This is a repository copy of *New measurement of the body mass index with bioimpedance using a novel interpretable Takagi-Sugeno Fuzzy NARX predictive model*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/179759/

Version: Accepted Version

Proceedings Paper:

He, C., Gu, Y., Wei, H. orcid.org/0000-0002-4704-7346 et al. (1 more author) (2022) New measurement of the body mass index with bioimpedance using a novel interpretable Takagi-Sugeno Fuzzy NARX predictive model. In: Jiang, R., Zhang, L., Wei, H.L., Crookes, D. and Chazot, P., (eds.) Recent Advances in AI-enabled Automated Medical Diagnosis. AI4MED 2021 : International Symposium on Artificial Intelligence for Medical Applications, 19-23 Aug 2021, Virtual Conference. Taylor & Francis , pp. 253-267. ISBN 9781032008431

This is an Accepted Manuscript of a book chapter published by CRC Press in Recent Advances in AI-enabled Automated Medical Diagnosis on 20/10/22, available online: https://www.routledge.com/Recent-Advances-in-AI-enabled-Automated-Medical-Diagnosis/Jiang-Zhang-Wei-Crookes-Chazot/p/book/9781032008431.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Accepted for publication by "Recent Advances in AI-enabled Automated Medical Diagnosis", Edited by Jiang, R., Crookes, D., Wei, H. L., Zhang, L., and Chazot, P. (Accepted 8 August 2021)

Chapter xxx

New Measurement of the Body Mass Index with Bioimpedance Using a Novel Interpretable Takagi-Sugeno Fuzzy NARX Predictive Model

Changjiang He [1], Yuanlin Gu [2*], Hua-Liang Wei [3], Qinggang Meng [4]

[1] *Department of Mathematics and Statistics, Lancaster University, UK.*

[2*] *Department of Computer Science, Roehampton University, UK. (Corresponding author: guyuanlin@hotmail.com)*

[3] *Department of Automatic Control and Systems Engineering, University of Sheffield, UK.*

[4] *Department of Computer Science, Loughborough University, UK.*

Abstract

Body Mass Index (BMI) is an important and useful indicator for medical diagnoses, accurate monitoring and forecasting of BMI are therefore crucial. However, the current measurement of BMI, which is usually highly correlated with the environmental and individual conditions, is inaccurate. Recent developments of bioelectrical impedance show that there is a great potential to improve the measurement of BMI. In this paper, we propose a novel interpretable Takagi-Sugeno Fuzzy NARX (TSF-NARX) model to predict BMI values from bioimpedance signals and anthropometric factors. The proposed model integrates the Nonlinear Auto Regressive Moving Average with Exogenous Input (NARMAX) method and Takagi-Sugeno fuzzy inference. An obvious novelty and advantage of the proposed method is that it provides a new framework, combining the capabilities of fuzzy inference and NARX representation empowered by nonlinear membership functions. The experimental results show that the TSF-NARX model outperforms other models in prediction accuracy and consistency. More importantly, the model identifies both the key

frequency bands and anthropometric factors that highly affect the BMI. The proposed model provides a tool for obtaining accurate, interpretable and robust measurement against the intra and extra uncertainty within the clinical diagnosis.

Keyword
Body Mass Index, bioimpedance, Takagi-Sugeno fuzzy logic, NARMAX method.

1. Introduction

Body Mass Index (BMI) describes the relationship between the mass and height of human. It is a critical indicator for the medical research and diagnosis, especially in the field of obesity, maturity and heritability [1-3]. The conventional measurement of human BMI is derived from the weight and the square of body height. Such a direct measurement can be severely interfered by many factors, such as race, gender and abdominal body structure. Those individual discrepancies can increase the uncertainty and error in decision making, hence, the conventional BMI measurement is not recommended for clinical judgement [4].

Bioelectrical impedance measures the body impedance via a weak electric current. According to the design, it fundamentally depends on the body water measurement within the tissue. Therefore, conventional body composition analysis rarely employs bioelectrical impedance for its high variability. However, with current technological developments, bioelectrical impedance has showed a great potential in BMI analysis [5].

A recent research has achieved the significant improvement of BMI measurement via integrating multi-frequency bioelectrical impedance into the conventional method [6]. However, similar to all other biomarkers, bioelectrical impedance is under the regulation of multiple bioprocesses and its correlation with the BMI is dependent on various conditions [7, 8]. Meanwhile, it is also worth noting that the most frequency bands of bioelectrical impedance share a strong collinearity and may not be directly associated with the BMI.

In order to establish effective and efficient models with the bioelectrical impedance, it is important to explore the frequency bands that are consistently correlated with human BMI. Such filtering process can significantly limit the interference of the collinearity within bioelectrical impedance and the noise from irrelevant psychophysiological activities. In addition, it can also identify the key frequency bands, which can be used for further signal analysis to increase the BMI measurement reliability.

Meanwhile, the distribution and quantity of body water in human is subject to psychophysiological state and individual conditions, and it is a dynamic process under the influence of external factors. As it is shown in the previous research, the direct correlations between some frequency bands of bioelectrical impedance and human BMI are quite low [6]. Therefore, it is necessary to apply feature extraction methods before integrating the bioelectrical impedance signals into the model.

