# Fuzzy logic as a Computational Tool for Quantitative Modelling of Biological Systems with Uncertain Kinetic Data

J. Bordon, M. Moškon, N. Zimic and M. Mraz

## **IEEE copyright notice:**

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

## Citation to the original IEEE publication:

J. Bordon, M. Moškon, N. Zimic and M. Mraz, **Fuzzy Logic as a Computational Tool for Quantitative Modelling of Biological Systems with Uncertain Kinetic Data**, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 12, No. 5, pp. 1199-1205, Sept.-Oct. 1 2015 (doi: 10.1109/TCBB.2015.2424424)

## Available at IEEE Xplore

http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7089201

# Fuzzy Logic as a Computational Tool for Quantitative Modelling of Biological Systems with Uncertain Kinetic Data

Jure Bordon, Miha Moškon, Nikolaj Zimic, Member, IEEE and Miha Mraz, Member, IEEE

Abstract—Quantitative modelling of biological systems has become an indispensable computational approach in the design of novel and analysis of existing biological systems. However, kinetic data that describe the system's dynamics need to be known in order to obtain relevant results with the conventional modelling techniques. These data are often hard or even impossible to obtain. Here we present a quantitative fuzzy logic modelling approach that is able to cope with unknown kinetic data and thus produce relevant results even though kinetic data are incomplete or only vaguely defined. Moreover, the approach can be used in the combination with the existing *state-of-the-art* quantitative modelling techniques only in certain parts of the system, i.e. where kinetic data are missing. The case study of the approach proposed here is performed on the model of 3-gene repressilator.

Index Terms—Fuzzy Logic, Uncertain Kinetic Data, Ordinary Differential Equations, Computational Biology, Gene Regulatory Networks, Modelling and Simulation, Synthetic Biology.

### I. INTRODUCTION

**R** ECENT advances in systems and synthetic biology have given detailed insight on the dynamics and structure of several biological systems. This knowledge has made the design and construction of novel biological systems with predefined functionalities more straightforward [1]–[3]. Among others, several synthetic gene regulatory networks (GRNs), such as genetic toggle-switches and oscillators, have caught the attention of the research community due to their occurrence in nature as well as their vast potential in different synthetic applications, e.g. bi-stable switch for gene therapy, repressilator, metabolator etc. [4]–[8]. However, experimental realization of these systems still presents a time-consuming and costly trial and error process.

Recently computational models present an indispensable tool that can be used for the design, optimization and *in silico* verification of a novel biological system before its experimental realization [9], [10]. Choosing an appropriate modelling technique depends on the complexity of the observed GRN, desired accuracy of simulation results and the availability of accurate kinetic data, which describe the dynamical properties of the system. Existing quantitative methods are mostly based on the numerical simulations of the system of ordinary differential equations (ODEs) or chemical master equation (CME). While describing systems' dynamics accurately, these approaches require accurate kinetic data in order to produce useful simulation results [11]–[14].

The dynamics of an arbitrary GRN can be roughly described with three different processes, i.e. transcription, translation and degradation. Each of these processes can be presented with at least one chemical reaction and its belonging kinetic rate(s). Kinetic rates can be sometimes (accurately) determined by using various parameter prediction and estimation techniques. If experimental data for a given biological system is available, these methods can estimate missing kinetic data, which can then be used in an ODE model [15]–[17]. However, experimental data are often hard or even impossible to obtain. In those cases parameter estimation techniques cannot be used and a different approach is needed.

In recent years fuzzy logic has been established as an alternative approach for the quantitative modelling of biological systems [18]. Fuzzy models consist of linguistic expressions (e.g. *Concentration* is *High* or *Promoter activity* is *Low*) and are straightforward to construct as well as easy to understand. When kinetic data are known the accuracy of fuzzy modelling approaches is equal to the existing deterministic approaches, such as ODE based models [19]. Moreover, existing fuzzy logic approaches can be used to obtain a qualitative response of the system's dynamics even though the kinetic data are unknown [20]. Uncertain kinetic data however still present a major obstacle for obtaining the quantitative response using existing modelling approaches [21].

Existing fuzzy logic approaches mostly consist of the key events descriptions only (e.g. gene activated or repressed, binding of a transcription factor probable or not etc.) [22], [23], but are as such unable to cope with the quantitative response of the system, such as protein concentration changes. On the other hand, Fuzzy Cognitive Maps can be constructed to describe a metabolic or gene regulatory network, but are used as a qualitative overview of the network (e.g. when are nodes actived/deactivated, how species interact with eacho ther over time, etc.) [24], [25]. Here we present a new approach that comprehensively exploits the advantages of fuzzy logic to obtain the quantitative simulation results. The approach is able to

The authors are with the Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia (e-mail: jure.bordon@fri.unilj.si; miha.moskon@fri.uni-lj.si; nikolaj.zimic@fri.uni-lj.si; miha.mraz@fri.uni-lj.si).

