Creating Fuzzy Models from Limited Data

SAŠO BLAŽIČ
Faculty of Electrical Engineering,
University of Ljubljana,
Tržaška 25, Ljubljana,
SLOVENIA

Abstract: - The design of experiments is a methodological approach in which measurement experiments are carefully planned to obtain highly informative data. This paper addresses the challenge of constructing mathematical models for complex nonlinear processes when the available measurement data have low information content. This problem often arises when data are collected without the guidance of an experimental modeling expert. We examine two practical examples to illustrate this issue: a textile wastewater decolorization process and atmospheric corrosion of structural metal materials. In both cases, the measured data were insufficient to construct highly accurate models. It is, therefore, necessary to make a trade-off between model complexity and accuracy by adapting modeling techniques to work effectively with the limited data available. The main aim of the paper is, therefore, to focus on simple but effective techniques that allow as much information as possible to be extracted from low-quality measurements and to maximize the usefulness of the model for its intended purpose.

Key-Words: - design of experiments, low information content, decolorization process, atmospheric corrosion, fuzzy model, radial basis network.

Received: May 24, 2023. Revised: May 21, 2024. Accepted: June 22, 2024. Published: July 24, 2024.

1 Introduction

This paper addresses a crucial issue in the field of modeling from measurement data. High-quality, informative data are essential for the construction of accurate mathematical models. Achieving this level of data quality requires careful experimental design. The challenge is even greater when working with nonlinear models, such as fuzzy models or artificial neural networks. These types of models involve a large number of parameters that need to be estimated or trained, making the need for highly informative data even more critical. Without sufficient and well-structured data, the reliability and accuracy of the resulting models can be significantly compromised.

Design of Experiments (DOE) is a systematic approach to the challenge of obtaining high-quality data for modeling. The process begins by identifying the key influencing (input) variables and the corresponding consequence (output) variables. Once these variables are identified, the next step is to define the range of their variation. DOE then involves strategically spreading the input variables across the entire experimental space to ensure that the model covers all potential scenarios it will encounter in real-world applications.

By applying DOE, researchers can minimize the number of experiments required while still capturing the essential properties of the system. This efficient approach not only reduces cost and time but also increases the reliability of the model over a wide range of input variable variations. Effective DOE ensures that the experimental design is comprehensive and robust, resulting in models that are both accurate and generalizable.

Design of experiments is a well-established methodology, dating back to the seminal work of [1]. Since then, the benefits of DOE have been well documented in the literature. For example, [2] highlights that DOE provides a powerful tool for maximizing information extraction from experimental data while minimizing resource expenditure. Similarly, [3] emphasizes that carefully designed experiments can lead to significant improvements in the accuracy and efficiency of model building. These principles are particularly important when dealing with complex, nonlinear models such as fuzzy models and artificial neural networks, where the large number of parameters requires carefully structured data to ensure successful training and estimation.

Design of experiments can be applied in various fields, including computer-aided circuit design [4],

production system design [5], evaluation of experiments [6], fuzzy modeling [7], [8], control design [9], tableting process optimization [10], servo system control design [11], to name a few. However, DOE becomes significantly more challenging when developing dynamic models of processes, as the frequency content of the excitation must be carefully considered to ensure accurate results.

Unfortunately, it is often necessary to construct a mathematical model of a process using data collected without the supervision of an experimental modeling expert. This typically results in data that is less informative, making the modeling process significantly more challenging. In such cases, it is crucial to reduce expectations about the level of precision and accuracy that the model can achieve. In practice, it is important for modeling to focus on the dominant dynamics in such cases. This can be achieved by selecting simple but robust modeling algorithms and parameterizing them appropriately. An algorithm with a relatively small number of tunable and design parameters can often cope well with high levels of uncertainty and noise in the data. When expectations for the model are adjusted and suitable methods are chosen, it is possible to develop a model that captures the properties of the system despite the inherent limitations of the data.

This paper presents two practical problems. Both cases are similar in that they involve complex nonlinear chemical processes, the measurement databases are small and include significant uncertainty, and new data could no longer be collected due to practical reasons. The objective is to build a nonlinear model that extracts the available information despite the measurement database being much smaller than typically expected for problems of this complexity.

