RAIRO-Oper. Res. 50 (2016) 413–435
DOI: 10.1051/ro/2015044

RAIRO Operations Research
www.rairo-ro.org

FUZZY PREDICTION STRATEGIES FOR GENE-ENVIRONMENT NETWORKS – FUZZY REGRESSION ANALYSIS FOR TWO-MODAL REGULATORY SYSTEMS

Erik Kropat[1], Ayşe Özmen[2], Gerhard-Wilhelm Weber[2,*], Silja Meyer-Nieberg[3] and Ozlem Defterli[4]

Abstract. Target-environment networks provide a conceptual framework for the analysis and prediction of complex regulatory systems such as genetic networks, eco-finance networks or sensor-target assignments. These evolving networks consist of two major groups of entities that are interacting by unknown relationships. The structure and dynamics of the hidden regulatory system have to be revealed from uncertain measurement data. In this paper, the concept of fuzzy target-environment networks is introduced and various fuzzy possibilistic regression models are presented. The relation between the targets and/or environmental entities of the regulatory network is given in terms of a fuzzy model. The vagueness of the regulatory system results from the (unknown) fuzzy coefficients. For an identification of the fuzzy coefficients' shape, methods from fuzzy regression are adapted and made applicable to the bi-level situation of target-environment networks and uncertain data. Various shapes of fuzzy coefficients are considered and the control of outliers is discussed. A first numerical example is presented for purposes of illustration. The paper ends with a conclusion and an outlook to future studies.

Mathematics Subject Classification. 92-08, 92D10, 62J86.

Received September 8, 2015. Accepted September 21, 2015.

1. Introduction

Interconnected networks with multiple connected groups of entities arise in many applications ranging from the prediction of genetic regulatory patterns in computational biology and the modelling and simulation of eco-finance networks to the formation of multisensor-multitarget networks in NBC-tracking scenarios [19]. In this

Keywords. Fuzzy evolving networks, fuzzy target-environment networks, uncertainty, fuzzy theory, fuzzy regression analysis, possibilistic regression, forecasting.

[1] Institute for Applied Computer Science, Universität der Bundeswehr München, 85577 Neubiberg, Germany. erik.kropat@unibw.de

[2] Institute of Applied Mathematics, Middle East Technical University, 06531 Ankara, Turkey. ayseozmen19@gmail.com; gweber@metu.edu.tr

[3] Institute for Theoretical Computer Science, Mathematics and Operations Research, Universität der Bundeswehr München, 85577 Neubiberg, Germany. silja.meyer-nieberg@unibw.de

[4] Faculty of Arts and Sciences, Department of Mathematics and Computer Science, Çankaya University, 06810 Ankara, Turkey. defterli@cankaya.edu.tr

* Honorary positions: Faculty of Economics, Business and Law, University of Siegen, Germany; Center for Research on Optimization and Control, University of Aveiro, Portugal; University of North Sumatra, Indonesia.

paper, we are focusing on the important group of so-called *target-environment networks under uncertainty* [16]. These *two-modal regulatory systems* are composed of two distinct groups of data which define different but strongly related levels of the model. The first group comprises the entities or *targets* under observation, which clearly are the most important variables of the regulatory system. The second group consists of a certain number of additional *environmental factors* that can have a strong impact on the targets regulatory patterns. These factors can act as controls and/or disturbances. The hidden interactions between the entities of the system have to be revealed from measurement data. Here, data uncertainty plays an important role with regard to modelling and prediction of the future states of the two-modal regulatory system.

An important example of two-modal regulatory systems are the so-called *gene-environment networks*, which were introduced in the genetic context by Weber *et al.* [4,7,28,32,35,40,41]. Here, the expression values of genes or proteins are the target variables under consideration. Additional environmental factors like toxins, transcription factors or other components of the metabolic pathways may take a strong influence on the targets. Since microarray experiments as well as environmental observations usually result in uncertain data, this approach has been further extended in order to deal with errors and data uncertainty. The papers [29,33,34,36–39] focus on gene-environment networks where noise and uncertainty are represented in terms *error intervals*. For an estimation of the unknown system parameters, a *regression analysis* based on *interval-arithmetics* is applied leading to *generalized Chebychev approximation problems* and regression problems to be solved by methods of *generalized semi-infinite optimization* [30,31]. Recently, gene-environment networks under *ellipsoidal uncertainty* have been introduced in [14–17]. In this approach, functionally related groups of variables are identified with data mining methods and the uncertain states of targets and environmental clusters are represented in terms of ellipsoids. An affine-linear model based on *ellipsoidal calculus* is applied to predict the future ellipsoidal states of the system and the estimation of system parameters is based on a *set-theoretic regression analysis*.

In the last decade, the concept of target-environment networks has been continuously developed and now provides a conceptual framework for many regulatory systems in computational biology and life sciences. In addition, *target-environment networks* have also been applied to financial sciences, where so-called *eco-finance networks* are introduced in [13, 39].

The quality of the regression models depends heavily on the quality of the available data sets. For example, modern high-throughput technologies can be used to measure the expression profiles of a large number of genes simultaneously, but at a limited number of reading points. *Regression analysis* can be applied to identify the functional relationship between independent and dependent variables, where both variables are given as real numbers [8]. Nevertheless, for classical regression analysis, measurements have to be taken at a high number of reading points in order to obtain valid statistical relations between the dependent and independent variables, which can be considered as too expensive in the genetic context. In addition, in classical regression analysis the linearity assumption has to be fulfilled, so that gene-environment networks are clearly out of the scope of classical regression.

In situations where these assumptions are not fulfilled, where imprecise data with not normally distributed errors have to be considered or where a vagueness in the relationship between input and output variables exists, *fuzzy-regression analysis* offers a more general viewpoint and provides means for tackling problems failing to satisfy these assumptions. Unlike classical regression, deviations between observed values and estimated values are assumed to be due to *system fuzziness* or *fuzziness of regression coefficients* [2].

In this paper, we introduce the new concept of *fuzzy target-environment networks* and discuss the related fuzzy regression models. The vagueness of the relation between the targets and/or environmental factors of such a regulatory network results from the (unknown) fuzzy coefficients of the underlying *fuzzy model* and it is no longer determined by precise crisp coefficients. For an identification of the shape of the fuzzy coefficients, methods from fuzzy regression have to be adapted and made applicable to the bi-level situation of target-environment systems and data.

Fuzzy regression as a variation of classical regression has been studied by many authors and we refer to [10] for a recent literature review on fuzzy regression approaches and applications. In general, there are two types of fuzzy regression methods – *possibilistic regression*, which is based on Tanaka's linear programming approach [25] and *fuzzy least-squares regression* [6]. In this paper, we focus on possibilistic regression and adapt various extensions of the *fuzzy regression problem* introduced by Tanaka *et al.* [25]. This model was based on crisp input vectors as well as fuzzy output vectors and used fuzzy coefficients, which were represented by symmetric triangular fuzzy numbers. The underlying idea was to minimize the fuzziness of the model by minimizing the spread of the fuzzy output or the total support of the fuzzy coefficients subject to all the given data. This basic model has been further extended in several directions in order to deal with potential limitations of possibilistic regression. For example, in possibilistic regression based on symmetric triangular fuzzy numbers, only the extremal data points determine the structure of the model. All others data points have no impact on the structure what results in a high sensitivity to outliers [20, 21].

This problem can be resolved by using asymmetric triangular or trapezoidal fuzzy numbers [3,9]. Since Tanaka *et al.* have introduced the concept of fuzzy regression, several fuzzy regression approaches have been proposed, often referring to a particular nature of input-output data. Some authors focus on crisp input-crisp output data [23], others use mixed crisp input-fuzzy output data [25] or fuzzy input-fuzzy output data [22]. Although possibilistic regression has been successfully applied in many areas of engineering sciences and Operations Research, methods involving fuzzy concepts have been rarely applied to genetics [1].

In this study, we consider fuzzy possibilistic regression for target-environment networks affected by errors and uncertainty. We present various fuzzy regression algorithms for target-environment data based on different representations of the fuzzy coefficients of the underlying fuzzy model. The algorithms are applied to crisp input-crisp output data. In addition, by assigning individual membership grades to input-output samples, the influence of outliers can be softened and controlled.

The paper is organized as follows: In Section 1, the concept of fuzzy target-environment networks and the corresponding fuzzy regression model with fuzzy coefficients are introduced. In Section 2, we adapt Tanaka's possibilistic regression model for crisp target-environment data and introduce various fuzzy regression algorithms. To overcome the limitations of this approach, we consider different shapes of fuzzy coefficients in terms of symmetric and asymmetric triangular fuzzy sets as well as symmetric and asymmetric trapezoidal fuzzy numbers. In addition, we consider models where membership grades are assigned to input-output data in order to deal with outliers. Section 4 presents an illustrative example of fuzzy prediction for a gene-environment regulatory network. We conclude with an outlook on potential directions of research.

2. Fuzzy target-environment networks and fuzzy regression

In this section, the concept of *fuzzy target-environment networks* is introduced. A *linear fuzzy model* determines the synergistic connections between the targets and the additional environmental entities. Various algorithms for an estimation of the unknown fuzzy coefficients of the fuzzy model are discussed in Section 3.

2.1. The fuzzy model

Target-environment networks and their inherent dynamics are often modeled by time-discrete systems

$$X^{(k+1)} = F\left(X^{(k)}, E^{(k)}\right),$$
$$E^{(k+1)} = G\left(X^{(k)}, E^{(k)}\right),$$

for $k \geq 0$, where the time-dependent n-vector $X^{(k)} = (X_1^{(k)}, \dots, X_n^{(k)})^T$ denotes the expression values of the n targets and the m-vector $E = (E_1^{(k)}, \dots, E_m^{(k)})^T$ represents the values of the m environmental items. Both linear and nonlinear models are available, where $F : \mathbb{R}^{n+m} \to \mathbb{R}^n$ and $G : \mathbb{R}^{n+m} \to \mathbb{R}^m$ describe the linear or nonlinear dynamics of the system.