Manuscript received 22 Sep. 2014; revised 20 Mar. 2015.



Fig. 1. Processing the data with fuzzy logic. Crisp variables are fuzzified to their corresponding fuzzy values on which fuzzy rules are applied. Fuzzy rules produce output fuzzy variables, which are defuzzified to their corresponding crisp values.

quantitatively describe the behaviour of a certain biological system even though the kinetic data are uncertain or known only partially. In addition, the proposed method can be used in a combination with existing *state-of-the-art* quantitative modelling approaches only in the parts of the system, which are vaguely defined, i.e. where kinetic data are missing. We demonstrate the introduced approach on the establishment and analysis of a fuzzy model of the 3-gene repressilator [7].

Section II describes the application of fuzzy logic to biological systems modelling. Establishment of a fuzzy logic model of the 3-gene repressilator as an use-case is presented in Section III. Simulation results and their analysis are given in Section IV and concluding remarks in Section V.

### II. FUZZY LOGIC AS A COMPUTATIONAL APPROACH FOR QUANTITATIVE MODELLING

Processing the data with the use of fuzzy logic can be also referred to as computing with words. In order to describe a process with fuzzy logic, input and output fuzzy variables (e.g. ProteinConcentration and ConcentrationChange) and their fuzzy values (e.g. Low and High) have to be defined. Calculation of the values of the output fuzzy variables is performed with the evaluation of *if-then* fuzzy rules on the input variables and their values (e.g. IF *ProteinConcentration* IS *Low* THEN *ConcentrationChange* IS *High*). Fuzzy logic can be used in the combination with the ordinary, i.e. crisp logic. However, fuzzy values of input fuzzy variables and crisp values of output fuzzy variables need to be calculated in order to combine the fuzzy computation with the crisp one (see Figure 1). We refer to these two processes as *fuzzification* and *defuzzification*. Fuzzification and defuzzification are defined on the basis of membership functions which characterize each fuzzy value regarding the value of its corresponding crisp variable [26].

Construction of a general fuzzy model therefore consists of the following steps:

- identification of input and output fuzzy variables (e.g. *ProteinConcentration, ConcentrationChange*, etc.),
- determination of fuzzy values that define each fuzzy variable (e.g. ProteinConcentration =

Low, Medium, High),

- determination of fuzzy rules that describe the dependence of the output fuzzy variables on the input fuzzy variables,
- fuzzification definition of transformation of a crisp variable to a fuzzy variable,
- defuzzification definition of transformation of a fuzzy variable to a crisp variable.

Knowledge obtained from existing modelling approaches can help us with the establishment of the fuzzy description of the observed process. For example, even though some kinetic data might be unknown, we can use an ODE based model to determine input and output variables and to make a rough estimation on the relations among the inputs and outputs (e.g. linear, exponential, etc.).

### A. Fuzzy description of a biological process

Current state of the biological system is usually described with the vector of concentrations of observed chemical species. Fuzzy description of the current state can be on the other hand defined by linguistic terms, i.e. with the fuzzy values that describe the fuzzy variables. The formal description of a fuzzy value is determined with its membership function, which defines the membership value from 0 (completely not a member) to 1 (completely a member) of a crisp value to a fuzzy one. Most common membership functions have a triangular or trapezoidal shape (see Figure 2), but different shapes may also be used. However, those are in rare cases required to achieve the correct description [27].

The number of fuzzy values used for the description of a fuzzy variable depends on the nature of the process we are modelling. Some processes require more accurate descriptions which can be achieved with a larger number of fuzzy values. On the other hand other processes express the activity only under certain conditions, e.g. when the input variable is very low, and can be described accurately with a relatively small number of fuzzy values.

If-then fuzzy rule base can be established once the fuzzy variables and their possible values are defined. Fuzzy rules present the linguistic expressions that define the relations



Fig. 2. Fuzzy values that define the fuzzy variable *Protein concentration* are formally defined by the membership functions with a triangular (a) or trapezoidal shapes (b). Different shapes may also be used, but are in rare cases required.

between the input and output fuzzy variables and can be usually established intuitively with the linguistic description of the system's dynamics.