Based on the above considerations, this paper proposes a novel Takagi-Sugeno Fuzzy NARX Model (TSF-NARX) model for prediction of BMI. The main contributions of this work are:

- Establish a hybrid model with interpretable structure for nonlinear BMI modelling.
- Identify the key frequency bands of bioimpedance and anthropometric factors that are correlated with BMI.
- Explore inter- and intra- uncertainty within the BMI measurement using fuzzy logic inference.

The remainder of this paper is as follows. Section 2 presents related works. The proposed TSF-NARX model is introduced in Section 3. The experimental results are presented in Section 4. Finally, the paper is concluded in Section 5.

2. Related Work

The interpretability of machine learning models has become an important topic in recent years [9]. For many real applications, the predictive model should not only achieve good prediction accuracy but also be able to reveal the insights, e.g., how the predictions are produced and what are the effective and most important features. For example, in the BMI prediction, it is crucial to identify the frequency bands and anthropometric factors that are highly correlated with BMI. The transparency of these features and model structures are important for obtaining an insightful understanding of human BMI. Roughly speaking, there are two common types of interpretable models, that is, post-explainable models [9] (e.g. most advanced machine learning models) and intrinsically-explainable models [10] (e.g. nonlinear parametric polynomial regression models). Post-explainability techniques aim to interpret an established machine learning model after the completion of training process. For example, SHapley Additive exPlanations (SHAP) [11] is used to measure the influence and importance of each feature for the prediction purpose. However, the model itself remains to be opaque and it cannot reveal the process of how the prediction was generated. On the other hand, a regression-based model employs fully interpretable structures and features to generate predictions [10]. A limitation of traditional regression-based models may be that they lack prediction accuracy when the required information is incomplete. The Nonlinear Auto Regressive Moving Average with Exogenous Input (NARMAX) model was developed for data modelling and system identification in the time, frequency, and spatiotemporal domains [10]. The NARMAX model can derive nonlinear terms from data of original variables and identify the most effective variables and their interactions (i.e. the product-terms) to build the model. These terms can describe the nonlinear dynamics of complex systems and can usually be linked to the original physical system or process. More importantly, the selected terms and model structure are fully interpretable. The Nonlinear Auto Regressive with Exogenous Input (NARX) model is a special case of the NARMAX model, and it has been successfully applied to solve real data modelling problems in many areas [10, 22-24]. In this study, we take advantage of the NARX model, especially its TIPS (transparent, interpretable, parsimonious and simple/sparse/simulatable) properties [25, 26] and integrate it into our proposed model framework for BMI prediction.

Fuzzy logic combines objective knowledge and subjective knowledge via computing on “degree of truth” rather than the traditional Boolean logic “true or false”. Fuzzy logic handles uncertainty, small size data and data sparsity better than other machine learning paradigms. Compared with the conventional modelling approaches such as neural network, models based on fuzzy logic share significant advantages such as flexible, simple and intuitive [12]. Fuzzy logic-based models have been wildly applied in the biological research for the high complexity and uncertainty within these systems, especially in the area related to medical diagnosis [13-15]. These fuzzy logic-based systems have been proved to be efficient and effective tools to support health condition analyses and clinical decisions. However, current existing fuzzy logic models developed for the BMI prediction purpose are limited to a few specific groups, such as athlete and obesity. The findings from the bioimpedance may allow for new and ground-breaking path-opening to the prediction of general human BMI with fuzzy logic-based model.

3. The proposed TSF-NARX model

The TSF-NARX model combines the type-1 Sugeno fuzzy inference system with the NARX representation. The NARX models are built based on multiple subsets resampled from the original dataset. the means and standard deviations of each identified terms are utilised to create the Gaussian distribution membership functions of each fuzzy rule correspondingly. Given an output variable y(t) and a number of R input variables $u_1(t), u_2(t), \dots, u_R(t)$, the general NARX model can be represented as [10]:

$$y(t) = F[u_1(t-1), u_1(t-2), \dots, u_1(t-n_x), u_2(t-1), u_2(t-2), \dots, u_2(t-n_u), \dots u_R(t-1), u_R(t-2), \dots, u_R(t-n_u), y(t-1), y(t-2), \dots, y\big(t-n_y\big)] + e(t) \quad (1)$$

where t indicates the index of the data samples, $F[\cdot]$ is some nonlinear function, n_u and n_y are the maximum time lags of the input and output variables, $e(t)$ is the noise sequence. Note that equation (1) is the general NARX model for time series prediction with a delay of 1. In some situations, the system is static and there is no delay or time dependencies between the input and output variables. In these cases, the NARX model can be simplified as follows:

$$y(t) = F[u_1(t), u_2(t), \dots, u_R(t)] + e(t) \quad (2)$$

where t indicates the index of the data samples. Usually, the R input variables $u_1(t), u_2(t), \dots, u_R(t)$, and their interactions such as $u_1(t)u_2(t)$ (the nonlinear degree of the term is 2) and $[u_2(t)]^2u_3(t)$ (the nonlinear degree of the term is 3) are used to build a NARX model. The initial