In this paper, we focus on the linear dynamics of single targets and environmental items. The time-discrete dynamics of each target, X_j $(j = 1, \dots, n)$, is represented by a $(n + m)$-input and single-output linear fuzzy

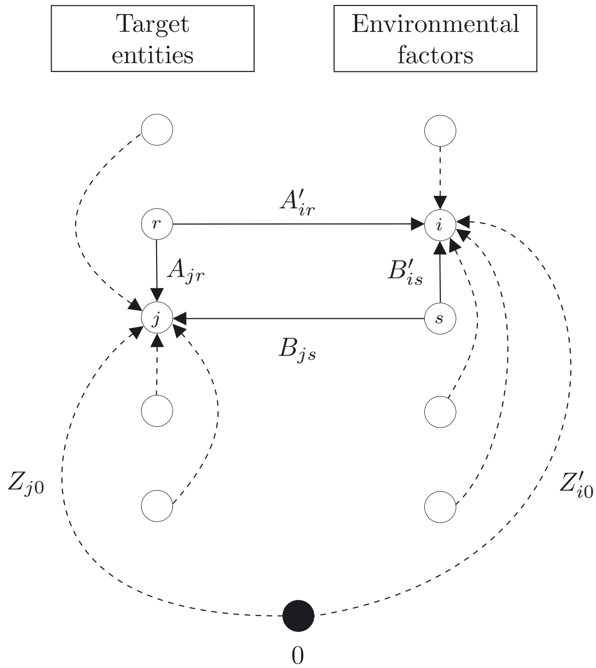

FIGURE 1. The fuzzy target-environment network. The nodes are the targets and environmental factors. The branches are weighted by the fuzzy coefficients of the fuzzy models $\mathcal{F}_j$ and $\mathcal{G}_i$.

system

$$X_j^{(k+1)} := \mathcal{F}_j\left(X^{(k)}, E^{(k)}\right) = Z_{j0} + \sum_{r=1}^{n} A_{jr} X_r^{(k)} + \sum_{s=1}^{m} B_{js} E_s^{(k)} \quad (k \in \mathbb{N}_0).$$

Similarly, the states of the environmental items, E_i $(i = 1, \ldots, m)$, are given by

$$E_i^{(k+1)} := \mathcal{G}_i\left(X^{(k)}, E^{(k)}\right) = Z'_{i0} + \sum_{r=1}^{n} A'_{ir} X_r^{(k)} + \sum_{s=1}^{m} B'_{is} E_s^{(k)} \quad (k \in \mathbb{N}_0).$$

The unknown fuzzy coefficients $Z_{j0}, A_{jr}, B_{js}, Z'_{i0}, A'_{ir}, B'_{is}$ of the fuzzy models $\mathcal{F}_j$ and $\mathcal{G}_i$ have to be determined from *crisp data vectors*

$$\overline{\mathbb{X}}^{(\kappa)} = \left(\overline{X}_1^{(\kappa)}, \ldots, \overline{X}_n^{(\kappa)}\right)^T \quad \text{and} \quad \overline{\mathbb{E}}^{(\kappa)} = \left(\overline{E}_1^{(\kappa)}, \ldots, \overline{E}_m^{(\kappa)}\right)^T,$$

with $\kappa = 0, 1, \ldots, T+1$, obtained from measurements taken at reading points $t_0 < t_1 < \ldots < t_{T+1}$. For the initial states of the linear fuzzy system we assume $X_r^{(0)} = \overline{X}_r^{(0)}$ and $E_s^{(0)} = \overline{E}_s^{(0)}$ $(r = 1, \ldots, n;\ s = 1, \ldots, m)$.

2.2. Fuzzy target-environment networks

The uncertain relations between the targets and environmental factors of the fuzzy model can be represented in terms of a highly interconnected regulatory network (*cf*. Fig. 1). The nodes of this *fuzzy target-environment network* are given by the targets and environmental items. The branches between targets and/or environmental factors are weighted by the corresponding fuzzy coefficients that define the coupling rules of the fuzzy model. In order to include the intercepts Z_{j0} and Z'_{i0} in our network, we introduce an additional node **0**. We note that also weights can be assigned to the nodes of the fuzzy network. This can be, *e.g.*, the outputs (or some measure of the outputs) of the fuzzy model. Although the weights of the branches are static, the evolution of the states of the targets and environmental items turns the system into a time-dependent *fuzzy evolving network*. Hereby, fuzzy-discrete mathematics and its network algorithms in both versions, statically and dynamically, becomes applicable on subjects such as connectedness, components, clusters, cycles, shortest paths or further subnetworks [12, 18]. Beside these discrete-combinatorial aspects, combinatorial relations between graphs and (nonlinear) optimization problems as well as topological properties of regulatory networks can be analyzed.

2.3. Fuzzy regression

The basic idea of fuzzy regression is to minimize the fuzziness of the fuzzy models $\mathcal{F}_j$ and $\mathcal{G}_i$. In case of *non-fuzzy data* they have to include all the given input-output data in their *level sets*[5], *i.e.*,

$$\overline{X}_j^{(\kappa+1)} \in \left[\mathcal{F}_j\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)\right]_\alpha, \quad \overline{E}_i^{(\kappa+1)} \in \left[\mathcal{G}_i\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)\right]_{\alpha'} \quad (\kappa = 0, 1, \ldots, T).$$

The *inclusion relations* for target and environmental data sets depend on the level sets of the fuzzy models $\mathcal{F}_j$ and $\mathcal{G}_i$ with parameters $\alpha, \alpha' \in (0,1]$, which have to be given by the practitioner according to the desired spread of the fuzzy models. They are usually unequal what refers to the individual behaviour of the two distinct groups of data.

In the following sections, we introduce various *fuzzy regression models* for fuzzy target-environment networks. These models are based on crisp measurement data as well as many different kinds of fuzzy coefficients.

3. Fuzzy regression analysis for target-environment data

In this section, we focus on fuzzy regression models for non-fuzzy target-environment data. The fuzzy coefficients of the linear fuzzy models $\mathcal{F}$ and $\mathcal{G}$ have to be determined from *non-fuzzy input* data vectors

$$\overline{\mathbb{X}}^{(\kappa)} = \left(\overline{X}_1^{(\kappa)}, \ldots, \overline{X}_n^{(\kappa)}\right)^T \in \mathbb{R}^n \quad \text{and} \quad \overline{\mathbb{E}}^{(\kappa)} = \left(\overline{E}_1^{(\kappa)}, \ldots, \overline{E}_m^{(\kappa)}\right)^T \in \mathbb{R}^m,$$

with $\kappa = 0, 1, \ldots, T+1$.

3.1. Fuzzy regression based on symmetric triangular fuzzy coefficients

In the first fuzzy regression model, the coefficients of the fuzzy model are given by *symmetric triangular fuzzy numbers*. As we are interested in the dynamics of single targets and environmental factors, our regression analysis will be based on crisp data sets

$$\left(\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)^T; \overline{X}_j^{(\kappa+1)}\right), \left(\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)^T; \overline{E}_i^{(\kappa+1)}\right) \quad (\kappa = 0, 1, \ldots, T).$$

The symmetric triangular fuzzy coefficients can be represented in terms of their center (C) and width (W) (*cf*. Fig. 2):

$$\begin{aligned} Z_{j0} &= \left(Z_{j0}^C, Z_{j0}^W\right)^T, \quad A_{jr} = \left(A_{jr}^C, A_{jr}^W\right)^T, \quad B_{js} = \left(B_{js}^C, B_{js}^W\right)^T, \\ Z'_{i0} &= \left(Z'^C_{i0}, Z'^W_{i0}\right)^T, \quad A'_{ir} = \left(A'^C_{ir}, A'^W_{ir}\right)^T, \quad B'_{is} = \left(B'^C_{is}, B'^W_{is}\right)^T. \end{aligned}$$

[5] The *r-level* (or *r-cut*) of a fuzzy set $\mu : \mathbb{R} \to [0,1]$ is defined for $0 < r \leq 1$ as the set $[\mu]_r := \{x \in \mathbb{R} \,|\, \mu(x) \geq r\}$.

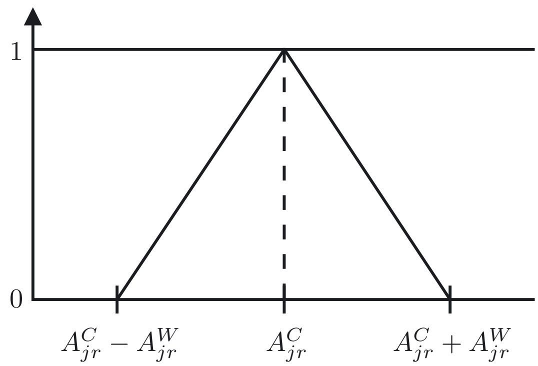

FIGURE 2. The symmetric triangular fuzzy coefficient $A_{jr} = (A^C_{jr}, A^W_{jr})^T$.

Applying interval arithmetic [11], the fuzzy model $\mathcal{F}_j$ can be rewritten as

$$\begin{aligned}
\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right) &= Z_{j0} + \sum_{r=1}^{n} A_{jr} X_r^{(k)} + \sum_{s=1}^{m} B_{js} E_s^{(k)} \\
&= \left(Z^C_{j0}, Z^W_{j0}\right)^T + \sum_{r=1}^{n} \left(A^C_{jr}, A^W_{jr}\right)^T X_r^{(k)} + \sum_{s=1}^{m} \left(B^C_{js}, B^W_{js}\right)^T E_s^{(k)} \\
&= \left(Z^C_{j0} + \sum_{r=1}^{n} A^C_{jr} \cdot X_r^{(k)} + \sum_{s=1}^{m} B^C_{js} \cdot E_s^{(k)},\ Z^W_{j0} + \sum_{r=1}^{n} A^W_{jr} \cdot \left| X_r^{(k)} \right| + \sum_{s=1}^{m} B^W_{js} \cdot \left| E_s^{(k)} \right| \right)^T .
\end{aligned}$$

Thus, $\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right)$ is a symmetric triangular fuzzy number

$$\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right) = \left(\mathcal{F}^C_j\left(X^{(k)}, E^{(k)}\right), \mathcal{F}^W_j\left(X^{(k)}, E^{(k)}\right)\right)^T$$

with

$$\begin{aligned}
\mathcal{F}^C_j\left(X^{(k)}, E^{(k)}\right) &= Z^C_{j0} + \sum_{r=1}^{n} A^C_{jr} \cdot X_r^{(k)} + \sum_{s=1}^{m} B^C_{js} \cdot E_s^{(k)}, \\
\mathcal{F}^W\left(X^{(k)}, E^{(k)}\right) &= Z^W_{j0} + \sum_{r=1}^{n} A^W_{jr} \cdot \left| X_r^{(k)} \right| + \sum_{s=1}^{m} B^W_{js} \cdot \left| E_s^{(k)} \right| .
\end{aligned}$$

Similarly, the fuzzy model $\mathcal{G}$ can be represented as the symmetric triangular fuzzy number

$$\mathcal{G}_i\left(X^{(k)}, E^{(k)}\right) = \left(\mathcal{G}^C_i\left(X^{(k)}, E^{(k)}\right), \mathcal{G}^W_i\left(X^{(k)}, E^{(k)}\right)\right)^T,$$

where

$$\begin{aligned}
\mathcal{G}^C_i\left(X^{(k)}, E^{(k)}\right) &= Z'^C_{i0} + \sum_{r=1}^{n} A'^C_{ir} \cdot X_r^{(k)} + \sum_{s=1}^{m} B'^C_{is} \cdot E_s^{(k)}, \\
\mathcal{G}^W_i\left(X^{(k)}, E^{(k)}\right) &= Z'^W_{i0} + \sum_{r=1}^{n} A'^W_{ir} \cdot \left| X_r^{(k)} \right| + \sum_{s=1}^{m} B'^W_{is} \cdot \left| E_s^{(k)} \right| .
\end{aligned}$$

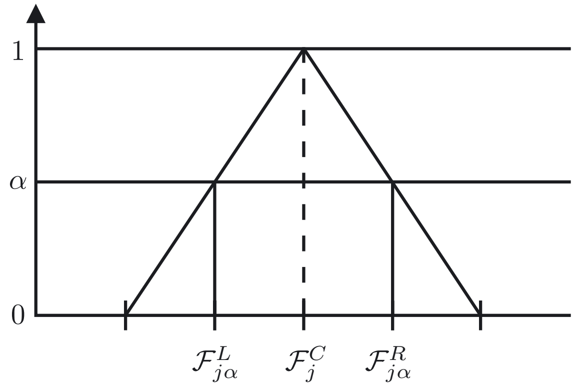

FIGURE 3. The α-cut of the fuzzy model $\mathcal{F}_j$.