Our approach will be used to quantitatively describe the system state changes caused by the reactions in observed GRN. Each fuzzy value will be defined with the concentrations in a certain interval and their corresponding membership values (e.g. if concentration of a protein can range from 0 to 1000nM, concentrations from 0 to 650nMcan be completely referred to as not High, i.e. membership value is 0; concentrations from 900 to 1000nM can be completely referred to as High, i.e. membership value is 1; and concentrations from 650 to 900nM as something in between, i.e. membership values linearly increase from 0 to 1; see Figure 2b). Output fuzzy variables will be defined as absolute changes of the concentrations caused by the processes that describe observed chemical reactions. Rule base will therefore have the form such as IF Protein-Concentration IS High THEN ConcentrationDecrease IS High.

Fuzzy logic can be as such used to quantitatively describe a biological process with only partial knowledge of the system's dynamics and without the direct use of kinetic data. In order to make the approach compatible with other modelling techniques that only operate with crisp values fuzzification and defuzzification processes are used in the input and output segment of a fuzzy model. This allows us to use the fuzzy logic only in the parts of the model in which kinetic data are unknown and to use conventional approaches elsewhere.

# B. Combining fuzzy logic with the existing modelling approaches

It is evident that the inputs and outputs of the fuzzy model will always be crisp values, i.e. current concentrations as inputs and concentration changes as outputs. This enables us to use our fuzzy model only as a replacement for a certain part of the conventional model in which kinetic data are unknown.

Current state in a biological system is usually described with the vector of concentrations of observed chemical species, i.e.  $\mathbf{x} = (x_1, x_2, ..., x_n)$ . System change can be described with the following set of differential equations:

$$\frac{d[x_i]}{dt} = \sum_{j=1}^{m} f_{i,j}(\mathbf{x}), \text{ for } i = 1, \dots, n,$$
(1)

where each function describes a different process (e.g. transcription, translation, etc.) and has its own set of kinetic parameters (e.g. transcription rate, translation rate, etc.). Using our fuzzy logic approach, we can replace of the any functions that are only partially known due to missing kinetic parameters:

$$\frac{d[x_i]}{dt} = FL_k(\mathbf{x}) + \sum_{j=1}^m f_{i,j}(\mathbf{x}),$$
  
for  $i = 1, \dots, n$  and  $k \neq j$ , (2)

where  $FL_k(\mathbf{x})$  is the fuzzy logic model of the process for which kinetic parameters are unknown. Inputs to our fuzzy logic model are crisp values of the concentrations of observed species, while the output is a crisp value of change in concentration of  $x_i$ . The output of our model is combined with the output of other functions to obtain the changes in concentrations of observed species for each time step of the simulation.

### III. CASE STUDY: 3-GENE REPRESSILATOR

Repressilator is a GRN that consists of an arbitrary number of genes, which are connected in a circular repression scheme. It has been shown that only the topologies of a repressilator with odd numbers of genes may exhibit oscillations for certain parameter values (kinetic rates) [7]. We will demonstrate the establishment of a quantitative fuzzy logic model on the 3-gene repressilator (see Figure 3). Even though the proposed approach could be used to describe the whole system, we will presume that the only process that is partially unknown due to the missing kinetic rates is transcription. Here we present the quantitative fuzzy description of transcription only. However, the fuzzy presentation of translation and degradation could be made in the same way straightforwardly.



Fig. 3. Circular repression scheme of the 3-gene repressilator.

Dynamics of the 3-gene repressilator is determined by the production of three different mRNA species (i.e. transcription), production of three different protein species (i.e. translation) and degradation of all mRNA and protein species. Transcription of  $mRNA_i$  is dampened by the presence of a protein  $P_j$ , where j represents the index of the protein that inhibits the production of  $mRNA_i$ , i.e.  $j = (i-1) \mod 3$ .

While we presume transcription rate to be unknown, it is impossible to find a numerical solution for the conventional deterministic model. However, we will demonstrate that the introduced quantitative fuzzy model is able to produce quantitatively relevant results, even though certain kinetic rates are unknown.

### A. Conventional deterministic model

Our reference model will be based on the system of ODEs. We assume that all genes of the repressilator have the same dynamical properties, i.e. equal kinetic rates [7]. The system of ODEs that defines the dynamics is as follows:

$$\frac{d[P_i]}{dt} = k_{tsl} \cdot [mRNA_i] - k_{deg_P} \cdot [P_i], \qquad (3)$$

$$\frac{d[mRNA_i]}{dt} = \frac{k_{tsk}}{1 + [P_j]^n} - k_{deg_{mRNA}} \cdot [mRNA_i], \quad (4)$$

where Eqn. (3) presents the protein concentration change (translation and degradation), in which  $P_i$  is the current protein concentration,  $mRNA_i$  current mRNA concentration,  $k_{tsl}$  translation rate and  $k_{deg_P}$  protein degradation rate. Eqn. (4) presents mRNA concentration change (transcription and degradation), in which  $P_j$  is the current repressor protein concentration, n Hill coefficient,  $k_{tsk}$  transcription rate and  $k_{deg_{mRNA}}$  mRNA degradation rate. Table I shows the values of all kinetic rates that will be used in our simulations and are derived from [7].

