Metrics and Statistics to Accommodate Machine Learning Algorithms in Granular Feature Spaces**, Fuzzy Economic Review, XII/2, 2007, pp. 45-74.
13. GEORGESCU, V., **A Time Series Knowledge Mining Framework Exploiting the Synergy between Subsequence Clustering and Predictive Markovian Models**, Fuzzy Economic Review, XIV/1, 2009, pp. 41-66
14. HUKUHARA, M., **Integration des applications mesurables dont la valeur est un compact convexe**, Funkcialaj Ekvacioj 10, 1967, pp. 205–223.
15. KÖRNER, R., W. NÄTHER, **Linear Regression with Random Fuzzy Variables: Extended Classical Estimates, Best Linear Estimates, Least Squares Estimates**, Inform. Sci. 109, 1998, pp. 95–118.
16. MALLAT, S. G., PEYRÉ, G., **A Wavelet Tour of Signal Processing: The Sparse Way**, Academic Press, 3<sup>rd</sup> Edition, 2009.
17. MÖLLER, B., U. REUTER, **Uncertainty Forecasting in Engineering**, Springer-Verlag, Berlin Heidelberg, 2007.
18. PURI, M.L., D.A. RALESCU, **Fuzzy random variables**, J. Math. Anal. Appl. 114, 1986, pp. 409–422.
19. RĂDULESCU F., **Cyclic Hilbert Spaces**, Studies in Informatics and Control, 18/1, 2009
20. STEFANINI, L., **A Generalization of Hukuhara Difference and Division for Interval and Fuzzy Arithmetic**, Fuzzy Sets and Systems (2009), doi: 10.1016/j.fss.2009.06.009