NARX model that contains all the specified model terms can be written in a vector form:

$$\boldsymbol{y} = \theta_1\boldsymbol{\varphi}_1 + \theta_2\boldsymbol{\varphi}_2 + \cdots + \theta_M\boldsymbol{\varphi}_M + \boldsymbol{e} \tag{3}$$

where $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_M$ are a number of M terms derived from the original inputs $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_R$, and $\theta_1, \theta_2, \dots, \theta_M$ are the estimated parameters of these terms. The nonlinear terms are derived using the original inputs through some nonlinear conversions. These derived terms together can approximate the nonlinear relationship between the input and output. The details of the term generation procedures can be found in [10, 16, 17]. Note that usually the initial full model contains a huge number of candidate model terms to guarantee that the true or exact nonlinearities of the system to be studied can be sufficiently approximated using the specified model terms. However, more than often not all of the candidate model terms are equally important and useful. In practical, the final identified model only contains a relatively small number of the most significant model terms, meaning that the final model is sparse. This is important to avoid overfitting; otherwise, the model may show poor generalization ability and more importantly fail to reveal the key drivers. For example, for the case of the BMI prediction, it is crucial to identify the key bioimpedance frequencies that are consistently correlated with BMI. The NARX model employs an orthogonal forward regression (OFR) algorithm to select the most important model terms from all the candidate model terms. The OFR algorithm chooses the model terms in a stepwise manner, by ranking the contributions of the candidate model terms to explaining the output variable. The Error Reduction Ratio (ERR) is employed to measure the contribution of each term, which is defined as follows [10, 16, 27]:

$$ERR_i = \frac{[\boldsymbol{y}^T \boldsymbol{q}_i^{(s)}]^2}{(\boldsymbol{y}^T \boldsymbol{y})(\boldsymbol{q}_i^{(s)T} \boldsymbol{q}_i^{(s)})} \tag{4}$$

where $\boldsymbol{q}_i^{(s)}$ is the basis of the *i*-th model term at *s*-th selection iteration. The $\boldsymbol{q}_i^{(s)}$ is re-calculated and updated during each selection iteration by an orthogonalization procedure. The details of the OFR algorithm and term selection procedure can be found in [10,16,17]. Assume that a number of n model terms are selected, the final NARX model can be represented as:

$$\boldsymbol{y} = \theta_{l_1}\boldsymbol{\varphi}_{l_1} + \theta_{l_2}\boldsymbol{\varphi}_{l_2} + \cdots + \theta_{l_n}\boldsymbol{\varphi}_{l_n} + \boldsymbol{e} \tag{5}$$

where $\boldsymbol{\varphi}_{l_1}, \boldsymbol{\varphi}_{l_2}, \ldots, \boldsymbol{\varphi}_{l_n}$ are the selected terms and $\theta_{l_1}, \theta_{l_2}, \ldots, \theta_{l_n}$ are the estimated parameters.

In this study, the size of original datasets is limited (there are only a total of 135 samples), and the target system is highly nonlinear. Thus, a single NARX model may not sufficiently summary the underlying uncertainty, e.g., the different scales of bioimpedance will determine different correlation degree with BMI. Therefore, following the method in [28, 29] the original dataset is resampled for several times, in this way, more possible patterns within the system can be explored. The inputs and output of the k-th resampled sub-dataset can be represented as $y^{(k)}$ and $u_1^{(k)}, u_2^{(k)}, \ldots u_R^{(k)}$. Following the procedures of OFR algorithm, the NARX model for the k-th sub-dataset can be built as:

$$\boldsymbol{y}^{(k)} = \theta_{l_1}{}^{(k)}\boldsymbol{\varphi}_{l_1}{}^{(k)} + \theta_{l_2}{}^{(k)}\boldsymbol{\varphi}_{l_2}{}^{(k)} + \cdots + \theta_{l_n}{}^{(k)}\boldsymbol{\varphi}_{l_n}{}^{(k)} + \boldsymbol{e} \tag{6}$$

where $\theta_{l_1}{}^{(k)} \ldots \theta_{l_n}{}^{(k)}$ are the estimated weights of the selected terms $\boldsymbol{\varphi}_{l_1}{}^{(k)}, \ldots, \boldsymbol{\varphi}_{l_n}{}^{(k)}$ for the k-th sub-dataset.

These multiple NARX models are combined as a whole by designing and implementing the Takagi-Sugeno fuzzy inference. Compared to the other fuzzy inferences, Takagi-Sugeno fuzzy logic systems, due to their good inference properties, are able to integrate the polynomial NARX models well. Specifically, the linguistic form of a typical Takagi-Sugeno fuzzy rule reads as follows:

IF *x_1 is A_1, x_2 is A_2, … and x_n is A_n,* ***THEN*** $y = f(x_1, x_2, \ldots, x_n)$,

Where x_m represents the m-th (m=1,2, …, n) input variable for the antecedence, and in the consequent the output variable y is obtained from a linear function based on the n variables. $A_1, A_2, \ldots, A_n$ are the fuzzy labels with membership functions for calculating the intensity of the rule for final output. Meanwhile, the fuzzification of the inputs also expend the feasible range of each NARX models. For example, a Gaussian membership function can provide a membership degree over all feasible range rather

than a few limited points. This results in a soft partition between NARX models and a finer tuned surface for prediction.
In this study, the fuzzy rules are generated from the pattern extraction based on the NARX regression. For each fuzzy rule, a NARX model defines the output function in the consequent, and the corresponding training data ranges of each input features define the fuzzy labels and the correlated membership functions in the antecedence.