TABLE I KINETIC PARAMETERS THAT WILL BE USED IN THE REFERENCE MODEL.

n.	Process	Parameter	Value
(1)	Transcription	$k_{tsk}$	$1.6min^{-1}$
(2)	Transcription	n	2
(3)	Translation	$k_{tsl}$	$2.6min^{-1}$
(4)	mRNA degradation	$k_{deg_{mRNA}}$	$0.12min^{-1}$
(5)	Protein degradation	$k_{deq_P}$	$0.06min^{-1}$

### B. Quantitative fuzzy transcription model

The mRNA concentration changes are described in Eqn. (4). First part of the equation presents the changes in the concentration caused by transcription. To demonstrate the proposed fuzzy approach, we will assume that transcription rate  $(k_{tsk})$  is unknown. Other parameter values will be derived from Table I. Quantitative fuzzy transcription model construction procedure can be described with the following steps:

- 1) *identification of known and unknown parameter values* (in our case only the value of  $k_{tsk}$  is unknown).
- analysis of the correlation between the kinetic rates and transcription dynamics,

- 3) establishment of the linguistic description of transcription dynamics,
- 4) *quantitative fuzzy model construction* (fuzzification and defuzzification, establishment of if-then rules, membership functions and potential scaling).

1) Identification of known and unknown parameter values: The equation that describes transcription can be derived from Eqn. (4) and has the following form:

$$\frac{d[mRNA_i]}{dt} = \frac{k_{tsk}}{1 + [P_j]^n},\tag{5}$$

where

- [P<sub>j</sub>] is current concentration of repressor protein (j = 1, 2, 3) an input variable to a quantitative fuzzy logic model, which will be fuzzified to a fuzzy variable,
- d[mRNA<sub>i</sub>] presents an increase of mRNA<sub>i</sub> concentration (i = 1, 2, 3) in time step dt an output fuzzy variable from a quantitative fuzzy logic model, which will be transformed to a crisp variable as an absolute concentration change for Eqn. (4),
- *n* is Hill coefficient the correlation between transcription and Hill coefficient will be analysed; its results will be used in a combination with the values from Table I to construct the quantitative fuzzy logic model,
- $k_{tsk}$  is transcription rate transcription rate is assumed to be unknown. Its effects on transcription, based on the step 2 of a quantitative fuzzy transcription model construction procedure, will be considered when constructing fuzzy model.

2) Analysis of the correlation between the kinetic rates and transcription dynamics: To understand how parameters n and  $k_{tsk}$  affect the mRNA concentration increase due to transcription, the correlation between their values and system's dynamics are analysed on the basis of Eqn. (5). Figure 4 shows the dependence of the mRNA concentration increase on transcription rate values from  $k_{tsk} = 0.5$  to  $k_{tsk} = 5$ , if n equals 1, 2, 3 or 4. Figure 5 shows the dependence of the mRNA concentration increase on Hill coefficient from n = 0 to n = 3, if  $k_{tsk}$  equals 0.5, 1, 3 or 10.

3) Establishment of the linguistic description of transcription dynamics: Figures 4 and 5 indicate that the repressor proteins drastically affect transcription even when their concentrations are relatively low, i.e. gene is completely silenced in most cases (e.g. when the repressor concentrations reach the threshold of 15nM). Transcription may therefore increase the concentration of mRNA only when the repressor protein concentrations are in the interval [0, 15nM]. The linguistic description can be constructed on the basis of our observations:

• The represor concentrations can be divided in two parts: a small interval of low concentrations, where transcription is active (e.g. repressor concentration is lower than 15nM at  $k_{tsk} = 5$ ; see Figure 4(d)) and the rest of the interval, where transcription is silenced (e.g. repressor concentration is higher than 15nM at  $k_{tsk} = 5$ ; see Figure 4(d)).



Fig. 4. Different colours indicate the mRNA concentration change. Figures present how the mRNA concentration changes at different values of  $k_{tsk}$ , where n = 4 (a), n = 3 (b), n = 2 (c) and n = 1 (d).