According to the basic idea of fuzzy regression, we have to determine fuzzy models $\mathcal{F}_j$ and $\mathcal{G}_i$ which include all the given input-output sets $\left(\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)^T ; \overline{X}_j^{(\kappa+1)}\right)$ and $\left(\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)^T ; \overline{E}_i^{(\kappa+1)}\right)$ in their level sets. The α-cut of $\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right)$ with $\alpha \in (0,1]$ as depicted in Figure 3 is given by the interval

$$\left[\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right)\right]_\alpha = \left[\mathcal{F}^L_{j\alpha}\left(X^{(k)}, E^{(k)}\right), \mathcal{F}^R_{j\alpha}\left(X^{(k)}, E^{(k)}\right)\right],$$

where

$$\begin{aligned}\mathcal{F}^L_{j\alpha}\left(X^{(k)}, E^{(k)}\right) &= \mathcal{F}^C_j\left(X^{(k)}, E^{(k)}\right) - (1-\alpha)\cdot\mathcal{F}^W_j\left(X^{(k)}, E^{(k)}\right),\\ \mathcal{F}^R_{j\alpha}\left(X^{(k)}, E^{(k)}\right) &= \mathcal{F}^C_j\left(X^{(k)}, E^{(k)}\right) + (1-\alpha)\cdot\mathcal{F}^W_j\left(X^{(k)}, E^{(k)}\right).\end{aligned}$$

Similarly, the α'-cut of $\mathcal{G}_i\left(X^{(k)}, E^{(k)}\right)$ with $\alpha' \in (0,1]$ takes the form

$$\left[\mathcal{G}_i\left(X^{(k)}, E^{(k)}\right)\right]_{\alpha'} = \left[\mathcal{G}^L_{i\alpha'}\left(X^{(k)}, E^{(k)}\right), \mathcal{G}^R_{i\alpha'}\left(X^{(k)}, E^{(k)}\right)\right],$$

where

$$\begin{aligned}\mathcal{G}^L_{i\alpha'}\left(X^{(k)}, E^{(k)}\right) &= \mathcal{G}^C_i\left(X^{(k)}, E^{(k)}\right) - (1-\alpha')\cdot\mathcal{G}^W_i\left(X^{(k)}, E^{(k)}\right),\\ \mathcal{G}^R_{i\alpha'}\left(X^{(k)}, E^{(k)}\right) &= \mathcal{G}^C_i\left(X^{(k)}, E^{(k)}\right) + (1-\alpha')\cdot\mathcal{G}^W_i\left(X^{(k)}, E^{(k)}\right).\end{aligned}$$

Therefore, the states $\overline{X}_j^{(\kappa+1)}$ and $\overline{E}_i^{(\kappa+1)}$ have to fulfill the constraints

$$\begin{aligned}\mathcal{F}^L_{j\alpha}\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) \le \overline{X}_j^{(\kappa+1)} \le \mathcal{F}^R_{j\alpha}\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right),\\ \mathcal{G}^L_{i\alpha'}\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) \le \overline{E}_i^{(\kappa+1)} \le \mathcal{G}^R_{i\alpha'}\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right),\end{aligned}$$

for all $\kappa \in \{0,1,\ldots,T\}$. As mentioned before, the inclusion relations for target and environmental data sets depend on level sets with (unequal) parameters $\alpha, \alpha' \in (0,1]$. We introduce also some additional conditions on the size of the coefficients of the fuzzy models. The constraints

$$Z^W_{j0}, A^W_{jr}, B^W_{js}, Z'^W_{i0}, A'^W_{ir}, B'^W_{is} \ge 0$$

ensure that the spread of a fuzzy coefficient is non-negative.

Now, we introduce two linear regression models for determining the symmetric triangular fuzzy coefficients of the linear fuzzy model $\mathcal{F}_j$ and $\mathcal{G}_i$. The first model is based on the idea used in [25]. The parameters are determined by solving a linear programming problem with an objective function of minimizing the total spread of the fuzzy coefficients:

Fuzzy-regression for target-environment data (FR 1)

$$\text{Minimize} \quad \sum_{j=1}^{n}\Big(Z_{j0}^{W}+\sum_{r=1}^{n}A_{jr}^{W}+\sum_{s=1}^{m}B_{js}^{W}\Big)+\sum_{i=1}^{m}\Big(Z_{i0}^{\prime W}+\sum_{r=1}^{n}A_{ir}^{\prime W}+\sum_{s=1}^{m}B_{is}^{\prime W}\Big),$$

$$\begin{aligned}
\text{subject to} \quad & \mathcal{F}_{j\alpha}^{L}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\le \overline{X}_j^{(\kappa+1)}\le \mathcal{F}_{j\alpha}^{R}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right),\\
& \mathcal{G}_{i\alpha'}^{L}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\le \overline{E}_i^{(\kappa+1)}\le \mathcal{G}_{i\alpha'}^{R}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\\
& (j=1,\ldots,n;\ i=1,\ldots,m;\ \kappa=0,1,\ldots,T),\\
& Z_{j0}^{W},\ Z_{i0}^{\prime W}\ \ge 0,\\
& A_{jr}^{W},\ A_{ir}^{\prime W}\ \ge 0\ (r=1,\ldots,n),\\
& B_{js}^{W},\ B_{is}^{\prime W}\ \ge 0\ (s=1,\ldots,m)\\
& (j=1,\ldots,n;\ i=1,\ldots,m).
\end{aligned}$$

Other objective functions for fuzzy regression are given in the literature. For example, the total spread of the fuzzy outputs can be used to define an alternative objective function ($cf.$ [9,24,25,27]). In our model, such kind of objective functions are given by

$$\sum_{\kappa=0}^{T}\left\{\sum_{j=0}^{n}\mathcal{F}_j^{W}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)+\sum_{i=0}^{m}\mathcal{G}_i^{W}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\right\},$$

and we obtain the following regression problem:

Fuzzy-regression for target-environment data (FR 2)

$$\begin{aligned}
\text{Minimize} \quad & \sum_{\kappa=0}^{T}\Bigg\{\sum_{j=1}^{n}\Big(Z_{j0}^{W}+\sum_{r=1}^{n}A_{jr}^{W}\cdot\left|\overline{X}_r^{(\kappa)}\right|+\sum_{s=1}^{m}B_{js}^{W}\cdot\left|\overline{E}_s^{(\kappa)}\right|\Big)\\
& \qquad +\sum_{i=1}^{m}\Big(Z_{i0}^{\prime W}+\sum_{r=1}^{n}A_{ir}^{\prime W}\cdot\left|\overline{X}_r^{(\kappa)}\right|+\sum_{s=1}^{m}B_{is}^{\prime W}\cdot\left|\overline{E}_s^{(\kappa)}\right|\Big)\Bigg\}
\end{aligned}$$

$$\begin{aligned}
\text{subject to} \quad & \mathcal{F}_{j\alpha}^{L}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\le \overline{X}_j^{(\kappa+1)}\le \mathcal{F}_{j\alpha}^{R}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right),\\
& \mathcal{G}_{i\alpha'}^{L}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\le \overline{E}_i^{(\kappa+1)}\le \mathcal{G}_{i\alpha'}^{R}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\\
& (j=1,\ldots,n;\ i=1,\ldots,m;\ \kappa=0,1,\ldots,T),\\
& Z_{j0}^{W},\ Z_{i0}^{\prime W}\ \ge 0,\\
& A_{jr}^{W},\ A_{ir}^{\prime W}\ \ge 0,(r=1,\ldots,n),\\
& B_{js}^{W},\ B_{is}^{\prime W}\ \ge 0,(s=1,\ldots,m)\\
& (j=1,\ldots,n;\ i=1,\ldots,m).
\end{aligned}$$

3.2. Fuzzy regression based on symmetric triangular fuzzy coefficients with membership grades

Data sets obtained by experiments (*e.g.*, microarray data) and environmental measurements are always affected by noise and uncertainty. In a preprocessing step, a statistical analysis of the measurement values can be performed in order to guarantee the quality of the observed data. In particular, outliers have to be detected and deleted from the sample. However, it is not always possible to split this sample unambiguously. For this reason membership grades $\alpha_{j\kappa}, \alpha'_{i\kappa} \in (0,1]$ are assigned to the data sets

$$\left(\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)^T; \overline{X}_j^{(\kappa+1)}\right) \quad \text{and} \quad \left(\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)^T; \overline{E}_i^{(\kappa+1)}\right) \quad (\kappa = 0, 1, \ldots, T).$$

When we include the membership grades in the objective function and the inclusion relations of the linear fuzzy regression model (FR2), we obtain the following method:

Fuzzy-regression for target-environment data (FR 3)

Minimize
$$\sum_{\kappa=0}^{T} \left\{ \sum_{j=1}^{n} \alpha_{j\kappa} \cdot \left(Z_{j0}^W + \sum_{r=1}^{n} A_{jr}^W \cdot \left|\overline{X}_r^{(\kappa)}\right| + \sum_{s=1}^{m} B_{js}^W \cdot \left|\overline{E}_s^{(\kappa)}\right| \right) + \sum_{i=1}^{m} \alpha'_{i\kappa} \cdot \left(Z_{i0}'^W + \sum_{r=1}^{n} A_{ir}'^W \cdot \left|\overline{X}_r^{(\kappa)}\right| + \sum_{s=1}^{m} B_{is}'^W \cdot \left|\overline{E}_s^{(\kappa)}\right| \right) \right\}$$

subject to
$$\mathcal{F}^L_{j\,\alpha_{j\kappa}}\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) \le \overline{X}_j^{(\kappa+1)} \le \mathcal{F}^R_{j\,\alpha_{j\kappa}}\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right),$$
$$\mathcal{G}^L_{i\,\alpha'_{i\kappa}}\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) \le \overline{E}_i^{(\kappa+1)} \le \mathcal{G}^R_{i\,\alpha'_{i\kappa}}\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)$$
$$(j = 1, \ldots, n;\ i = 1, \ldots, m;\ \kappa = 0, 1, \ldots, T),$$

$$Z_{j0}^W,\ Z_{i0}'^W \ge 0,$$
$$A_{jr}^W,\ A_{ir}'^W \ge 0\ (r = 1, \ldots, n),$$
$$B_{js}^W,\ B_{is}'^W \ge 0\ (s = 1, \ldots, m)$$
$$(j = 1, \ldots, n;\ i = 1, \ldots, m).$$

3.3. Fuzzy regression based on asymmetric triangular fuzzy coefficients

Limitations of fuzzy regression models based on symmetric triangular fuzzy coefficients were pointed out in [9]. One major drawback is that obviously different data sets may lead to the same linear fuzzy model. This is due to the fact that extremal data points mainly determine the spread of the models $\mathcal{F}_j$ and $\mathcal{G}_i$. As linear fuzzy regression models with symmetric triangular fuzzy coefficients are not flexible enough to represent the difference between data sets, Ishibuchi and Nii proposed *asymmetric triangular* or *trapezoidal fuzzy coefficients* [9]. In this section, we adapt this approach for a regression analysis of target-environment data based on asymmetric triangular fuzzy numbers. An algorithm for trapezoidal fuzzy coefficients is presented in Section 3.4.