Table 1. Linguistic labels of selected features for TSF-NARX model

Linguistic Labels	$\boldsymbol{\varphi}_{l_1}$	$\boldsymbol{\varphi}_{l_2}$	$\boldsymbol{\varphi}_{l_3}$	$\boldsymbol{\varphi}_{l_4}$	$\boldsymbol{\varphi}_{l_5}$	$\boldsymbol{\varphi}_{l_6}$
Small	<145	<850	<900	<40.0	<200	<2500
Medium	145~165	850~1050	900~1100	40.0~60.0	200~300	2500~3500
Large	>165	>1050	>1100	>60.0	>300	>3500

Table 1 summaries the linguistic labels used for the model. The proposed fuzzy model uses the Gaussian basis function for representing the membership functions within the fuzzy rules.

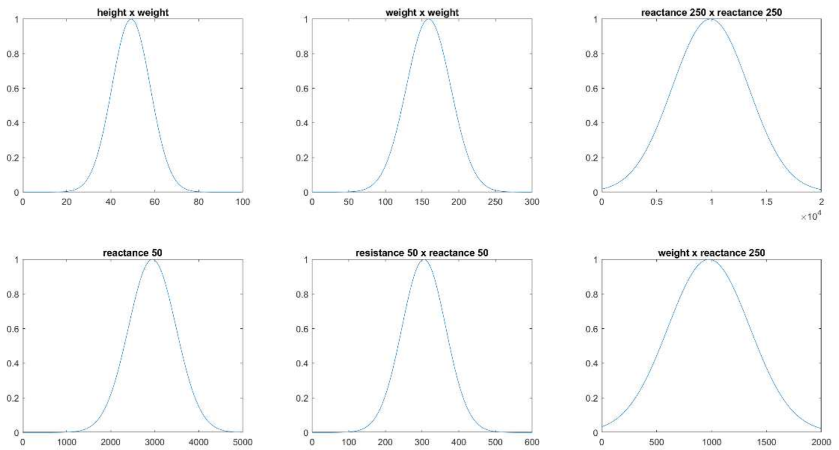

Figure 1. Membership functions for Rule no.5

A typical rule (rule no.5) implemented in the TSF-NARX reads as following:

IF $\boldsymbol{\varphi}_{l_1}$ *is medium and* $\boldsymbol{\varphi}_{l_2}$ *is medium and* $\boldsymbol{\varphi}_{l_3}$ *is medium and* $\boldsymbol{\varphi}_{l_4}$ *is medium and* $\boldsymbol{\varphi}_{l_5}$ *is medium and* $\boldsymbol{\varphi}_{l_6}$ *is medium,* ***THEN*** *BMI* $\boldsymbol{y}^{(5)} = \theta_{l_1}{}^{(5)}\boldsymbol{\varphi}_{l_1}{}^{(5)} + \theta_{l_2}{}^{(5)}\boldsymbol{\varphi}_{l_2}{}^{(5)} + \cdots + \theta_{l_n}{}^{(5)}\boldsymbol{\varphi}_{l_n}{}^{(5)}$,

and its corresponding membership functions are shown in the Figure 1.

Table 2. Data description

	variable	description	mean	min	max
y	BMI	Body Mass Index [kg]	55.1473	36.5900	74.9500
u1	height	height [m]	1.6216	1.4500	1.8000
u2	weight	weight [kg]	97.7400	56.2000	136.8000
u3	age	age [year]	44.8593	18.0000	69.0000
u4	R5	logarithm of resistance at 5 kHz	6.3090	5.9440	6.6840
u5	R10	logarithm of resistance at 10 kHz	6.2842	5.9220	6.6540
u6	R50	logarithm of resistance at 50 kHz	6.1761	5.8310	6.5390
u7	R100	logarithm of resistance at 100 kHz	6.1253	5.7900	6.4910
u8	R250	logarithm of resistance at 250 kHz	6.0516	5.7250	6.4170
u9	X5	reactance at 5 kHz	25.7164	9.9300	41.9200
u10	X10	reactance at 10 kHz	35.7914	18.8400	59.6900
u11	X50	reactance at 50 kHz	49.3887	29.8100	74.4100
u12	X100	reactance at 100 kHz	44.0784	26.1400	62.0200
u13	X250	reactance at 250 kHz	30.6726	17.1000	44.9700

4. Experimental Results

The data used in this study is described in Section 4.1. Section 4.2 presents the term generation and selection results. Section 4.3 shows the TSF-

NARX model and performance comparison. Discussions are in Section 4.4.