Fig. 5. Different colours indicate the mRNA concentration change. Figures present how the mRNA concentration changes at different values of n, where  $k_{tsk} = 10$  (a),  $k_{tsk} = 3$  (b),  $k_{tsk} = 1$  (c) and  $k_{tsk} = 0.5$  (d).

- Transcription is active when the repressor concentrations are low (even though we only show the lower part of protein concentration interval ([0, 30nM]), it is enough to demonstrate that transcription completely stops when the concentrations are higher than the threshold).
- Transcription rate increases when the repressor concentrations are relatively low. (e.g. mRNA concentra-

tion change increases when the repressor concentrations decrease from 15nM to 0nM; see Figure 4(d) at  $k_{tsk} = 5$ ).

- By increasing the value of transcription rate  $k_{tsk}$  the mRNA concentrations change increase linearly (the edge of the non-zero mRNA concentration change is linearly proportional to  $k_{tsk}$  in Figure 4).
- By decreasing the Hill coefficient n the repressor concentrations interval where transcription is still active widens (e.g. interval [0, 5nM] at n = 3 increases to [0, 15nM] if n = 1; see Figure 5(c)).

4) Quantitative fuzzy model construction: The quantitative fuzzy logic model presented in Figure 6 can be constructed on the basis of the linguistic description given above.

Crisp input variable presents the repressor concentration and will be transformed to a fuzzy variable Repressor-Concentration with fuzzy values Low and High. While Low describes the lower part of possible concentrations, where transcription is active, High describes the rest of the possible concentrations, where transcription is silenced. Output fuzzy variable mRNAIncrease will also be described with the fuzzy values denoted Low and High. However, these values will have different membership functions than the ones describing input fuzzy variable. Fuzzy variable values should include all possible values of respective crisp variables (e.g. we can presume that the repressor concentrations always lie between 0 and 500nM). While we presume that transcription rate parameter is not exactly known, it is impossible to presume the maximal repressor concentration as well as the maximal mRNA concentration increase. Therefore both input and output fuzzy variable intervals are normalized to interval [0, 1], where values close to 0 correspond to value Low, while values close to 1 correspond to value *High* (see Figure 7).



Fig. 7. Membership functions for fuzzy sets *Low* and *High*, which describe the concentration of repressor regulating the observed transcription process (a) and mRNA concentration change (b). The concentrations are normalized to interval [0,1] and are therefore unitless.

In the case of output fuzzy variable, value *Low* describes small or almost negligible increase of the mRNA concentration and is active when the repressor concentrations are *High*. Respectively, *High* presents a maximal increase of the mRNA concentrations, and becomes active as the repressor concentrations go towards 0. If-then rule set, which describes the observed behaviour, can be established once the fuzzy variable values are defined with their cor-



Fig. 6. The quantitative fuzzy logic model of transcription will have a crisp input variable  $P_j$ , which will be fuzzified to fuzzy variable *RepressorConcentration* and an output fuzzy variable *mRNAIncrease*, which will be defuzzified to a crisp variable defining mRNA increase due to transcription.

responding membership functions. The number of fuzzy rules is bounded by the number of values from the input fuzzy variable. In our case we need two rules to describe all possibilities:

```
    IF RepressorConcentration IS Low
THEN mRNAIncrease IS High,
    IF RepressorConcentration IS High
THEN mRNAIncrease IS Low.
```

It is convenient to use the proposed approach only for the parts of the system that lack the exact kinetic data. For this purpose input and output variable values need to be scaled to the ranges of the concentrations observed in other parts of the system in which conventional modelling techniques are used. Scaling is performed with the functions prescale (see Algorithm III.1), which maps the repressor concentrations from a crisp value to an interval [0, 1] and *postscale* (see Algorithm III.2), which maps the mRNA concentration changes from the interval [0, 1] to a crisp value. These two functions dynamically adjust the maximal protein concentrations and the maximal mRNA concentration changes according to the crisp values of the concentrations in other segments of the model. Both parameters ( $prescale_{in}$ ,  $postscale_{in}$ ) that determine scaling are initially set to 1 (no scaling). While prescale<sub>out</sub> is increased by input variable  $P_j$  (prescale<sub>out</sub> is assigned to  $prescale_{in}$  in the next iteration),  $postscale_{out}$  is incrementally increased by parameter multiplier until the maximal concentration change is reached. The parameter is set to 1.01 in our simulation (1% increase for every iteration). Increasing *multiplier* will cause faster convergence to the final value of *postscale*, however, it might also introduce bigger error due to larger changes of *postscale* in every iteration. Introduced functions allow us to describe the unknown processes quantitatively with the knowledge we obtain from the known parts of the system.