This paper is organized as follows: Sections 2 and 3 deal with the decolorization process and the atmospheric corrosion process, while Sections 4 and 5 provide the discussion and conclusions.

2 Decolorization Process

This section examines the process of decolorization of textile wastewater from industrial dyeing. There are many techniques for decolorization, [12]. In our case, we measure the absorbance of the wastewater (A), which indicates its "dirtiness" – zero means that the water is completely transparent, while higher values indicate increased opacity. The decolorization process involves adding hydrogen peroxide (H_2O_2) and exposing the wastewater to ultraviolet light (UV) for a certain duration.

We developed a model of this process that was later used for control design. While a fuzzy model approach to model and control this process was used in [13], our work focuses on model development based on a smaller database.

Before proceeding with control design, it is essential to have a robust model of the process to be controlled. This is particularly important when using model-based control methods. The main challenge in this scenario is the quality of the data from which the model is derived. It is well known that various model forms (e.g. fuzzy models, artificial neural networks, spline models, piecewise linear models, etc.) can describe nonlinear processes with arbitrary small modeling errors. However, none of these models can reliably extrapolate information to parts of space where little or no measured data is available.

2.1 Model Structure

When deciding on the model structure, the first consideration is the choice of model inputs and outputs. A primary concern is the determination of the model output, which presents at least two possible choices in this particular example:

- The decolorization factor D. It is given by the formula

$$D = \frac{A_i - A_f}{A_i} \tag{1}$$

  Where A_i represents the initial absorbance (before the decolorization process), and A_f represents the final absorbance (after the decolorization process). Therefore, this means that the process was completely unsuccessful, while $D = 1$ indicating perfect decolorization.
- The final absorbance A_f.

As the model is intended for control design purposes, the focus of the decision-making process should be on the control aspect of the problem. The objective of the control system is to achieve effective decolorization of the effluent. It is therefore essential to establish a specific reference point. In our view, the definition of a target final absorbance serves as an appropriate control objective.

If this approach were not adopted, and the decolorization factor was chosen as the system output (and consequently the control reference), all wastewater treatments would be treated equally regardless of their initial "dirtiness". This would lead to almost identical control actions in different scenarios, undermining the rationale for developing a complex nonlinear model.

In conclusion, the optimal solution for the system output is the final absorbance, A_f.

The choice of model inputs may seem straightforward, but there are still some open questions. Potential candidates for system inputs include: the concentration of the peroxide H_2O_2 (C_p), the power of the UV lamp (P_{UV}), and the time of exposure to the UV lamp (T).

Since the final absorbance is likely to be non-linearly dependent on the initial absorbance, the initial absorbance must be included as one of the system inputs. However, adding a fourth input instead of three increases the amount of data required to build the model. Therefore, to streamline the process, we avoid a 4-dimensional input space. Consequently, the model will take the following form:

$$D = f(C_p, P_{UV}, T) \tag{2}$$

Note that the decolorization factor is used as the output, as we do not include the dependence on the initial absorbance in our model. The final output A_f is then derived from equation (1):

$$A_f = (1-D)A_i = (1-f(C_p, P_{UV}, T))A_i \tag{3}$$

It is important to note the inclusion of the exposure time T in the model described by equation (2). Because the measurements are taken at somewhat arbitrary intervals, the model is structured to explicitly include time as a model input.

2.2 RBN Model of the Decolorization Process

The Radial Basis Network (RBN) is used to approximate the mapping described by (2). The model for the process is based on measured data of the dye Irgalan Gelb 3R KWL: P_{UV} [W], C_p [mg/l], T [min], and the absorbance A (measured only at a few different times T). The data collected include measurements from three replicate sets of 15 experiments each. The sample size is relatively small due to the manual collection and analysis of wastewater samples, which also contributes to higher measurement errors. Obvious outliers were manually removed from the data set. With three inputs to the network and one output, the results are plotted as three-dimensional planes: the x and y axes correspond to two inputs, while the third input is fixed, and the decolorization factor is plotted on the z axis.