We now assume that the coefficients of the fuzzy regression model are *asymmetric triangular fuzzy coefficients* (*cf*. Fig. 4). Therefore, they can be represented in terms of their lower limit (L), center (C) and upper limit (U) as follows:

$$Z_{j0} = (Z_{j0}^L, Z_{j0}^C, Z_{j0}^U)^T, \quad A_{jr} = (A_{jr}^L, A_{jr}^C, A_{jr}^U)^T, \quad B_{js} = (B_{js}^L, B_{js}^C, B_{js}^U)^T,$$
$$Z'_{i0} = (Z_{i0}'^L, Z_{i0}'^C, Z_{i0}'^U)^T, \quad A'_{ir} = (A_{ir}'^L, A_{ir}'^C, A_{ir}'^U)^T, \quad B'_{is} = (B_{is}'^L, B_{is}'^C, B_{is}'^U)^T,$$

where $r = 1, \ldots, n$ and $s = 1, \ldots, m$.

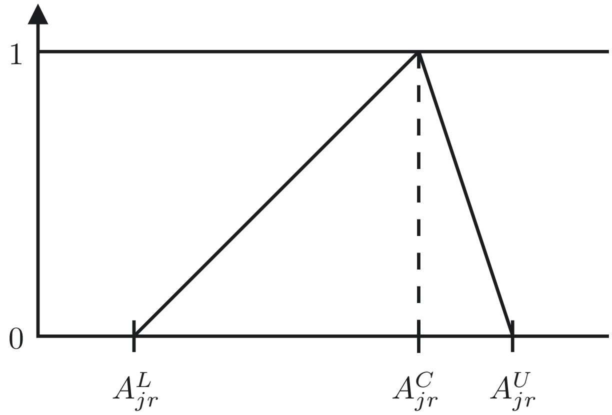

FIGURE 4. The asymmetric triangular fuzzy coefficient $A_{jr} = (A_{jr}^L, A_{jr}^C, A_{jr}^U)^T$.

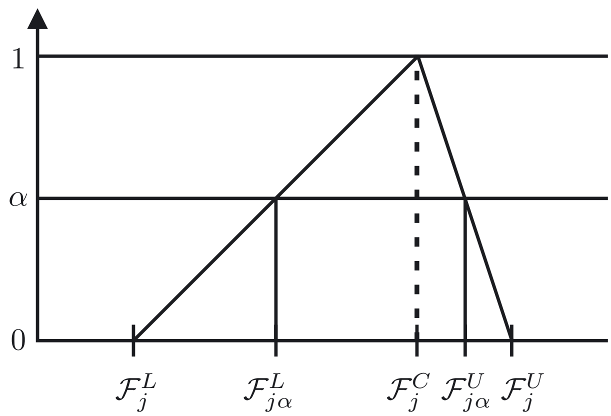

FIGURE 5. The α-cut of the asymmetric triangular fuzzy model $\mathcal{F}_j$.

When all the fuzzy coefficients are asymmetric triangular, the fuzzy models $\mathcal{F}_j$ and $\mathcal{G}_i$ are also asymmetric triangular fuzzy numbers (*cf*. Fig. 5). Therefore, $\mathcal{F}_j$ is given by

$$\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right) = \left(\mathcal{F}_j^L\left(X^{(k)}, E^{(k)}\right), \mathcal{F}_j^C\left(X^{(k)}, E^{(k)}\right), \mathcal{F}_j^U\left(X^{(k)}, E^{(k)}\right)\right)^T,$$

where

$$\begin{aligned}
\mathcal{F}_j^L\left(X^{(k)}, E^{(k)}\right) &= Z_{j0}^L + \sum_{r=1}^{n} \delta^L(X_r^{(k)})X_r^{(k)} + \sum_{s=1}^{m} \rho^L(E_s^{(k)})E_s^{(k)},\\
\mathcal{F}_j^C\left(X^{(k)}, E^{(k)}\right) &= Z_{j0}^C + \sum_{r=1}^{n} A_{jr}^C X_r^{(k)} + \sum_{s=1}^{m} B_{js}^C E_s^{(k)},\\
\mathcal{F}_j^U\left(X^{(k)}, E^{(k)}\right) &= Z_{j0}^U + \sum_{r=1}^{n} \delta^U(X_r^{(k)})X_r^{(k)} + \sum_{s=1}^{m} \rho^U(E_s^{(k)})E_s^{(k)}
\end{aligned}$$

with

$$\delta^L(X_r^{(k)}) = \begin{cases} A_{jr}^L, & \text{if } X_r^{(k)} \geq 0 \\ A_{jr}^U, & \text{if } X_r^{(k)} < 0 \end{cases}, \qquad \rho^L(E_s^{(k)}) = \begin{cases} B_{js}^L, & \text{if } E_s^{(k)} \geq 0 \\ B_{js}^U, & \text{if } E_s^{(k)} < 0 \end{cases},$$

and

$$\delta^U(X_r^{(k)}) = \begin{cases} A_{jr}^U, & \text{if } X_r^{(k)} \geq 0 \\ A_{jr}^L, & \text{if } X_r^{(k)} < 0 \end{cases}, \qquad \rho^U(E_s^{(k)}) = \begin{cases} B_{js}^U, & \text{if } E_s^{(k)} \geq 0 \\ B_{js}^L, & \text{if } E_s^{(k)} < 0 \end{cases}.$$

Similarly,

$$\mathcal{G}_i\left(X^{(k)}, E^{(k)}\right) = \left(\mathcal{G}_i^L\left(X^{(k)}, E^{(k)}\right), \mathcal{G}_i^C\left(X^{(k)}, E^{(k)}\right), \mathcal{G}_i^U\left(X^{(k)}, E^{(k)}\right)\right),$$

where

$$\begin{aligned}
\mathcal{G}_i^L\left(X^{(k)}, E^{(k)}\right) &= Z_{i0}'^L + \sum_{r=1}^{n} \delta'^L(X_r^{(k)}) X_r^{(k)} + \sum_{s=1}^{m} \rho'^L(E_s^{(k)}) E_s^{(k)}, \\
\mathcal{G}_i^C\left(X^{(k)}, E^{(k)}\right) &= Z_{i0}'^C + \sum_{r=1}^{n} A_{ir}'^C X_r^{(k)} + \sum_{s=1}^{m} B_{is}'^C E_s^{(k)}, \\
\mathcal{G}_i^U\left(X^{(k)}, E^{(k)}\right) &= Z_{i0}'^U + \sum_{r=1}^{n} \delta'^U(X_r^{(k)}) X_r^{(k)} + \sum_{s=1}^{m} \rho'^U(E_s^{(k)}) E_s^{(k)},
\end{aligned}$$

with

$$\delta'^L(X_r^{(k)}) = \begin{cases} A_{ir}'^L, & \text{if } X_r^{(k)} \geq 0 \\ A_{ir}'^U, & \text{if } X_r^{(k)} < 0 \end{cases}, \qquad \rho'^L(E_s^{(k)}) = \begin{cases} B_{is}'^L, & \text{if } E_s^{(k)} \geq 0 \\ B_{is}'^U, & \text{if } E_s^{(k)} < 0 \end{cases},$$

and

$$\delta'^U(X_r^{(k)}) = \begin{cases} A_{ir}'^U, & \text{if } X_r^{(k)} \geq 0 \\ A_{ir}'^L, & \text{if } X_r^{(k)} < 0 \end{cases}, \qquad \rho'^U(E_s^{(k)}) = \begin{cases} B_{is}'^U, & \text{if } E_s^{(k)} \geq 0 \\ B_{is}'^L, & \text{if } E_s^{(k)} < 0 \end{cases}.$$

The α-cut of

$$\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right) = \left(\mathcal{F}_j^L\left(X^{(k)}, E^{(k)}\right), \mathcal{F}_j^C\left(X^{(k)}, E^{(k)}\right), \mathcal{F}_j^U\left(X^{(k)}, E^{(k)}\right)\right)^T$$

is the interval

$$\left[\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right)\right]_\alpha = \left[\mathcal{F}_{j\alpha}^L\left(X^{(k)}, E^{(k)}\right), \mathcal{F}_{j\alpha}^U\left(X^{(k)}, E^{(k)}\right)\right],$$

where

$$\begin{aligned}
\mathcal{F}_{j\alpha}^L\left(X^{(k)}, E^{(k)}\right) &= \alpha \cdot \mathcal{F}_j^C\left(X^{(k)}, E^{(k)}\right) + (1-\alpha) \cdot \mathcal{F}_j^L\left(X^{(k)}, E^{(k)}\right), \\
\mathcal{F}_{j\alpha}^U\left(X^{(k)}, E^{(k)}\right) &= \alpha \cdot \mathcal{F}_j^C\left(X^{(k)}, E^{(k)}\right) + (1-\alpha) \cdot \mathcal{F}_j^U\left(X^{(k)}, E^{(k)}\right).
\end{aligned}$$

Similarly, the α'-cut of

$$\mathcal{G}_i\left(X^{(k)}, E^{(k)}\right) = \left(\mathcal{G}_i^L\left(X^{(k)}, E^{(k)}\right), \mathcal{G}_i^C\left(X^{(k)}, E^{(k)}\right), \mathcal{G}_i^U\left(X^{(k)}, E^{(k)}\right)\right)^T$$

is the interval

$$\left[\mathcal{G}_i\left(X^{(k)}, E^{(k)}\right)\right]_{\alpha'} = \left[\mathcal{G}_{i\alpha'}^L\left(X^{(k)}, E^{(k)}\right), \mathcal{G}_{i\alpha'}^U\left(X^{(k)}, E^{(k)}\right)\right],$$

where

$$\mathcal{G}^L_{i\alpha'}\left(X^{(k)},E^{(k)}\right)=\alpha'\cdot\mathcal{G}^C_i\left(X^{(k)},E^{(k)}\right)+(1-\alpha')\cdot\mathcal{G}^L_i\left(X^{(k)},E^{(k)}\right),$$
$$\mathcal{G}^U_{i\alpha'}\left(X^{(k)},E^{(k)}\right)=\alpha'\cdot\mathcal{G}^C_i\left(X^{(k)},E^{(k)}\right)+(1-\alpha')\cdot\mathcal{G}^U_i\left(X^{(k)},E^{(k)}\right).$$

In order to determine the centers as well as the upper and lower limits of the asymmetric triangular fuzzy coefficients, we adapt the following *hybrid method of least-squares regression and fuzzy regression* [9]:

Fuzzy-regression for target-environment data (FR 4)

(1) Apply least squares regression in order to determine the centers $\mathcal{F}^C_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$ and $\mathcal{G}^C_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$.