Algorithm III.1 Function *prescale* that applies the quantitative context to the process described with the fuzzy logic model.

 $prescale_{in}$  = maximal protein concentration  $prescale_{out}$  = adjusted maximal protein concentration input = input protein concentration (crisp input of the fuzzy model)  $input_{prescaled}$  = scaled input protein concentration (mapped to the interval [0,1])

function PRESCALE( prescale<sub>in</sub>, input )

```
prescale_{out} \leftarrow prescale_{in}
```

```
\begin{array}{ll} \text{if } input > prescale_{in} & \text{then} \\ prescale_{out} \leftarrow input \\ \text{end if} \end{array}
```

 $input_{prescaled} \leftarrow input/prescale_{out}$ return  $[input_{prescaled}, prescale_{out}]$ end function

Algorithm III.2 Function *postscale* that applies the quantitative context to the process described with the fuzzy logic model.

 $postscale_{in}$  = maximal mRNA concentration change  $postscale_{out}$  = adjusted maximal mRNA concentration change output = normalized mRNA concentration increase (output of the fuzzy model)

 $output_{postscaled}$  = scaled mRNA concentration increase multiplier = multiplication factor for  $postscale_{out}$ 

**function** POSTSCALE( *postscale*<sub>in</sub>, *output*, *multiplier* )

 $postscale_{out} \leftarrow postscale_{in}$ 

 $\begin{array}{l} \text{if } output \cdot multiplier > postscale_{out} \ \text{ then } \\ postscale_{out} \leftarrow output \cdot multiplier \\ \text{end if} \end{array}$ 

 $output_{postscaled} \leftarrow output \cdot postscale_{out}$ return  $[output_{postscaled}, postscale_{out}]$ end function

### IV. RESULTS AND DISCUSSION

Conventional model (see Section III-A) and quantitative fuzzy logic model (see Section III-B) were constructed in MATLAB Simulink<sup>1</sup>. *Fuzzy toolbox* was used for the construction of fuzzy logic model of transcription. Parameter values used in both models were derived from Table I. Simulations were performed with the ode4 *Runge-Kutta* engine for numerical solving using a fixed time step of 0.1 minutes. The system's dynamics was simulated for 2000 minutes. An example of a simulation run on both models is presented in Figure 8.



Fig. 8. Simulation results of the fuzzy logic (a) and the conventional model (b) of the 3-gene repressilator using the parameter set from Table I. Transcription rate is omitted from the fuzzy model. Both simulation runs reveal stable oscillations with comparable frequencies and amplitudes.

We analysed the presence of oscillations for different parameter values, i.e. protein degradation and translation rates were varied. The quantitative accuracy of the fuzzy approach was measured with the agreement of the frequencies and amplitudes of oscillations between the conventional and fuzzy model. Frequencies were determined using Fast Fourier transform (FFT) analysis. We expected some dissimilarities between the conventional and the fuzzy model, while transcription rate was not used in the latter. However, quantitative relevance should be retained in the fuzzy model. The results presented in Figure 9 indicate that the parameter range for which the system exhibits oscillatory behaviour is wider when transcription is modelled with fuzzy logic. However, the amplitude and the frequency of oscillations are comparable to those obtained with the conventional model. Even if we run the ODE model with different values of transcription rate, frequency of oscillations stays the same. On the other hand amplitude changes proportionally to increase or decrease of transcription rate. Nevertheless, for biologically relevant values of transcription rate, changes to amplitude are not significant and remain comparable to those obtained by our fuzzy approach.

Dissimilarities between the conventional and the fuzzy model arise especially in the bifurcation regions, i.e. in the parameter space where the system transitions from the convergence to a steady state to oscillatory behaviour. However, frequency analysis confirms that the fuzzy logic



Fig. 9. Frequency analysis for different values of translation  $(k_{tsl})$  and degradation rates  $(k_{deg_{Protein}})$  for both models. Figures demonstrate that the frequencies of oscillations obtained by both models are comparable and that the Fuzzy approach retains the quantitative relevance of the ODE model.

model preserves the quantitative relevance of simulation results despite the missing kinetic data.