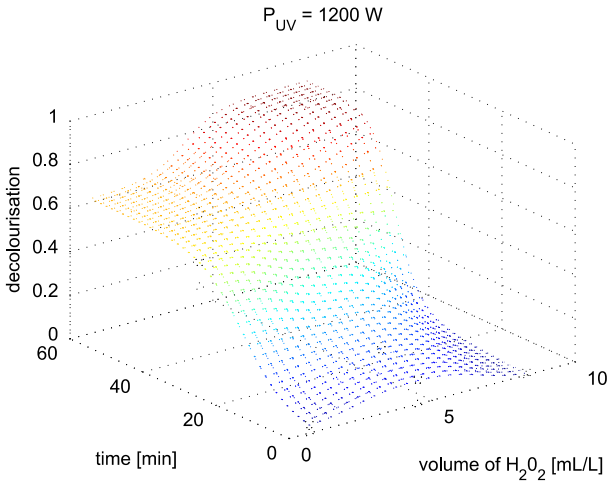

Fig. 1: RBN approximation for $P_{UV} = 1200$ W

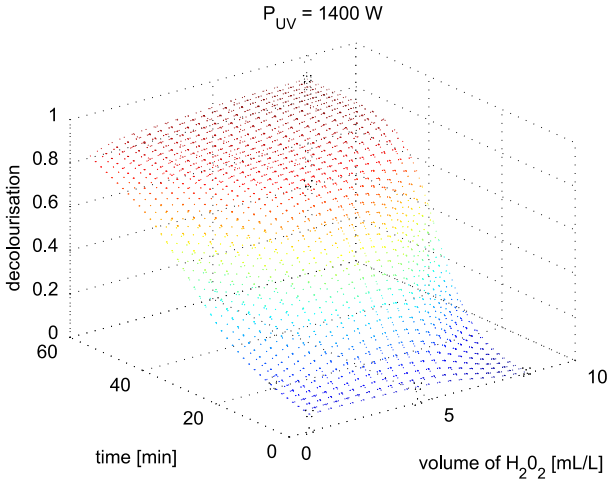

Fig. 2: RBN approximation for $P_{UV} = 1400$ W

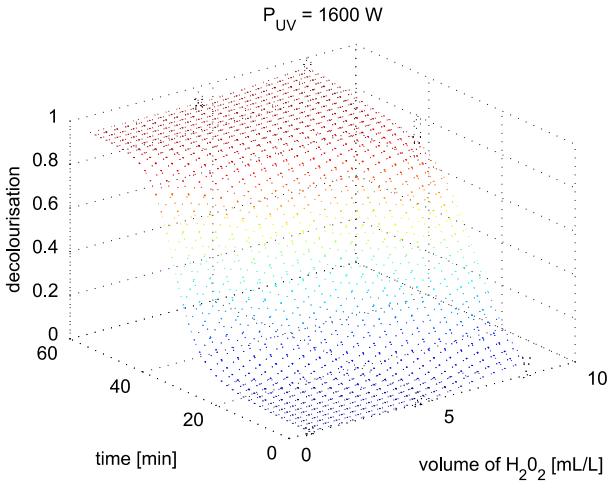

Fig. 3: RBN approximation for $P_{UV} = 1600$ W

Figure 1, Figure 2 and Figure 3 show the decolorization factor plotted against peroxide concentration and time, with the UV lamp power held constant at 1200, 1400, and 1600 W, respectively. The measured data points are

represented by circles, while the network outputs at these inputs are represented by crosses. Figure 4, Figure 5 and Figure 6 show similar plots for fixed values of peroxide concentration. Despite the obvious limitations of the data set (missing important data points under certain operating conditions and high variability in repeated experiments), some qualitative characteristics of the process are visible.