(2) Determine the lower limits $\mathcal{F}^L_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$, $\mathcal{G}^L_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$ and the upper limits $\mathcal{F}^U_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$, $\mathcal{G}^U_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$ by solving the following linear programming problem:

$$\text{Minimize}\quad \sum_{\kappa=0}^{T}\left\{\sum_{j=1}^{n}\left[\mathcal{F}^U_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)-\mathcal{F}^L_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\right]+\sum_{i=1}^{m}\left[\mathcal{G}^U_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)-\mathcal{G}^L_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\right]\right\}$$

$$\text{subject to}\quad \mathcal{F}^L_{j\alpha}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\le\overline{X}^{(\kappa+1)}_j\le\mathcal{F}^U_{j\alpha}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right),$$
$$\mathcal{G}^L_{i\alpha'}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\le\overline{E}^{(\kappa+1)}_i\le\mathcal{G}^U_{i\alpha'}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$$
$$(j=1,\ldots,n;\ i=1,\ldots,m;\ \kappa=0,1,\ldots,T),$$

$$Z^L_{j0}\le Z^C_{j0}\le Z^U_{j0},\ Z'^L_{i0}\le Z'^C_{i0}\le Z'^U_{i0},$$
$$A^L_{jr}\le A^C_{jr}\le A^U_{jr},\ A'^L_{ir}\le A'^C_{ir}\le A'^U_{ir},$$
$$B^L_{js}\le B^C_{js}\le B^U_{js},\ B'^L_{is}\le B'^C_{is}\le B'^U_{is}$$
$$(j,r=1,\ldots,n;\ i,s=1,\ldots,m;\ \kappa=0,1,\ldots,T).$$

In step (1), the centers of the fuzzy coefficients are determined while in step (2) the lower limits and upper limits of the asymmetric triangular fuzzy coefficients are calculated. The objective function is defined as the total spread of the fuzzy outputs from the linear fuzzy models $\mathcal{F}_j$ and $\mathcal{G}_i$, *i.e.*, the difference between the upper limit and lower limit of $\mathcal{F}_j$ and $\mathcal{G}_i$, respectively.

3.4. Fuzzy regression based on trapezoidal fuzzy coefficients

Fuzzy regression models with *asymmetric trapezoidal fuzzy coefficients* are proposed in [9] in order to reduce unnecessary fuzziness of the output and to avoid linear programming problems with no feasible solution (*cf.* Fig. 6). In this section, we will extend our model in this direction and we will use non-fuzzy data sets and asymmetric trapezoidal fuzzy coefficients.

Given the two crisp data sets

$$\left(\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)^T;\overline{X}^{(\kappa+1)}_j\right),\ \left(\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)^T;\overline{E}^{(\kappa+1)}_i\right)\quad(\kappa=0,1,\ldots,T),$$

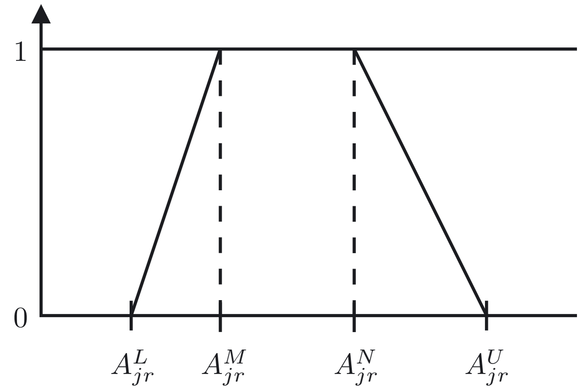

FIGURE 6. The asymmetric triangular fuzzy coefficient $A_{jr} = (A^L_{jr}, A^M_{jr}, A^N_{jr}, A^U_{jr})^T$.

we denote the coefficients as

$$\begin{aligned} Z_{j0} &= \left(Z^L_{j0}, Z^M_{j0}, Z^N_{j0}, Z^U_{j0}\right)^T, \quad Z'_{i0} = \left(Z'^L_{i0}, Z'^M_{i0}, Z'^N_{i0}, Z'^U_{i0}\right)^T, \\ A_{jr} &= \left(A^L_{jr}, A^M_{jr}, A^N_{jr}, A^U_{jr}\right)^T, \quad A'_{ir} = \left(A'^L_{ir}, A'^M_{ir}, A'^N_{ir}, A'^U_{ir}\right)^T, \\ B_{js} &= \left(B^L_{js}, B^M_{js}, B^N_{js}, B^U_{js}\right)^T, \quad B'_{is} = \left(B'^L_{is}, B'^M_{is}, B'^N_{is}, B'^U_{is}\right)^T, \end{aligned}$$

where $r = 1, \ldots, n$ and $s = 1, \ldots, m$.

The fuzzy models $\mathcal{F}_j$ and $\mathcal{G}_i$ are asymmetric trapezoidal fuzzy numbers. Therefore, $\mathcal{F}_j$ is given by:

$$\mathcal{F}_j\left(X^{(k)}, E^{(k)}\right) = \left(\mathcal{F}^L_j\left(X^{(k)}, E^{(k)}\right), \mathcal{F}^M_j\left(X^{(k)}, E^{(k)}\right), \mathcal{F}^N_j\left(X^{(k)}, E^{(k)}\right), \mathcal{F}^U_j\left(X^{(k)}, E^{(k)}\right)\right)^T,$$

where

$$\begin{aligned} \mathcal{F}^L_j\left(X^{(k)}, E^{(k)}\right) &= Z^L_{j0} + \sum_{r=1}^{n} \delta^L(X^{(k)}_r)X^{(k)}_r + \sum_{s=1}^{m} \rho^L(E^{(k)}_s)E^{(k)}_s, \\ \mathcal{F}^M_j\left(X^{(k)}, E^{(k)}\right) &= Z^M_{j0} + \sum_{r=1}^{n} \delta^M(X^{(k)}_r)X^{(k)}_r + \sum_{s=1}^{m} \rho^M(E^{(k)}_s)E^{(k)}_s, \\ \mathcal{F}^N_j\left(X^{(k)}, E^{(k)}\right) &= Z^N_{j0} + \sum_{r=1}^{n} \delta^N(X^{(k)}_r)X^{(k)}_r + \sum_{s=1}^{m} \rho^N(E^{(k)}_s)E^{(k)}_s, \\ \mathcal{F}^U_j\left(X^{(k)}, E^{(k)}\right) &= Z^U_{j0} + \sum_{r=1}^{n} \delta^U(X^{(k)}_r)X^{(k)}_r + \sum_{s=1}^{m} \rho^U(E^{(k)}_s)E^{(k)}_s, \end{aligned}$$

with

$$\delta^L(X_r^{(k)}) = \begin{cases} A_{jr}^L, & \text{if } X_r^{(k)} \geq 0 \\ A_{jr}^U, & \text{if } X_r^{(k)} < 0 \end{cases}, \quad \rho^L(E_s^{(k)}) = \begin{cases} B_{js}^L, & \text{if } E_s^{(k)} \geq 0 \\ B_{js}^U, & \text{if } E_s^{(k)} < 0 \end{cases},$$

$$\delta^M(X_r^{(k)}) = \begin{cases} A_{jr}^M, & \text{if } X_r^{(k)} \geq 0 \\ A_{jr}^N, & \text{if } X_r^{(k)} < 0 \end{cases}, \quad \rho^M(E_s^{(k)}) = \begin{cases} B_{js}^M, & \text{if } E_s^{(k)} \geq 0 \\ B_{js}^N, & \text{if } E_s^{(k)} < 0 \end{cases},$$

$$\delta^N(X_r^{(k)}) = \begin{cases} A_{jr}^N, & \text{if } X_r^{(k)} \geq 0 \\ A_{jr}^M, & \text{if } X_r^{(k)} < 0 \end{cases}, \quad \rho^M(E_s^{(k)}) = \begin{cases} B_{js}^N, & \text{if } E_s^{(k)} \geq 0 \\ B_{js}^M, & \text{if } E_s^{(k)} < 0 \end{cases},$$

$$\delta^U(X_r^{(k)}) = \begin{cases} A_{jr}^U, & \text{if } X_r^{(k)} \geq 0 \\ A_{jr}^L, & \text{if } X_r^{(k)} < 0 \end{cases}, \quad \rho^U(E_s^{(k)}) = \begin{cases} B_{js}^U, & \text{if } E_s^{(k)} \geq 0 \\ B_{js}^L, & \text{if } E_s^{(k)} < 0 \end{cases}.$$

Similarly,

$$\mathcal{G}_i\left(X^{(k)}, E^{(k)}\right) = \left(\mathcal{G}_i^L\left(X^{(k)}, E^{(k)}\right), \mathcal{G}_i^M\left(X^{(k)}, E^{(k)}\right), \mathcal{G}_i^N\left(X^{(k)}, E^{(k)}\right), \mathcal{G}_i^U\left(X^{(k)}, E^{(k)}\right)\right)^T,$$

where

$$\begin{aligned}
\mathcal{G}_i^L\left(X^{(k)}, E^{(k)}\right) &= Z_{i0}'^L + \sum_{r=1}^{n} \delta'^L(X_r^{(k)}) X_r^{(k)} + \sum_{s=1}^{m} \rho'^L(E_s^{(k)}) E_s^{(k)}, \\
\mathcal{G}_i^M\left(X^{(k)}, E^{(k)}\right) &= Z_{i0}'^M + \sum_{r=1}^{n} \delta'^M(X_r^{(k)}) X_r^{(k)} + \sum_{s=1}^{m} \rho'^M(E_s^{(k)}) E_s^{(k)}, \\
\mathcal{G}_i^N\left(X^{(k)}, E^{(k)}\right) &= Z_{i0}'^N + \sum_{r=1}^{n} \delta'^N(X_r^{(k)}) X_r^{(k)} + \sum_{s=1}^{m} \rho'^N(E_s^{(k)}) E_s^{(k)}, \\
\mathcal{G}_i^U\left(X^{(k)}, E^{(k)}\right) &= Z_{i0}'^U + \sum_{r=1}^{n} \delta'^U(X_r^{(k)}) X_r^{(k)} + \sum_{s=1}^{m} \rho'^U(E_s^{(k)}) E_s^{(k)},
\end{aligned}$$

with

$$\delta'^L(X_r^{(k)}) = \begin{cases} A_{ir}'^L, & \text{if } X_r^{(k)} \geq 0 \\ A_{ir}'^U, & \text{if } X_r^{(k)} < 0 \end{cases}, \quad \rho'^L(E_s^{(k)}) = \begin{cases} B_{is}'^L, & \text{if } E_s^{(k)} \geq 0 \\ B_{is}'^U, & \text{if } E_s^{(k)} < 0 \end{cases},$$

$$\delta'^M(X_r^{(k)}) = \begin{cases} A_{ir}'^M, & \text{if } X_r^{(k)} \geq 0 \\ A_{ir}'^N, & \text{if } X_r^{(k)} < 0 \end{cases}, \quad \rho'^M(E_s^{(k)}) = \begin{cases} B_{is}'^M, & \text{if } E_s^{(k)} \geq 0 \\ B_{is}'^N, & \text{if } E_s^{(k)} < 0 \end{cases},$$

$$\delta'^N(X_r^{(k)}) = \begin{cases} A_{ir}'^N, & \text{if } X_r^{(k)} \geq 0 \\ A_{ir}'^M, & \text{if } X_r^{(k)} < 0 \end{cases}, \quad \rho'^N(E_s^{(k)}) = \begin{cases} B_{is}'^N, & \text{if } E_s^{(k)} \geq 0 \\ B_{is}'^M, & \text{if } E_s^{(k)} < 0 \end{cases},$$

$$\delta'^U(X_r^{(k)}) = \begin{cases} A_{ir}'^U, & \text{if } X_r^{(k)} \geq 0 \\ A_{ir}'^L, & \text{if } X_r^{(k)} < 0 \end{cases}, \quad \rho'^U(E_s^{(k)}) = \begin{cases} B_{is}'^U, & \text{if } E_s^{(k)} \geq 0 \\ B_{is}'^L, & \text{if } E_s^{(k)} < 0 \end{cases}.$$