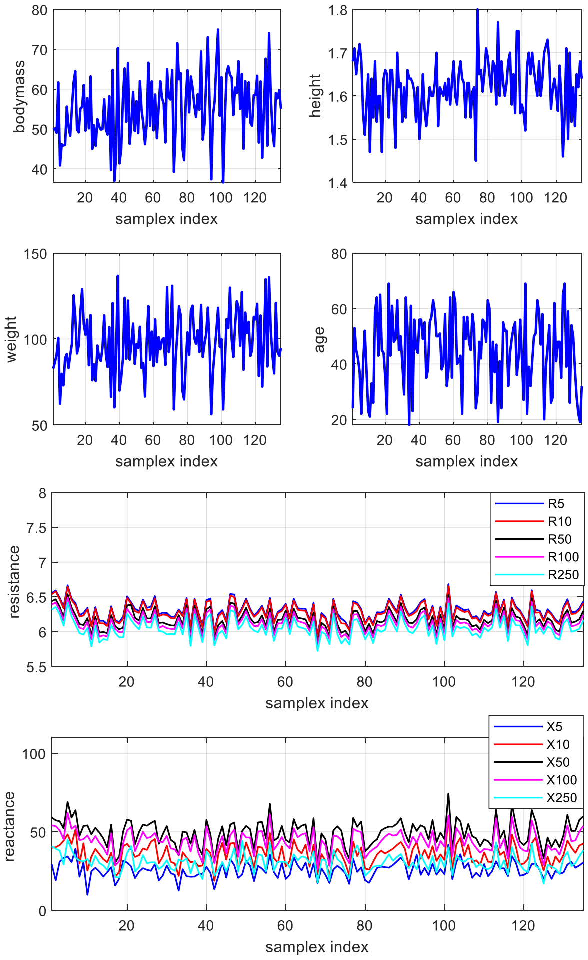

Figure 1. BMI and external variables

4.1 Data Description

We applied the proposed model on a BMI dataset with bioimpedance. The dataset was collected at the Food Science and Human Nutrition Research Unit of the Department of Experimental Medicine of Sapienza Rome University. A number of 135 data samples were obtained from overweight and obese women via dual-energy X-ray absorptiometry (DXA) examination (Hologic 4500 RDR) [18,19]. The BMI index was considered as model output, and the bioimpedance signals, along with some other external variables, were used as model inputs. Detailed descriptions of the data and variables are presented in Table 2. Bioimpedance data contains resistance and reactance signals, which were measured at frequencies 5, 10, 50, 100, 250 kHz, respectively. Height, weight and ages were used as external anthropometric variables for model construction [6].

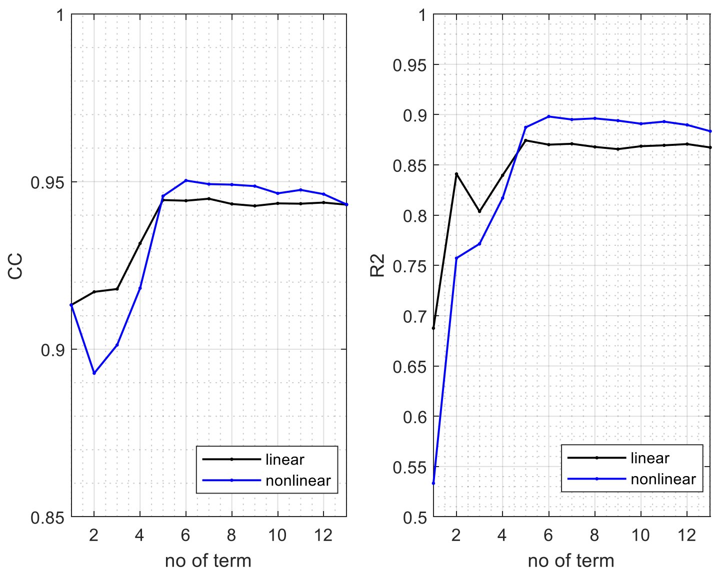

Figure 3. Linear models vs nonlinear NARX models

As can be seen from Table 2, the variation ranges of these variables are as follows. The age of participants ranges from 18 to 69 with an average of 45. The mean, minimum and maximum height is 1.62m, 1.45m and 1.80m, respectively. The cases within the dataset are comprehensive, covering a wide range of anthropometric conditions. A graphical illustration of seven variables is given in Figure 2, which presents large variation in each variable. No missing value or outlier was detected in the dataset. Around 80% of the samples were used for model training and the remaining 20% samples were used for testing the model performance.

4.2 The Identified NARX model

Previous studies have investigated the effectiveness of linear models [6]. However, the nonlinear effect among the input variables has not been analysed. To explore the impact of nonlinear interactions, we performed a number of trial experiments and compared the performances of linear models and nonlinear NARX models. First, we applied the OFR (orthogonal forward regression) algorithm to generate a number of nonlinear NARX models with different number of model terms. Second, we built the associated linear models using exactly the same input variables. The only difference is that the nonlinear NARX models contain nonlinear model terms, but the linear models only use original linear inputs. The performances of the models were evaluated with the correlation coefficient (CC) and the prediction efficiency (PE) as presented in Figure 3. For the definition of PE, interested readers are referred to [21]. From the results, the linear models have better performances when the model contains up to 4 model terms. However, as the number of terms increases, the nonlinear NARX models achieve significantly better performances than linear models. This may be understood that there are indeed some nonlinear model terms (interaction product model terms) that cab better explain the variation of BMI. In other words, the value of BMI nonlinearly depends on these specified variables.