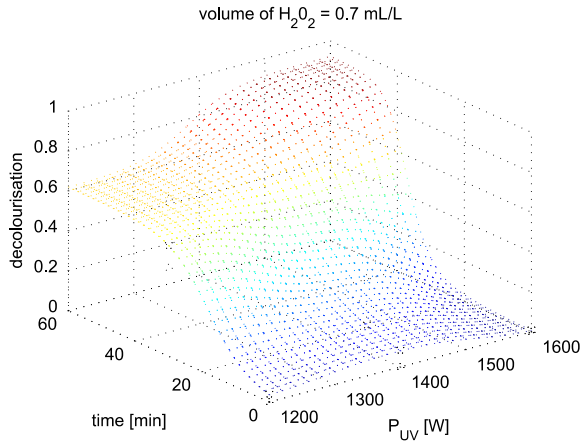

Fig. 4: RBN approximation for C_p = 0.7 mg/l

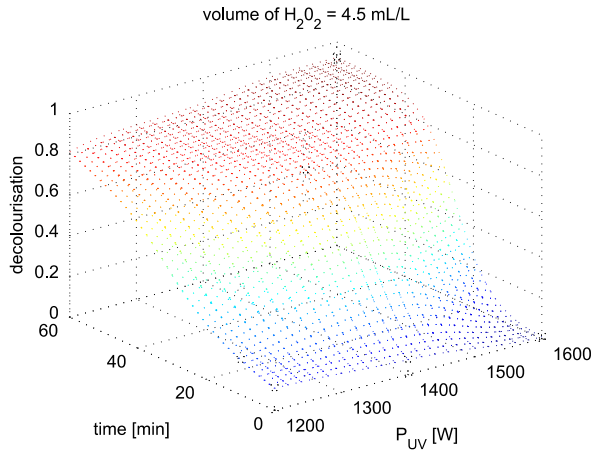

Fig. 5: RBN approximation for C_p = 4.5 mg/l

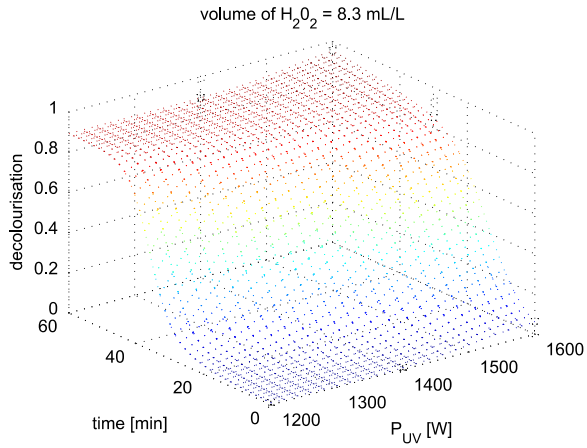

Fig. 6: RBN approximation for C_p = 8.3 mg/l

2.3 Control Design

In this case, the control problem is to determine the actions that lead to purified water. Increasing these actions – C_p, $P_{UV,}$ and T – results in better control results, but on the other hand it also contributes to higher control costs.

We propose optimal control as the solution to the control problem. A crucial aspect of applying optimal control is defining an appropriate cost function that comprehensively captures all relevant aspects of the problem. In our approach we use the following cost function.

$$J = k_d g(A_f) + k_p \frac{C_p}{C_{max}} + k_e \frac{P_{UV} T}{P_{max} T_{max}} \tag{4}$$

Here, $g(A_f)$ represents the function defining the cost associated with unsatisfactory final absorbance:

$$g(A_f) = \begin{cases} 0 & A_f < A_{satis} \\ A_f - A_{satis} & A_f \geq A_{satis} \end{cases} \tag{5}$$

If the final absorbance falls below a certain threshold A_{satis} (the acceptable level), the first term of the cost function, which penalizes unsatisfactory decolorization, will be zero. Conversely, if the final absorbance exceeds this threshold, the first term will have a positive penalty. The second and third terms in (4) represent the costs associated with the consumption of H_2O_2 and energy, respectively, where C_{max}, $P_{max,}$ and T_{max} are the maximum values for C_p, $P_{UV,}$ and T, respectively.

Choosing the appropriate weights k_d, $k_{p,}$ and k_e for the decolorization cost, peroxide cost, and energy cost, respectively, is a delicate task. However, this specific problem is beyond the scope of this paper.