The α-cut of $\mathcal{F}_j\left(X^{(k)},E^{(k)}\right)$ is the interval

$$\left[\mathcal{F}_j\left(X^{(k)},E^{(k)}\right)\right]_\alpha=\left[\mathcal{F}^L_{j\alpha}\left(X^{(k)},E^{(k)}\right),\,\mathcal{F}^R_{j\alpha}\left(X^{(k)},E^{(k)}\right)\right],$$

where

$$\mathcal{F}^L_{j\alpha}\left(X^{(k)},E^{(k)}\right)=\alpha\cdot\mathcal{F}^M_j\left(X^{(k)},E^{(k)}\right)+(1-\alpha)\cdot\mathcal{F}^L_j\left(X^{(k)},E^{(k)}\right),$$
$$\mathcal{F}^R_{j\alpha}\left(X^{(k)},E^{(k)}\right)=\alpha\cdot\mathcal{F}^N_j\left(X^{(k)},E^{(k)}\right)+(1-\alpha)\cdot\mathcal{F}^U_j\left(X^{(k)},E^{(k)}\right),$$

and the α'-cut of $\mathcal{G}_i\left(X^{(k)},E^{(k)}\right)$ is the interval

$$\left[\mathcal{G}_i\left(X^{(k)},E^{(k)}\right)\right]_{\alpha'}=\left[\mathcal{G}^L_{i\alpha'}\left(X^{(k)},E^{(k)}\right),\,\mathcal{G}^R_{i\alpha'}\left(X^{(k)},E^{(k)}\right)\right],$$

where

$$\mathcal{G}^L_{i\alpha'}\left(X^{(k)},E^{(k)}\right)=\alpha'\cdot\mathcal{G}^M_i\left(X^{(k)},E^{(k)}\right)+(1-\alpha')\cdot\mathcal{G}^L_i\left(X^{(k)},E^{(k)}\right),$$
$$\mathcal{G}^R_{i\alpha'}\left(X^{(k)},E^{(k)}\right)=\alpha'\cdot\mathcal{G}^N_i\left(X^{(k)},E^{(k)}\right)+(1-\alpha')\cdot\mathcal{G}^U_i\left(X^{(k)},E^{(k)}\right).$$

Now, we can state the fuzzy regression model for non-fuzzy target-environment data with asymmetric trapezoidal fuzzy coefficients. In the objective function, we minimize the sum of total spread and inner spread of the fuzzy models which is given by

$$\sum_{j=1}^{n}\left[(\mathcal{F}^U_j-\mathcal{F}^L_j)+(\mathcal{F}^N_j-\mathcal{F}^M_j)\right]$$

and

$$\sum_{i=1}^{m}\left[(\mathcal{G}^U_i-\mathcal{G}^L_i)+(\mathcal{G}^N_i-\mathcal{G}^M_i)\right],$$

respectively. Beside the inclusion relations, we impose additional constraints in order to preserve the trapezoidal shape of the fuzzy coefficients:

Fuzzy-regression for target-environment data (FR 5)

Minimize
$$\sum_{\kappa=0}^{T}\Bigg\{\sum_{j=1}^{n}\Big[\ \mathcal{F}^U_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)-\mathcal{F}^L_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$$
$$+\mathcal{F}^N_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)-\mathcal{F}^M_j\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\Big]$$
$$+\sum_{i=1}^{m}\Big[\ \mathcal{G}^U_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)-\mathcal{G}^L_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$$
$$+\mathcal{G}^N_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)-\mathcal{G}^M_i\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\Big]\Bigg\}$$

subject to
$$\mathcal{F}^L_{j\alpha}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\le\overline{X}^{(\kappa+1)}_j\le\mathcal{F}^R_{j\alpha}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right),$$
$$\mathcal{G}^L_{i\alpha'}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)\le\overline{E}^{(\kappa+1)}_i\le\mathcal{G}^R_{i\alpha'}\left(\overline{\mathbb{X}}^{(\kappa)},\overline{\mathbb{E}}^{(\kappa)}\right)$$
$$(j=1,\dots,n;\ i=1,\dots,m;\ \kappa=0,1,\dots,T),$$
$$Z^L_{j0}\le Z^M_{j0}\le Z^N_{j0}\le Z^U_{j0},\ Z'^L_{i0}\le Z'^M_{i0}\le Z'^N_{i0}\le Z'^U_{i0},$$
$$A^L_{jr}\le A^M_{jr}\le A^N_{jr}\le A^U_{jr},\ A'^L_{ir}\le A'^M_{ir}\le A'^N_{ir}\le A'^U_{ir},$$
$$B^L_{js}\le B^M_{js}\le B^N_{js}\le B^U_{js},\ B'^L_{is}\le B'^M_{is}\le B'^N_{is}\le B'^U_{is}$$
$$(j,r=1,\dots,n;\ i,s=1,\dots,m;\ \kappa=0,1,\dots,T).$$

3.5. Fuzzy regression based on trapezoidal fuzzy coefficients with membership grades

In this section, we assume that individual membership grades $\alpha_{j\kappa}, \alpha'_{i\kappa} \in (0,1]$ are assigned to the data sets

$$\left(\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)^T; \overline{X}_j^{(\kappa+1)}\right) \quad \text{and} \quad \left(\left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)^T; \overline{E}_i^{(\kappa+1)}\right) \quad (\kappa = 0, 1, \ldots, T).$$

In this way, the quality of data obtained from a statistical analysis in a preprocessing step can also be reflected in the fuzzy regression with trapezoidal fuzzy coefficients. As in the case of symmetric triangular fuzzy coefficients in Section 3.2, the fuzzy regression model (FR5) with trapezoidal fuzzy coefficients can now be further extended and improved with regard to individual membership grades:

Fuzzy-regression for target-environment data (FR 6)

Minimize
$$\sum_{\kappa=0}^{T} \left\{ \sum_{j=1}^{n} \alpha_{j\kappa} \cdot \left[\mathcal{F}_j^U \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) - \mathcal{F}_j^L \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) + \mathcal{F}_j^N \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) - \mathcal{F}_j^M \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) \right] + \sum_{i=1}^{m} \alpha'_{i\kappa} \cdot \left[\mathcal{G}_i^U \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) - \mathcal{G}_i^L \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) + \mathcal{G}_i^N \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) - \mathcal{G}_i^M \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) \right] \right\}$$

subject to
$$\mathcal{F}_{j\alpha_{j\kappa}}^L \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) \le \overline{X}_j^{(\kappa+1)} \le \mathcal{F}_{j\alpha_{j\kappa}}^R \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right),$$
$$\mathcal{G}_{i\alpha'_{i\kappa}}^L \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right) \le \overline{E}_i^{(\kappa+1)} \le \mathcal{G}_{i\alpha'_{i\kappa}}^R \left(\overline{\mathbb{X}}^{(\kappa)}, \overline{\mathbb{E}}^{(\kappa)}\right)$$
$$(j = 1, \ldots, n;\ i = 1, \ldots, m;\ \kappa = 0, 1, \ldots, T),$$
$$Z_{j0}^L \le Z_{j0}^M \le Z_{j0}^N \le Z_{j0}^U,\ Z_{i0}'^L \le Z_{i0}'^M \le Z_{i0}'^N \le Z_{i0}'^U,$$
$$A_{jr}^L \le A_{jr}^M \le A_{jr}^N \le A_{jr}^U,\ A_{ir}'^L \le A_{ir}'^M \le A_{ir}'^N \le A_{ir}'^U,$$
$$B_{js}^L \le B_{js}^M \le B_{js}^N \le B_{js}^U,\ B_{is}'^L \le B_{is}'^M \le B_{is}'^N \le B_{is}'^U$$
$$(j, r = 1, \ldots, n;\ i, s = 1, \ldots, m;\ \kappa = 0, 1, \ldots, T).$$

Finally, Table 1 summarizes the regression models together with the corresponding type of coefficients and model outputs as well as the specific form of the objective function.

TABLE 1. Fuzzy regression algorithms for target-environment data.

Algorithm	Coefficients/Fuzzy Output	Objective Function (min.)
(FR1)	symmetric triangular	total spread of fuzzy coefficients
(FR2)	symmetric triangular	total spread of fuzzy model output
(FR3)	symmetric triangular	total spread of fuzzy model output with membership grades
(FR4)	asymmetric triangular	total spread of fuzzy model output
(FR5)	asymmetric trapezoidal	sum of total spread and inner spread of fuzzy model output
(FR6)	asymmetric trapezoidal	sum of total spread and inner spread of fuzzy model output with membership grades

4. Numerical example

As a first illustrative example a subset of a whole-genome mouse microarray data set [5] from Agilent Technologies (Santa Clara, CA) is chosen for illustration purposes. The data set was obtained after a single oral dose exposure to Tetrachlorodibenzo-*p*-dioxin (TCDD). Despite the large amount of genetic data available in many databases, pure gene-environment data has not been available previously. Typically, only a few number of expression values and environmental observations is available. A classical statistical regression requires a large number of measurements for a valid prediction. As discussed above, fuzzy regression approaches require a comparatively small number of measurements and they offer an alternative way for prediction of gene-environment regulatory systems. We intend to turn to an investigation of large data sets and a comparison of regression approaches in future studies.

4.1. Experimental data

The whole-genome microarray data that is used in our numerical application belongs to the hepatic tissue from immature, ovariectomized $C57BL/6$ mice, and the data set shows the changes in gene expression profiles observed at the time points 2, 4, 8, 12, 18, 24, 72 and 168 hours after TCDD treatment (see [5] for further details with references, and supplementary Table 4 for the complete data set).

TCDD (a kind of toxin) is an environmental factor that causes various kinds of species-specific effects (like tumor promotion, modulation of endocrine systems, hepatotoxicity, etc.) which are a result of changes in gene expression produced by a protein, the so called aryl hydrocarbon receptor (AhR). The AhR is a transcription factor from a specific kind of family of proteins that work as environmental sensors to different stimuli (see [5] and references therein).

In the experimental design of the mentioned data set in [5], the genome-wide effects of TCDD is analyzed by detecting the changes (or regulations) in gene expressions caused by a specific transcription factor AhR in different tissues of various species such as human, mouse, rat and mice.

Tables 2 and 3 list the information and expression values of totally 16 genes that we select as a subset of the whole-genome $C57BL/6$ Mice TCDD temporal microarray data given in [5]. This subnetwork of 16 genes contains 11 target genes and 5 genes who targeted them, therefore it is considered here as a target-environment network of totally 16 genes that we studied for our real-world application and corresponding numerical results.

4.2. Fuzzy prediction

As an illustrative example of the general prediction approaches described above we focus on the target gene Tom1l2 (ID 176687) and an environmental factor given by gene Myc (ID 145739). Figures 7–10 show the predictions of the corresponding expression values for seven samplings. The numerical results illustrate the qualitative behaviour of the fuzzy approaches discussed in the previous sections. As we can expect from our theoretical considerations, the fuzzy regression approaches (FR1) and (FR2) that are based on symmetric triangular fuzzy numbers result in less precise predictions (see Fig. 7). This is because the extremal values govern the prediction.