Table 3. The identified terms, the associated ERR values, estimated weights and t values.

Terms	Description	Weights	ERR	T value
'u1×u2'	height × weight	0.4143	99.3188	15.0239
'u2×u2'	weight × weight	-0.0015	0.2939	4.6767
'u13×u13'	reactance 250kHz × reactance 250kHz	0.0023	0.0199	0.9794
'u11'	reactance 50 kHz	2.8189	0.0275	7.7502
'u6×u11'	resistance 50kHz × reactance 50 kHz	-0.4219	0.1080	7.3365
'u2×u13'	weight × reactance 250 kHz	-0.0029	0.0102	2.1716

It can be observed that the performances vary with the number of terms. A too simple model (e.g., with a too small number of terms) cannot sufficiently represent the relation between the inputs and output, whereas the inclusion of excessive terms increases the model complexity, leading to overfitting and deteriorating the model's generalization ability. To identify the optimal number of terms, we applied the term length selection criteria called adjustable prediction error sum of squares (APRESS) [20]. According to the APRESS, the optimal number of the model terms is 6. We applied the OFR algorithm and identified 6 most effective model terms, see Table 3 for details.

The associate NARX model in Table 3 reads as:

$$y = 0.4143 \times u1 \times u2 - 0.0015 \times u2 \times u2 + 0.0023 \times u13 \times u13 + 2.8189 \times u11 - 0.4219 \times u6 \times u11 - 0.0029 \times u2 \times u13 \quad (7)$$

The ERR values indicate the importance of the selected terms. Note that the ERR value of the first term is much higher than those of the following terms. This indicates that the combined effect of the two factors, height and weight, plays a dominant role in determining a person's BMI value. The finding here accords closely with that currently used in hospitals, i.e., BMI = weight/(height square), where the units of weight and height are kg and m, respectively.

The significance of the selected terms was validated with the student t-test, which is used to determine if there is a significant difference between two groups [16]. The confidence interval is chosen to be 95%. If the t value is larger than 1, it means that this term contributes significantly to the regression model. From the results, all but one term are significant, and the t value of the term $u13 \times u13$ is slightly smaller than (but very close to) 1; it is reasonable to keep the term in the model.

From the results, some insights can be revealed from the identified model. First, the selected terms indicate the most impactful factors on BMI. For example, $u1 \times u2$ and $u2 \times u2$ indicates that the interactions between height and weight and square of weight significantly affect BMI. This finding is in line with previous study which explored the effectiveness of anthropometric variables [6]. In addition, the model reveals the key bioimpedance frequency bands that impact BMI. For example, the selected terms $u11$, $u6 \times u11$ and $u2 \times u13$ are derived from the variable reactance at 50 kHz, resistance at 50kHz, and reactance at 250 kHz, indicating that these are the most important frequency bands.

4.3 TSF-NARX Results

Figure 4 shows the surface plots for TSF-NARX model membership functions. The BMI roughly increases with the values of height, weight and the reactance at 50 kHz and decreases with the reactance at 250 kHz. It can be observed that the surface is not flat. Therefore, these selected features are correlated with each other to a certain degree, and the relationships between them and BMI are nonlinear.

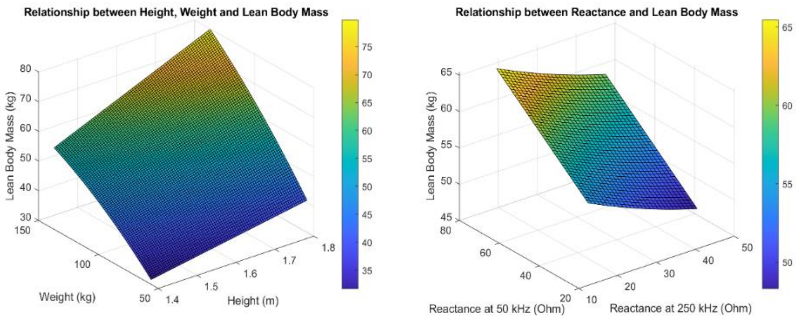

Figure 4. Membership function for height & weight, reactance at 50kHz & reactance at 250 kHz of fuzzy-NARX model

Table 4 summarises the prediction performances of all models measured using the following three metrics: correlation coefficient (CC), coefficient of determination (R2) and root mean square error (RMSE). CC measures the correlation between the target observations and the model predictions, R2 describes the match between the predictions and the observations, whereas RMSE measures the overall residuals between the estimations and the observations.

For comparison purpose, we tested the linear model, the NARX model, a neural network model and adaptive neuro fuzzy inference system (ANFIS) model on the same test dataset. The linear model was constructed by using all the original input variables listed in Table 2. The NARX model was constructed by the 6 selected important terms. The ANFIS model was generated from Matlab built-in function with default settings. As this is a static system which does not contain time series data, the tested neural network was constructed by one input layer, one output layer and several fully connected layers. ANFIS combines the fuzzy logic interference with neural network, and the fuzzy rules of the system are extracted from the data using the conventional neural network structure. The fuzzy rule-base of the ANIFS in this study was generated with subtractive clustering (range of influence 0.5, squash factor 1.25, accept ratio 0.5 and reject ratio 0.15), and then the fuzzy inference system was trained with hybrid optimisation (error tolerance 0 and epochs 3).