Substituting (3) into (4) gives the control cost function:

$$J(C_p, P_{UV}, T, A_i) = k_d g((1 - f(C_p, P_{UV}, T))A_i) + k_p \frac{C_p}{C_{max}} + k_e \frac{P_{UV} T}{P_{max} T_{max}} \tag{6}$$

The cost function is based on four variables, with the mapping function f approximated by the RBN model. To achieve optimal control, we minimize the cost function (6) in the space of the three control variables (C_p, $P_{UV,}$ and T). As the minimization is performed in a high-dimensional space (C_p, P_{UV}, T, and A_i), three optimal control commands depend only on the initial absorbance, which is known before the decolorization process. By minimizing the cost function (6) off-line, the control functions can be derived:

$$\begin{aligned} P_{UV} &= h_1(A_i) \\ T &= h_2(A_i) \\ C_p &= h_3(A_i) \end{aligned} \tag{7}$$

The control functions described can be readily implemented on dedicated hardware, for example, in tabular format. It is crucial to understand that the shape of these control functions is significantly influenced by the chosen design parameters k_d, k_p, and k_e. Experts in textile engineering have validated the correct behavior of the control system.

3 Atmospheric Corrosion Process

In this section, we model the process of atmospheric corrosion using a Takagi-Sugeno fuzzy model. The resulting model can be used to predict atmospheric corrosion if the near future under simulated climate changes, such as increased atmospheric acidification, lower SO_2 concentrations in Europe, higher O_3 and NO_x concentrations, and global warming.

Atmospheric corrosion of structural metals is a complex and nonlinear process influenced by several meteorochemical and material factors. The full extent of the effects of these factors on material degradation is not yet fully understood. Fortunately, long-term climate programs in Europe (e.g. ECE/EMEP and ICP) and Asia/Africa (e.g. RAPIDC) provide daily measurements of meteorochemical variables such as temperature, relative humidity, precipitation, and major pollutant concentrations. In addition, annual corrosion mass loss data are recorded for metals such as carbon steel, copper, zinc, and aluminum. Corrosion measurements for various metals are also available from many sites around the world.

Previous research has addressed the modeling of these phenomena, [14], [15]. However, the results from existing large databases are not permitted for publication. This paper uses a very small, publicly available database, but still provides significant insights. The analysis of this database also highlights the major challenges typically encountered in the study of atmospheric corrosion.

The main problem in this case is that the model is based on a small database that includes 32 measurements of corrosion taken in different parts of the world. Each set of measurements contains 7 variables, one of which is the corrosion mass. The aim is to build a model that predicts the corrosion mass based on the other 6 measurements, although this is difficult given the sparsity of the data in a high dimensional input space.

3.1 Linear Static Model

As mentioned above, this paper is based on a relatively small atmospheric corrosion database consisting of only 32 measurements. Each measurement consists of exposition time t (in years), temperature T (in ^{0}C), relative humidity H_R (in %), SO_2 concentration C_{SO2} (in $\mu g.m^{-3}$), precipitation p (in mm), pH p_H, and corrosion mass C (in $g.m^{-2}$). We aim to develop a model of this system with corrosion mass C as the output and the other six variables as inputs. Given the limited number of data points in this six-dimensional space and the potentially poor quality of the measurements, we started the modeling process by identifying a static linear system:

$$\hat{\tilde{C}}_i = \begin{bmatrix} \tilde{t}_i & \tilde{T}_i & \tilde{H}_R & \tilde{C}_{SO2} & \tilde{p}_i & \tilde{p}_{H_i} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_6 \end{bmatrix} = \boldsymbol{\psi}_i^T \boldsymbol{\theta}, \; i = 1 \ldots 32 \quad (8)$$

where the hat symbol represents the estimated output of the linear system, and the tilde symbol denotes a normalized variable, which is obtained by subtracting the mean and dividing by the standard deviation of that variable. Normalizing the variables ensures that all variables have a mean of zero and a variance of one, thereby equalizing their contributions to the final model.

The solution in this case can be obtained by a simple method of least squares:

$$\boldsymbol{\theta}^T = \begin{bmatrix} 0.8162 & 0.0564 & 0.0913 & 0.4277 & -0.0123 & -0.1084 \end{bmatrix} \quad (9)$$

It is important to note that the data set is very small, making it impractical to split it into separate identification and validation subsets. In addition, the data spans a large region of the input space, making extrapolation to other regions difficult. The key point is that while the identification of a linear model can be done with a smaller dataset, the identification of a Takagi-Sugeno model requires significantly more data.