Much better results can be achieved with the hybrid method of least squares regression and fuzzy regression (FR4). For the genetic data from Table 3 the prediction of the expression values shows nearly no deviation and center, center-width and center+width are nearly identical (see Fig. 8, left image). In general, fuzzy regression based on trapezoidal coefficients is considered as more flexible. With the fuzzy regression method (FR5) we obtain again much better predictions for the data set under consideration (see Fig. 8, right image).

We obtain a similar picture for the environmental factor gene Myc (ID 145739). Again, the fuzzy regression strategies (FR1) and (FR2) show a low prediction quality (see Fig. 9).

The hybrid method method (FR4) that is based on asymmetric triangular fuzzy numbers, and the fuzzy approach (FR5) lead again to superior prediction results (see Fig. 10).

The examples above illustrate the qualitative behaviour of the various fuzzy regression models for the target-environment networks introduced in this paper. General fuzzy regression theory already indicates that fuzzy

TABLE 2. A part of $C57BL/6$ Mice TCDD temporal microarray data (description of the genes in the selected subnetwork, see supplementary Table 4 in [5]).

CloneID	Probe	Entrez GeneID	Gene symbol	Gene name
150701	A51P202162	216810	Tom1l2	target of myb1-like 2
154262	A51P272646	21968	Tom1	target of myb1 homolog
155547	A51P299866	71943	Tom1l1	target of myb1-like 1
155980	A51P308814	68276	Toe1	target of EGR1, member 1
167060	A52P105913	216810	Tom1l2	target of myb1-like 2
168777	A52P13389	68632	Myct1	myc target 1
169219	A52P1480	21968	Tom1	target of myb1 homolog
176687	A52P383753	216810	Tom1l2	target of myb1-like 2
179423	A52P469458	68276	Toe1	target of EGR1, member 1
183011	A52P583563	71943	Tom1l1	target of myb1-like 1
184424	A52P630646	68276	Toe1	target of EGR1, member 1
158892	A51P367866	13653	Egr1	early growth response 1
145739	A51P102096	17869	Myc	myelocytomatosis oncogene
149613	A51P180761	17864	Mybl1	myeloblastosis oncogene-like 1
158047	A51P351144	17865	Mybl2	myeloblastosis oncogene-like 2
183171	A52P589622	17864	Mybl1	myeloblastosis oncogene-like 1

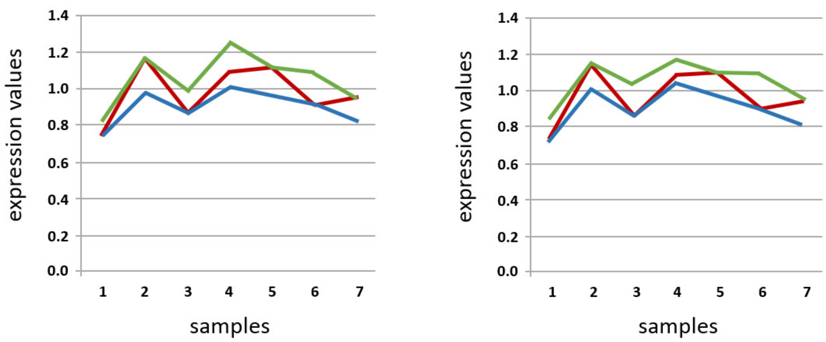

FIGURE 7. Expression values of gene Tom1l2 (ID 176687): Fuzzy regression results obtained with (FR1) – *left image* – and (FR2) – *right image*. The red line depicts the central value and the blue and green lines are center-width and center+width, respectively (Color online).

hybrid models and regression approaches based on asymmetric fuzzy numbers provide better results than linear models based on symmetric parameters. This is also the case in the numerical example presented above. The obtained results show that fuzzy approaches are applicable to gene-environment data. Compared to other prediction approaches like neural networks, the parameters of the fuzzy regression models can be even determined when only a few number of measurements is available. This is an advantage in the current situation where gene-environment data is only available with a small number of samples. Typically, the number of genes and environmental factors in gene-environment data is much higher than the number of samples. Whenever predictions have to made on the basis of small data sets where imprecise data becomes involved, fuzzy regression can support the prediction. This even more important, when each gene and environmental factor can be monitored several times under varying conditions. Nevertheless, for a more detailed study and a comparison with other approaches, larger gene-environment data sets are required.

TABLE 3. A part of $C57BL/6$ Mice TCDD temporal microarray data. CloneIDs, genes and expression values in the selected subnetwork given in Table 2, see supplementary Table 4 in [5].

CloneID	Gene S.	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6	sample 7	sample 8
150701	Tom1l2	0.94461155	0.97502571	1.01504505	1.04515910	0.92826730	1.01328480	0.92135280	0.92759496
154262	Tom1	0.95806974	0.96664912	0.93885249	0.8489846	0.83275616	0.82753015	0.94185579	0.75484133
155547	Tom1l1	1.06827879	1.32651401	1.19834495	1.04832494	1.14479709	1.14369571	1.09890068	1.10287309
155980	Toe1	0.93451422	1.00863683	1.09234762	1.02561224	1.05001426	1.06493783	1.03766215	1.04626155
167060	Tom1l2	0.97148854	0.98046839	1.03474474	0.98632944	0.93387341	1.02215636	1.02764237	0.94588989
168777	Myct1	0.96981782	0.88317215	1.08502686	1.00037587	1.07871532	1.00419497	1.00730658	1.14439738
169219	Tom1	0.86958379	1.03319240	1.17927861	0.83565915	0.93835098	1.32708383	0.94818872	0.90235573
176687	Tom1l2	0.89905554	0.75318289	1.16383362	0.86540061	1.08676589	1.10582566	0.90585828	0.94663352
179423	Toe1	0.96205670	0.94183278	1.05489385	0.98735642	1.13256562	1.21424377	1.07656944	1.0033170
183011	Tom1l1	0.96602547	0.95331699	1.11521018	0.89680165	1.18518376	0.90491748	0.93765175	0.99103117
184424	Toe1	0.90346366	0.99773824	1.08958066	0.98045099	1.14603531	1.01464808	1.08568466	1.08008397
158892	Egr1	2.58529305	0.1958821	0.73433846	0.75315076	1.84934735	0.81747842	1.88940322	1.9654249
145739	Myc	1.14516723	2.7764957	1.64422393	3.49538755	2.14031959	2.60398889	1.83789301	2.63181448
149613	Mybl1	1.00000000	1.00000000	1.00000000	1.00000000	1.0000000	1.00000000	1.00000000	1.00000000
158047	Mybl2	0.92232084	0.92638117	1.06035864	0.97620231	1.07358599	0.92821044	1.22258306	1.04365885
183171	Mybl1	1.00000000	1.00000000	1.00000000	0.87805402	0.94255263	1.00000000	1.00000000	1.04507864

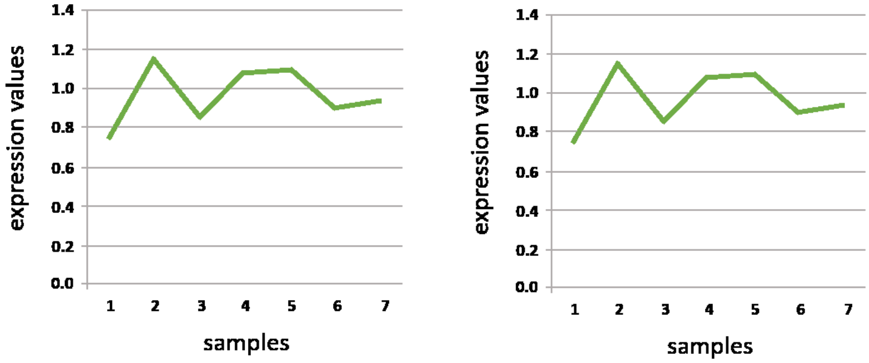

Figure 8. Expression values of gene Tom1l2 (ID 176687): Fuzzy regression results obtained with (FR4), *left image*, and (FR5), *right image*. Center, center-width and center+width are nearly identical.

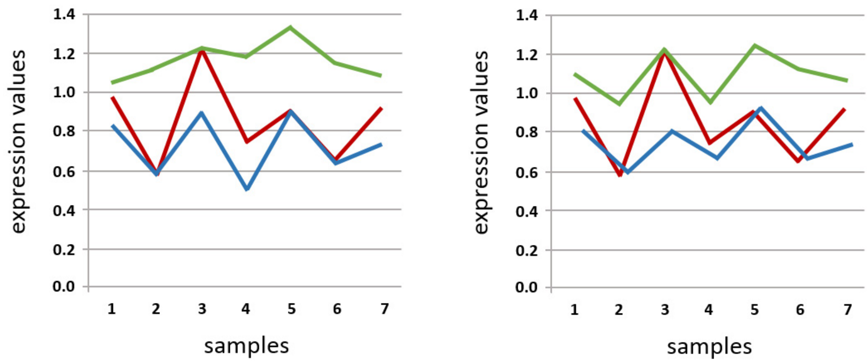

Figure 9. Expression values of environmental factor Myc (145739): Fuzzy regression results obtained with (FR1) – *left image* – and (FR2) – *right image*. The red line depicts the central value and the blue and green lines are center-width and center+width, respectively.

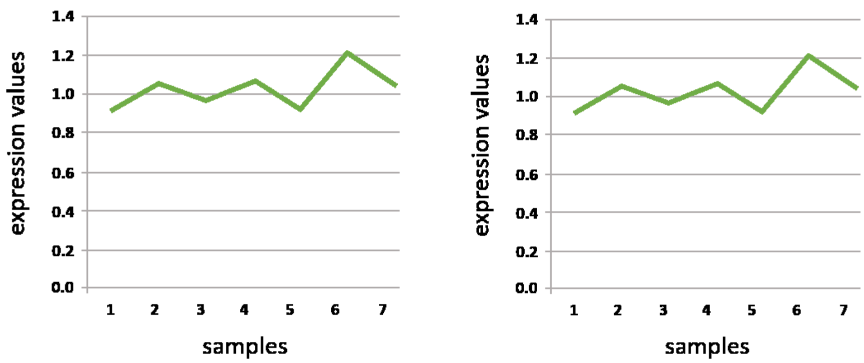

Figure 10. Expression values of environmental factor Myc (145739): Fuzzy regression results obtained with (FR4), *left image*, and (FR5), *right image*. Center, center-width and center+width are nearly identical.