Table 4. The summary of model performances

Model	CC	R2	RMSE
linear model	0.9417	0.8624	2.8013
NARX model	0.9503	0.8981	2.4653
TSF-NARX model	0.9484	0.8937	2.4620
neural network	0.9417	0.0000	57.4197
ANFIS	0.9153	0.7191	4.0012

As shown in Table 4, the nonlinear NARX and the TSF-NARX outperform the other models in terms of all the three metrics. It suggests that data relationship of such a biological dataset is severely nonlinear, thus, its modelling requires nonlinear elements. Meanwhile, compared to the NN and ANFIS, the NARX and the TSF-NARX models can deal with the small sample size problem very well. NARX model has achieved the highest CC and R2 while TSF-NARX has achieved the lowest root mean square error. This indicates that NARX is more specialised in the pattern extraction with existing data, whereas the TSF-NARX focuses on exploring the boundary of feasible data range.

Figure 5 summaries the residuals between the predictions and observations for the ANFIS and the TSF-NARX model. Compared to the ANFIS, the residuals of the TSF-NARX model are closer to the desired value 0 and are evenly distributed on both negative and positive side. It suggests that the TSF-NARX model has relatively high accuracy and high precision. Figure 6 shows the relations between the predictions and observations for the ANFIS and TSF-NARX model. It can be observed that compared to the ANFIS. the TSF-NARX model can provide more accurate prediction with less bias in general. This indicates that the pattern extraction based on the NARX model is more efficient and robust than the subtractive clustering method in this small sample data case.

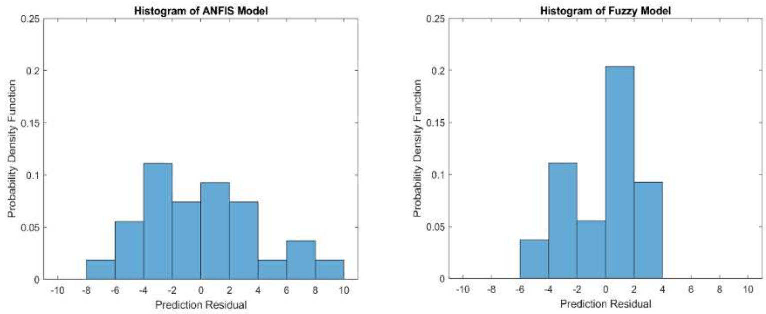

Figure 5. Residual histogram of ANFIS and TSF-NARX

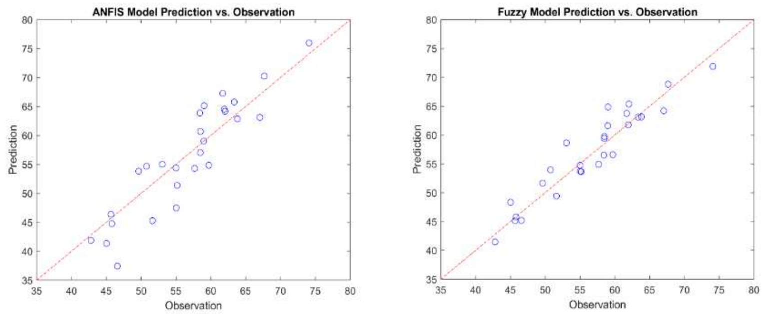

Figure 6. Scatter plot of ANFIS and TSF-NARX

The experimental results show that the TSF-NARX provides excellent generalization properties. The use of the Takagi-Sugeno fuzzy rule-base interference created a simple and intuitive linguistic-based model that can be easily interpreted by end-users of the model. The relationships within the features can be directly displayed in the surface plots as shown earlier and be implemented into quick estimation. Meanwhile, the fuzzy inference expands the prediction range with nonlinear membership functions of the rules. In addition, it ensures that the model is robust to the intra and extra uncertainty within the system. Therefore, fuzzy logic modelling approaches can normally achieve excellent performance even with small sample size. However, such models are heavily dependent on the quality of rule-base. Thus, the pattern extraction or the clustering method is important for fuzzy logic-based models and is of great importance to the current research especially for developing the associated fuzzy logic. Compared to the ANFIS, the pattern extraction of NARX has achieved better results than the neural network model and explored the full potential of fuzzy inference. Therefore, the combination of NARX and fuzzy logic has proved to be the most efficient and effective approach to model the clinical data.

5 Conclusion

This work proposed a novel hybrid TSF-NARX model for BMI prediction. It identified the key factors for the BMI measurement, i.e., reactance at 50 kHz, resistance at 50kHz, reactance at 250 kHz, weight and height. The prediction accuracy of the proposed model outperformed other existing state of the art models. The integration of fuzzy logic inference provides a model structure that can be fully interpreted by human, and this is important for further applications. Furthermore, such a paradigm creates the foundation for more advanced predictive approach for medical research and clinical diagnosis. The proposed method can be applied to the exploration of the underlying mechanism of complicated data modelling problems in other areas such as environment and space weather.