Due to the lack of a proper validation set, we will conduct a form of verification. The model error is defined as follows:

$$e_i = \tilde{C}_i - \boldsymbol{\psi}_i^T \hat{\boldsymbol{\theta}}, \; i = 1,2,\ldots 32 \quad (10)$$

The mean-square error (MSE) is then defined as:

$$M = \frac{1}{32} \sum_{i=1}^{32} e_i^2 \quad (11)$$

For the linear model (9), the mean-square error is

$$M = 0.1096 \quad (12)$$

3.2 Takagi-Sugeno Static Model

The first step in building the model is structure identification, which involves selecting the appropriate system structure. Due to the limited

number of measurements, we cannot afford a very complex structure. Therefore, we decided to test each of the six variables as potential antecedents of the fuzzy model. Thus, in the k-th iteration $(k=1,2,3,4,5,6)$, not only did x_k serve as an antecedent variable, but we also replaced it with three fuzzy variables f_1x_k, f_2x_k and f_3x_k ($f_1+f_2+f_3=1$ by design) in the regressor vector. Three Gaussian membership functions $\mu_i = e^{-(x_k-c_i)^2}$ were used for calculating f_i. Since all the signals are normalized and roughly span the interval [-3, 3], the membership functions had centers at $c_1=-2, c_2=0, c_3=2$, with the σ parameter set to 1 in all cases. Three fuzzy variables are therefore given by

$$\mu_1 = e^{-(x_k-c_1)^2} \quad \mu_2 = e^{-(x_k-c_2)^2} \quad \mu_3 = e^{-(x_k-c_3)^2}$$
$$f_1 = \frac{\mu_1}{\mu_1+\mu_2+\mu_3} \quad f_2 = \frac{\mu_2}{\mu_1+\mu_2+\mu_3} \quad f_3 = \frac{\mu_3}{\mu_1+\mu_2+\mu_3} \tag{13}$$

Finally, the regressor vector takes the following form:

$$\boldsymbol{\psi}_k^T = [x_1 \quad \cdots \quad x_{k-1} \quad x_k f_1 \quad x_k f_2 \quad x_k f_3 \quad x_{k+1} \quad \cdots \quad x_6] \tag{14}$$

Similar to the linear model approach, the signals used for identification were normalized. The vector $\hat{\boldsymbol{\theta}}_k$ consisting of eight estimated parameters (3 "fuzzy" and 5 "linear") was obtained by least squares. Of particular interest is the mean square error calculated over the identification data set, following a method analogous to that of the linear model. The results are detailed in Table 1 (second row), which shows the results for different values of k, representing different antecedent variables.

In addition, an experiment was conducted using 7 membership functions centred at -3, -2, -1, 0, 1, 2, 3, all with σ values of 1. The mean square errors from this experiment are presented in Table 1 (third row).

Table 1. Mean-Square Errors of the T-S Model (3 and 7 MFs) and Antecedent Variable x_k

k	1	2	3	4	5	6
M_{3k}	0.0790	0.0928	0.0975	0.1095	0.0907	0.1092
M_{7k}	0.0642	0.0669	0.0708	0.0945	0.0789	0.0959

Comparing the second and third rows of Table 1, the most significant reduction in variance occurs when x_1 is used as the antecedent variable (in both rows). In the second row x_5, x_2 and x_3 are the following candidates for the antecedent. In the third row, x_1 is followed by x_2, x_3 and x_5. Both sequences are similar but not identical due to the specific placement of the membership functions.

Based on the results of this simple analysis, the recommended antecedent variables would be x_1 when using a single antecedent variable, and x_1 and x_2 when using two. The final Takagi-Sugeno model based on the original data uses x_1 and x_2 as antecedent variables (each fuzzified with 3 membership functions, as in the case of a single antecedent variable), resulting in 9 fuzzy variables f_i. The next step was to select one of the six variables (x_k) to be replaced by nine regressors (f_1x_k to f_9x_k). It turned out that the best results were obtained by selecting x_2. In the end, the regressor vector is made up of the other original variables (x_1 and x_3 to x_6) and the nine variables mentioned above (f_1x_2 to f_9x_2). In this configuration, the MSE is 0.0582, with a total of 3x3 + 5 = 14 estimated parameters.