5. Conclusion

The objective of this paper is to introduce fuzzy target-environment networks and fuzzy evolving networks as further approaches for the analysis of two-modal regulatory systems affected by errors and uncertainty. The proposed method is based on a fuzzy model with fuzzy coefficients. Depending on the shape of these uncertain parameters, various possibilistic regression models are obtained. In future works, methods from fuzzy least-squares regression based on a minimization of the total square error of the output can be addressed [6]. In addition, the regression models can be coupled with different types of fuzzy input vectors. Beside the crisp input from measurements also fuzzy input data can be considered in the proposed algorithms which is of particular importance with regard to applications in case of critical operations. For an analysis of nonlinear systems, fuzzy neural networks approaches can be adapted to the bi-level situation of target-environment networks [9]. A further direction of research could discuss the parameter identification of regulatory systems with interacting groups of variables affected by fuzzy uncertainty. Such an approach could be based on the set-theoretic regression analysis of [14–16], where functionally related groups of targets and environmental entities under ellipsoidal uncertainty are considered. In general, fuzzy regression approaches are very flexible and can be adapted to a variety of regulatory systems where data uncertainty and model restrictions are involved. This is usually the case for interdependent networks in systems biology where the regulating effects are not known in detail and one has to refer to approximating models for a representation of the interaction at the system level. In such a situation, fuzzy regression strategies can be applied in order to evaluate the prediction quality of the model and to compare its results with real-world observations and measurements. The fuzzy regression approaches for target-environment networks discussed in this paper are not restricted to biological problems. They can be applied to any interdependent system, in particular when data uncertainty becomes involved. In addition, fuzzy models can to some extend model the inherent uncertainty of a system as they depend on uncertain fuzzy parameters. This can support the modeling of any biological, economic or technical system where the regulatory mechanisms are not completely understood and only a small number of uncertain measurements are available.

References

[1] C. Carleos, F. Rodriguez, H. Lamelas and J.A. Baro, Simulating complex traits influenced by genes with fuzzy-valued effects in pedigreed populations. *Bioinform.* **19** (2003) 144–148.

[2] Y.-H.O. Chang and B.M. Ayyub, Fuzzy regression methods – a comparative assessment. *Fuzzy Sets Syst.* **119** (2001) 187–203.

[3] S. Charfeddine, K. Zbidi and F. Mora-Camino, Fuzzy Regression Analysis Using Trapezoidal Fuzzy Numbers. In *Proc. of the Joint 4th Conference of the European Society for Fuzzy Logic and Technology and the 11th Rencontres Francophones sur la Logique Floue et ses Applications*, edited by E. Montseny, P. Sobrevilla. Barcelona, Spain (2005).

[4] O. Defterli, A. Fügenschuh and G.-W. Weber, Modern tools for the time-discrete dynamics and optimization of gene-environment networks. *Commun. Nonlinear Sci. Numer. Simul.* **16** (2011) 4768–4779.

[5] E. Dere, A.L. Forgacs, T.R. Zacharewski and L.D. Burgoon, Genome-wide computational analysis of dioxin response element location and distribution in the human, mouse, and rat genomes. *Chem. Res. Toxicol.* **24** (2011) 494–504.

[6] P. Diamond, Fuzzy least squares source. *Inform. Sci.* **46** (1988) 141–157.

[7] J. Gebert, M. Lätsch, E.M.P. Quek and G.-W. Weber, Analyzing and optimizing genetic network structure via path-finding. *J. Comput. Technol.* **9** (2004) 3–12.

[8] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning. Springer (2001).

[9] H. Ishibuchi and M. Nii, Fuzzy regression using asymmetric fuzzy coefficients and fuzzified neural networks. *Fuzzy Sets Syst.* **199** (2001) 273–290.

[10] C. Kahraman, A. Beskese and F. Tunc Bozbura, Fuzzy Regression Approaches and Applications. In Fuzzy Applications in Industrial Engineering. Vol. 201 of *Studies in Fuzziness and Soft Computing*. Springer, Berlin, Heidelberg (2006), 589–615.

[11] G.J. Klir and B. Yuan, Fuzzy Sets and Fuzzy Logic. Theory and Applications. Prentice Hall, New Jersey (1995).

[12] E. Kropat, Fuzzy-Unabhängigkeitssysteme und Fuzzy-Matroide. Wissenschaftlicher Verlag Berlin, Berlin (2002).

[13] E. Kropat, G.-W. Weber and B. Akteke-Öztürk, Eco-finance Networks Under Uncertainty. In *Proc. of the International Conference on Engineering Optimization, EngOpt 2008*, edited by J. Herskovits, A. Canelas, H. Cortes, M. Aroztegui. Rio de Janeiro, Brazil, 1-5 June 2008.

[14] E. Kropat, G.-W. Weber and J.-J. Rückmann, Regression analysis for clusters in gene-environment networks based on ellipsoidal calculus and optimization. *Dyn. Contin. Discrete Impulsive Syst. Ser. B* **17** (2010) 639–657.

[15] E. Kropat, G.-W. Weber and S. Belen, Dynamical Gene-environment Networks Under Ellipsoidal Uncertainty - Set-theoretic Regression Analysis Based on Ellipsoidal OR. In Vol. 1 of Dynamics, Games and Science I. *Springer Proc. Math.*, edited by M.M. Peixoto, A.A. Pinto, D.A. Rand. Springer-Verlag, Berlin, Heidelberg (2011) 545–571.

[16] E. Kropat, G.-W. Weber and C.S. Pedamallu, Regulatory Networks Under Ellipsoidal Uncertainty - Data Analysis and Prediction by Optimization Theory and Dynamical Systems. In Data Mining: Foundations and Intelligent Paradigms, edited by D.E. Holmes, L.S. Jain. Vol. 2 of *Statistical, Bayesian, Time Series and other Theoretical Aspects, ISRL 24*. Springer-Verlag, Berlin (2012) 27–56.

[17] E. Kropat, G.-W. Weber, S.Z. Alparslan-Gök and A. Özmen, Inverse Problems in Complex Multi-modal Regulatory Networks Based on Uncertain Clustered Data. In Modeling, Optimization, Dynamics and Bioeconomy, edited by A. Pinto and D. Zilberman. Springer (2014) 437–451.

[18] D.S. Malik and J.N. Mordeson, Fuzzy Discrete Structures. Physica-Verlag, Heidelberg (2000).

[19] S. Meyer–Nieberg and E. Kropat, Tracking Targets under Uncertainty: Natural Computing Approaches. In *Proc. of the 47th Hawaii International Conference on Systems Sciences, January 6-9, 2014*. Waikoloa, Big Island, Hawaii (2014) 1162–1171.

[20] G. Peters, Fuzzy linear regression with fuzzy intervals. *Fuzzy Sets Syst.* **63** (1994) 45–55.

[21] D.T. Redden and W.H. Woodall, Properties of certain fuzzy linear regression methods. *Fuzzy Sets Syst.* **64** (1994) 361–375.

[22] M. Sakawa and H. Yano, Multiobjective fuzzy linear regression analysis for fuzzy input-output data. *Fuzzy Sets Syst.* **47** (1992) 173–181.

[23] D.A. Savic and W. Pedrycz, Evaluation of fuzzy linear regression models. *Fuzzy Sets Syst.* **39** (1991) 51–63.

[24] H. Tanaka and H. Ishibuchi, Possibilistic Regression Analysis Based on Linear Programming. In: Fuzzy Regression Analysis, edited by J. Kacprzyk, M. Fedrizzi. Physica-Verlag, Heidelberg (1992) 47–60.

[25] H. Tanaka, K. Uejima and K. Asai, Linear regression analysis with fuzzy model. *IEEE Systems Trans. Systems Man Cybernet.* **12** (1982) 903–907.

[26] H. Tanaka, I. Hayashi and J. Watada, Possibilistic linear regression analysis for fuzzy data. *Eur. J. Oper. Res.* **40** (1989) 389–396.

[27] H. Tanaka, H. Ishibuchi and S.G. Hwang, Fuzzy model of the number of staff in local government by fuzzy regression analysis with similarity relations. *J. Jpn Indust. Management Assoc.* **41** (1990) 99–104.

[28] M. Taştan, *Analysis and prediction of gene expression patterns by dynamical systems, and by a combinatorial algorithm*. MSc thesis, Institute of Applied Mathematics, METU, Turkey (2005).

[29] Ö. Uğur and G.-W. Weber, Optimization and dynamics of gene-environment networks with intervals, in the Special Issue at the occasion of the 5th Ballarat Workshop on Global and Non-Smooth Optimization: Theory, Methods and Applications, November 28–30, 2006. *J. Ind. Manag. Optim.* **3** (2007) 357–379.

[30] G.-W. Weber, Charakterisierung struktureller Stabilität in der nichtlinearen Optimierung. In *Aachener Beiträge zur Mathematik 5*, edited by H.H. Bock, H.T. Jongen, and W. Plesken. Augustinus publishing house (now: Mainz publishing house), Aachen (1992).

[31] G.-W. Weber, Generalized semi-infinite optimization and related topics, edited by K.H. Hofmannn and R. Wille. In vol. 29 of *Research and Exposition in Mathematics*. Lemgo, Heldermann Publishing House (2003)

[32] G.-W. Weber, S.Z. Alparslan-Gök and N. Dikmen, Environmental and life sciences: gene-environment networks – optimization, games and control – a survey on recent achievements. Invited paper. *J. Organisational Transformation and Social Change* **5** (2008) 197–233.

[33] G.-W. Weber, P. Taylan, S.-Z. Alparslan-Gök, S. Özöğür and B. Akteke-Öztürk, Optimization of gene-environment networks in the presence of errors and uncertainty with Chebychev approximation. *TOP Oper. Res. J. SEIO* (Spanish Statistics and Operations Research Society) **16** (2008) 284–318.

[34] G.-W. Weber, A. Tezel, P. Taylan, A. Soyler and M. Çetin, Mathematical contributions to dynamics and optimization of gene-environment networks. In Special Issue: In Celebration of Prof. Dr. Hubertus Th. Jongen's 60th Birthday, edited by D. Pallaschke, O. Stein. *J. Optim.* **57** (2008) 353–377.

[35] G.-W. Weber, S.Z. Alparslan-Gök and B. Söyler, A new mathematical approach in environmental and life sciences: gene-environment networks and their dynamics. *Environ. Model. Assess.* **14** (2009) 267–288.

[36] G.-W. Weber, E. Kropat, B. Akteke-Öztürk and Z.-K. Görgülü, A Survey on OR and Mathematical Methods Applied on Gene-Environment Networks. Special Issue on "Innovative Approaches for Decision Analysis in Energy, Health, and Life Sciences" at the occasion of *EURO XXII 2007*, Prague, Czech Republic, July 8–11 (2007). *Central Eur. J. Oper. Res.* **17** (2009) 315–341.

[37] G.-W. Weber, S. Özögür-Akyüz and E. Kropat, A review on data mining and continuous optimization applications in computational biology and medicine. *Embryo Today, Birth Defects Research C* **87** (2009) 165–181.

[38] G.-W. Weber, Ö. Uğur, P. Taylan and A. Tezel, On optimization, dynamics and uncertainty: A tutorial for gene-environment networks. In the Special Issue Networks in Computational Biology. *Discrete Appl. Math.* **157** (2009) 2494–2513.

[39] G.-W. Weber, E. Kropat, A. Tezel and S. Belen, Optimization applied on regulatory and eco-finance networks – survey and new developments. Special Issue in memory of Professor Alexander Rubinov, edited by M. Fukushima. *Pacific J. Optim.* **6** (2010) 319–340.

[40] G.-W. Weber, O. Defterli, S.Z. Alparslan-Gök and E. Kropat, Modeling, inference and optimization of regulatory networks based on time series data. *Eur. J. Oper. Res.* **211** 1–14 (2011)

[41] F.B. Yılmaz, H. Öktem and G.-W. Weber, Mathematical Modeling and Approximation of Gene Expression Patterns and Gene Networks. In *Operations Research Proceedings*, edited by F. Fleuren, D. den Hertog and P. Kort (2005) 280–287.