### V. CONCLUSION

Missing kinetic data present a major obstacle in the quantitative modelling of biological systems. Even though some data are missing, various parameter estimation techniques may be used for their evaluation. These techniques however often require large sets of experimental data, which are sometimes very hard or even impossible to obtain. Here we introduced an alternative approach that exploits the properties of fuzzy logic and enables us to obtain quantitatively relevant simulation results even though the kinetic data are incomplete. While the accuracy of simulations is partially lost, they can still be used to produce results with biological relevance. Moreover, the approach presented here is compatible with conventional state-of-theart modelling approaches. We successfully demonstrated the establishment of proposed modelling methodology in the combination with ODE based model on fuzzy transcription in the reaction network of the 3-gene repressilator. Translation or degradation could be modelled in the same way straightforwardly. Since the proposed method relies on the knowledge unrelated to kinetic data as well as on the kinetic data that is known, dissimilarities between the fuzzy and the conventional model would increase with the number of processes modelled by fuzzy logic. However, the approach would still be able to produce quantitative results with biological relevance.

### ACKNOWLEDGMENTS

The research was supported by the national postgraduate programme Higher Education National Scheme, financed by the European Union (EU), University of Ljubljana and Slovenian Ministry of Higher Education, Science and Technology. Results presented here are in scope of PhD thesis that is being prepared by Jure Bordon, University of Ljubljana, Faculty of Computer and Information science.

#### REFERENCES

 U. Alon, An Introduction to Systems Biology: Design Principles of Biological Circuits, ser. Chapman and Hall/CRC mathematical & computational biology series. Chapman & Hall/CRC, 2007. [Online]. Available: http://books.google.si/books?id=pAUdPQICZ54C

<sup>&</sup>lt;sup>1</sup>The MATLAB and Simulink models are available at http://lrss.fri.unilj.si/bio/material/tcbb\_Bordon.zip under the Creative Commons Attribution license.

- [2] R. Kitney and P. Freemont, "Synthetic biology the state of play." FEBS Lett, vol. 586, no. 15, pp. 2029–2036, Jul. 2012.
- [3] E. Andrianantoandro, S. Basu, D. K. Karig, and R. Weiss, "Synthetic biology: new engineering rules for an emerging discipline," *Molecular Systems Biology*, vol. May, pp. 1–14, 2006.
- [4] E. Fung, W. W. Wong, J. K. Suen, T. Bulter, S.-g. Lee, and J. C. Liao, "A synthetic gene-metabolic oscillator," *Nature*, vol. 435, no. 7038, pp. 118–122, 2005.
- [5] S. L. Michael B. Elowitz, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, pp. 335–338, 2000.
- [6] J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty, "A fast, robust and tunable synthetic gene oscillator," *Nature*, vol. 456, pp. 516–520, 2008.
- [7] N. Strelkowa and M. Barahona, "Switchable genetic oscillator operating in quasi-stable mode," *Journal of The Royal Society Interface*, vol. 7, no. 48, pp. 1071–1082, 2010.
- [8] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in Escherichia coli," *Nature*, vol. 403, pp. 339– 342, 2000.
- [9] Y. Y. Chen, K. E. Galloway, and C. D. Smolke, "Synthetic biology: advancing biological frontiers by building synthetic systems," *Genome biology*, vol. 13, no. 2, p. 240, 2012.
- [10] L. Chen and R. Wang, "Designing gene regulatory networks with specified functions," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, pp. 2444–2450, 2006.
- [11] H. de Jong, "Modeling and simulation of genetic regulatory systems: A literature review," *Journal of Computational Biology*, vol. 9, pp. 67–103, 2002.
- [12] J. N. Weiss, "The hill equation revisited: uses and misuses," FASEB Journal, vol. 11, pp. 835–841, 1997.
- [13] D. T. Gillespie, "Stochastic simulation of chemical kinetics," *Annual Review of Physical Chemistry*, vol. 58, pp. 35–55, 2007.
- [14] A. Cheng and T. K. Lu, "Synthetic biology: An emerging engineering discipline," Annual Review of Biomedical Engineering, vol. 14, pp. 155–178, 2012.
- [15] C. G. Moles, P. Mendes, and J. R. Banga, "Parameter estimation in biochemical pathways: A comparison of global optimization methods," *Genome research*, vol. 13, no. 11, pp. 2467–2474, 2003.
- [16] G. Lillacci and M. Khammash, "Parameter estimation and model selection in computational biology," *PLoS Computational biology*, vol. 6, no. 3, p. e1000696, 2010.
- [17] J. Sun, J. M. Garibaldi, and C. Hodgman, "Parameter estimation using metaheuristics in systems biology: A comprehensive review," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 9, no. 1, pp. 185–202, 2012.
- [18] R. Küffner, T. Petri, L. Windhager, and R. Zimmer, "Petri nets with fuzzy logic (PNFL): Reverse engineering and parametrization," *PLoS One*, vol. 5, no. 9, p. e12807, 2010.
- [19] L. Windhager, "Modeling of dynamic systems with Petri nets and fuzzy logic," Ph.D. dissertation, Ludwig-Maximilians-Universität München, April 2013. [Online]. Available: http://nbn-resolving.de/ urn:nbn:de:bvb:19-156555
- [20] L. Windhager and R. Zimmer, "Intuitive modeling of dynamic systems with Petri nets and fuzzy logic," in *Proc German Conf Bioinf*, vol. 136. Citeseer, 2008, pp. 106–115.
- [21] A. P. Heath and L. E. Kavraki, "Computational challenges in systems biology," *Computer Science Review*, vol. 3, no. 1, pp. 1–17, 2009.
- [22] Y. Gendrault, M. Madec, V. Wlotzko, C. Lallement, and J. Haiech, "Fuzzy logic, an intermediate description level for design and simulation in synthetic biology," in *Biomedical Circuits and Systems Conference (BioCAS)*, 2013 IEEE. IEEE, 2013, pp. 370–373.
- [23] R. I. Hamed, S. Ahson, and R. Parveen, "A new approach for modelling gene regulatory networks using fuzzy Petri nets," *Journal* of *Integrative Bioinformatics*, vol. 7, no. 1, p. 113, 2010.
- [24] J. Dickerson, Z. Cox, E. Wurtele, and A. Fulmer, "Creating metabolic and regulatory network models using fuzzy cognitive maps," in *IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th*, vol. 4. IEEE, 2001, pp. 2171–2176.
- [25] P. Du, J. Gong, E. S. Wurtele, and J. A. Dickerson, "Modeling gene expression networks using fuzzy logic," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 35, no. 6, pp. 1351–1359, 2005.
- [26] L. Zadeh, "Fuzzy logic and approximate reasoning," Synthese, vol. 30, no. 3, pp. 407–428, 1975.
- [27] G. J. Klir and B. Yuan, Fuzzy sets and fuzzy logic. Prentice Hall New Jersey, 1995.