Figure 7 compares the measured outputs with the outputs of three models: the linear model (6 estimated parameters), the T-S model with x_1 as the antecedent variable with 7 membership functions (12 estimated parameters), and the T-S model with x_1 and x_2 as the antecedent variables (each fuzzified with 3 membership functions, for a total of 14 estimated parameters). Ideally, it would be preferable to plot in the original high-dimensional space, but as the input space is 6-dimensional, this is impractical. Therefore, Figure 7 shows the sample index i on the x-axis. We can see that the Takagi-Sugeno model reduces the error significantly as expected. Any validation of the model is impossible in this case due to the scarcity of data, which prevents the data set from being split into a training set and a validation set.

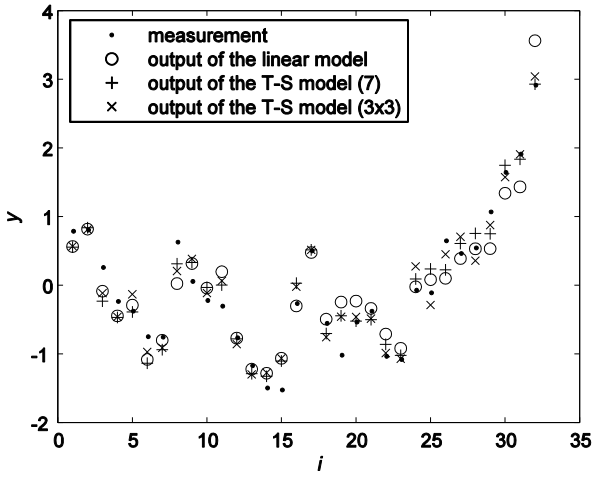

Fig. 7: The comparison between the measured output, the output of the linear model, and the outputs of two T-S models. The data are plotted against the sample index i. Each data set, represented by an index i, also contains 6 input variables, which are not plotted here

4 Discussion

Both cases of modeling nonlinear complex processes share some common features. In both cases, there was a relatively small number of measurements available and these measurements were of poor quality due to inherent problems in the measurement process.

Section 2 discusses the decolorization process. It was shown that the first crucial step is to define the structure that leads to more favorable model properties. Specifically, in this case, the model becomes "less" nonlinear when a particular input is chosen over another option. In addition, as with any modeling approach, the purpose of modeling is essential for validating the model obtained. In our case, the model is used for subsequent optimal control, which selects the control inputs to purify the water while minimizing the use of energy and chemicals. A Radial Basis Network model is therefore used to "filter out" the measurements and produce smooth control laws.

The aim of modeling the atmospheric corrosion process, described in Section 3, is to predict atmospheric corrosion in the near future based on measurements of some influential atmospheric parameters. In this case, a Takagi-Sugeno model was used. The main challenge is the selection of the antecedent variables. Working with an extremely small database in a high-dimensional space poses some problems, but a simple method of selecting one or two antecedent variables is proposed, where the structure of the regressor vector is adapted accordingly.

5 Conclusion

This paper addresses a well-known challenge: constructing a model for a nonlinear process using data with limited information content. While it does not provide a definitive solution, it does stimulate discussion of possible approaches. In such scenarios, the balance between model complexity and accuracy is crucial. Ideally, new experiments could improve data quality, but in the cases presented here, additional measurements were not feasible. This highlights the difficulty of modeling complex systems with sparse data and underlines the need for innovative strategies in such constrained environments.

As it is impossible to provide general guidelines for modeling an arbitrary process based on measured data, it is essential to consider the purpose of the model and adapt the techniques to the amount and quality of data available. In the examples discussed, the most sensible decision would have been to collect more data and focus on trustworthy data. However, as this was not possible, the methods had to be simplified and the resulting model had to be robust to the uncertainty of individual measurements by properly tuning the design parameters of radial basis networks and Takagi-Sugeno models. As a result, the output of the model changes slowly as the inputs vary; in other words, the output of the function has small derivatives with respect to individual inputs. Although other approaches could have been used, this basic philosophy should always be followed.

Acknowledgement:
This work has been supported by the Slovenian Research Agency (ARIS) under Research Program P2-0219.

References:

[1] R. A. Fisher, *The design of experiments*, Oliver & Boyd, Oxford, 1935.

[2] D. C. Montgomery, *Design and Analysis of Experiments*, Wiley, 2017.

[3] G. E. P. Box, J. S. Hunter, and W. G. Hunter, *Statistics for Experimenters: Design, Innovation, and Discovery*. Wiley, 2005.

[4] E. Rémond, E. Nercessian, C. Bernicot, and R. Mina, Mathematical approach based on a "Design of Experiment" to simulate process variations, *Design, Automation & Test in Europe*, Grenoble, 2011.

[5] C. Chaabane, T. Chaari, and D. Trentesaux, "Integrated sizing support system using simulation and design of experiments", *2009 International Conference on Computers & Industrial Engineering*, Troyes, pp. 735-741, 2009, https://doi.org/10.1109/ICCIE.2009.5223808.

[6] O. Tůmová, L. Kupka, and P. Netolický, "Design of Experiments approach and its application in the evaluation of experiments", *2018 International Conf. on Diagnostics in Electrical Engineering*, Pilsen, 2018, https://doi.org/10.1109/DIAGNOSTIKA.2018.8526104.

[7] R.-E. Precup, C.-A. Bojan-Dragos, E.-L. Hedrea, R.-C. Roman, and E. M. Petriu, "Evolving fuzzy models of shape memory alloy wire actuators", *Romanian Journal of Information Science and Technology*, vol. 24, no. 4, pp. 353-365, Dec. 2021, https://www.romjist.ro/abstract-698.html.

[8] Z. C. Johanyák, "Fuzzy rule interpolation based model for student result prediction", Journal of Intelligent & Fuzzy Systems, vol. 36, no. 2, pp. 999-1008, Apr. 2019, https://doi.org/10.3233/JIFS-169875.

[9] P. Aengchuan, and B. Phruksaphanrat, "A comparative study of design of experiments and fuzzy inference system for plaster process control", *Proceedings of the World Congress on Engineering 2017* Vol I, London, 2017.

[10] A. Belič, I. Škrjanc, D. Zupančič-Božič, and F. Vrečer, "Tableting process optimisation with the application of fuzzy models", *International Journal of Pharmaceutics*, vol. 389, no. 1/2, pp 86-93, 2010, https://doi.org/10.1016/j.ijpharm.2010.01.023.

[11] I. A. Zamfirache, R.-E. Precup, R.-C. Roman, and E. M. Petriu, "Reinforcement learning-based control using Q-learning and gravitational search algorithm with experimental validation on a nonlinear servo system", *Information Sciences*, vol. 583, pp. 99-120, Jan. 2022, https://doi.org/10.1016/j.ins.2021.10.070.

[12] A. B. dos Santos, F. J. Cervantes, and J. B. van Lier, "Review paper on current technologies for decolourisation of textile wastewaters: Perspectives for anaerobic biotechnology", *Bioresource Technology*, vol. 98, no. 12, pp. 2369-2385, 2007, https://doi.org/10.1016/j.biortech.2006.11.013

[13] K. Kumar, S. Deep, S. Suthar, M. G. Dastidar, and T. R. Sreekrishnan, "Application of fuzzy inference system (FIS) coupled with Mamdani's method in modelling and optimization of process parameters for biotreatment of real textile wastewater", *Desalination and Water Treatment*, Vol. 57, No. 21, pp. 9690-9697, 2016, https://doi.org/10.1080/19443994.2015.1042062

[14] H. Simillion, O. Dolgikh, H. Terryn, and J. Deconinck, "Atmospheric corrosion modelling", *Corrosion Reviews*, Vol. 32, No. 10, pp. 73-100, 2014, https://doi.org/10.1515/corrrev-2014-0023

[15] M. Halama, K. Kreislova, and J. Van Lysebettens, "Prediction of Atmospheric Corrosion of Carbon Steel Using Artificial Neural Network Model in Local Geographical Regions", *Corrosion*, Vol. 67, No. 6, 2011, https://doi.org/10.5006/1.3595099.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

This work has been supported by the Slovenian Research Agency (ARIS) under Research Program P2-0219.

Conflict of Interest

The authors have no conflicts of interest to declare.