**Jure Bordon** received his BSc degree in Computer Science from the Faculty of Computer and Information Science, University of Ljubljana, Slovenia in 2011.

He is a PhD student and currently employed as a teaching assistant in the Computer Structures and Systems Laboratory at the Faculty of Computer and Information Science, University of Ljubljana, Slovenia, where he is in charge of the several laboratory courses. He is also a member of Computational biology group at the

same faculty. His research interests are mainly directed towards Fuzzy logic modelling and computational biology. He has published his work in several conference and workshop proceedings.



**Miha Moškon** received his BSc degree in Computer Science from the Faculty of Computer and Information Science, University of Ljubljana, Slovenia in 2007 and his PhD in 2012.

He is currently employed as an assistant professor in the Computer Structures and Systems Laboratory at the Faculty of Computer and Information Science, University of Ljubljana, Slovenia, where he is in charge of the several laboratory courses. He is also a member of Computational biology group at the same faculty.

His main research interests are computational biology and unconventional processing methods. He has published his work in several international journals and conference proceedings.



Nikolaj Zimic received his BSc, MSc and PhD degree in computer science from the Faculty of Computer and Information Science, University of Ljubljana, Slovenia in 1984, 1990 and 1994.

Nikolaj Zimic is a full professor and a dean of the Faculty of Computer and information Science (Ljubljana). He leads the Laboratory of Computer Structures and Systems. His research work is mainly focused on the fields of unconventional computing, fuzzy logic and computer networks. He led several industrial and research

projects covering the topics from the optimization of network protocols to implementation of fuzzy algorithms for control and decision systems. He is an occasional reviewer for several international journals, such as Fuzzy sets and Mathematical and Computer Modeling of Dynamical Systems. His research work has been published in different international scientific journals and on several conferences. He is a member of IEEE professional society.



**Miha Mraz** received his BSc, MSc and PhD degree in computer science from the Faculty of Computer and Information Science, University of Ljubljana, Slovenia in 1992, 1995 and 2000.

He is currently employed as a full professor at the Faculty of Computer and Information Science, University of Ljubljana, Slovenia and is also a head of Computational biology group at the same faculty. His research interests are recently directed towards unconventional processing methods, such as fuzzy logic, synthetic

biological systems and QCA structures. He has published his work in many distinguished journals such as Nanotechnology, Journal of Theoretical Biology, International Journal on Unconventional computing, Japanese journal of applied physics, etc. He is also in the editorial board of newly founded Journal of Synthetic Biology. He is a member of IEEE professional society.