Acknowledge
This work was partially supported by Innovate UK under grant reference 26526.

References

[1] Shea, J.R., Henshaw, M.H., Carter, J. and Chowdhury, S.M., 2020. Lean body mass is the strongest anthropometric predictor of left ventricular mass in the obese paediatric population. Cardiology in the Young, 30(4), pp.476-481.
[2] Schultz, S.R. and Johnson, M.K., 1995. Effects of birth date and body mass at birth on adult body mass of male white-tailed deer. Journal of Mammalogy, 76(2), pp.575-579.
[3] Jaquish, C.E., Dyer, T., Williams-Blangero, S., Dyke, B., Leland, M. and Blangero, J., 1997. Genetics of adult body mass and maintenance of adult body mass in captive baboons (Papio hamadryas subspecies). American journal of primatology, 42(4), pp.281-288.

[4] Zuercher, G.L., Roby, D.D. and Rexstad, E.A., 1999. Seasonal changes in body mass, composition, and organs of northern red-backed voles in interior Alaska. Journal of Mammalogy, 80(2), pp.443-459.
[5] Borga, M., West, J., Bell, J.D., Harvey, N.C., Romu, T., Heymsfield, S.B. and Leinhard, O.D., 2018. Advanced body composition assessment: from body mass index to body composition profiling. Journal of Investigative Medicine, 66(5), pp.1-9.
[6] Cammarota, C. and Pinto, A., 2021. Variable selection and importance in presence of high collinearity: an application to the prediction of lean body mass from multi-frequency bioelectrical impedance. Journal of Applied Statistics, 48(9), pp.1644-1658.
[7] Kyle, U.G., Bosaeus, I., De Lorenzo, A.D., Deurenberg, P., Elia, M., Gómez, J.M., Heitmann, B.L., Kent-Smith, L., Melchior, J.C., Pirlich, M. and Scharfetter, H., 2004. Bioelectrical impedance analysis—part I: review of principles and methods. Clinical nutrition, 23(5), pp.1226-1243.
[8] Kyle, U.G., Bosaeus, I., De Lorenzo, A.D., Deurenberg, P., Elia, M., Gómez, J.M., Heitmann, B.L., Kent-Smith, L., Melchior, J.C., Pirlich, M. and Scharfetter, H., 2004. Bioelectrical impedance analysis—part II: utilization in clinical practice. Clinical nutrition, 23(6), pp.1430-1453.
[9] Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R. and Chatila, R., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, pp.82-115.
[10] Billings, S.A., 2013. Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains. John Wiley & Sons.
[11] Lundberg, S.M. and Lee, S.I., 2017, December. A unified approach to interpreting model predictions. In Proceedings of the 31st international conference on neural information processing systems (pp. 4768-4777).
[12] Mendel, J.M., 2017. Uncertain rule-based fuzzy systems. Introduction and new directions, pp.684.
[13] Rathore, H., Mohamed, A., Guizani, M. and Rathore, S., 2021. Neuro-fuzzy analytics in athlete development (NueroFATH): a machine learning approach. Neural Computing and Applications, pp.1-14.
[14] Keivanian, F., Chiong, R. and Hu, Z., 2019, April. A fuzzy adaptive binary global learning colonization-MLP model for body fat prediction. In 2019 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART) (pp. 1-4). IEEE.

[15] Ribeiro, A.C., Silva, D.P. and Araujo, E., 2014, July. Fuzzy breast cancer risk assessment. In 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (pp. 1083-1087). IEEE.
[16] Wei, H.L. and Billings, S.A., 2008. Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information. International Journal of Modelling, Identification and Control, 3(4), pp.341-356.
[17] Wei, H.L., Billings, S.A. and Liu, J., 2004. Term and variable selection for non-linear system identification. International Journal of Control, 77(1), pp.86-110.
[18] Ellis, K.J., 2000. Human body composition: in vivo methods. Physiological reviews, 80(2), pp.649-680.
[19] Earthman, C.P., 2015. Body composition tools for assessment of adult malnutrition at the bedside: a tutorial on research considerations and clinical applications. Journal of Parenteral and Enteral Nutrition, 39(7), pp.787-822.
[20] Billings, S.A. and Wei, H.L., 2008. An adaptive orthogonal search algorithm for model subset selection and non-linear system identification. International Journal of Control, 81(5), pp.714-724.
[21] Gu, Y., Wei, H.L., Boynton, R.J., Walker, S.N. and Balikhin, M.A., 2019. System identification and data-driven forecasting of AE index and prediction uncertainty analysis using a new cloud-NARX model. Journal of Geophysical Research: Space Physics, 124(1), pp.248-263.
[22] Spinelli, W., Piroddi, L. and Li, K., 2004, December. Nonlinear modeling of NO/sub x/emission in a coal-fired power generation plant. In 2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601) (Vol. 4, pp. 3850-3855). IEEE.
[23] Gu, Y., Yang, Y., Dewald, J.P., Van der Helm, F.C., Schouten, A.C. and Wei, H.L., 2020. Nonlinear modeling of cortical responses to mechanical wrist perturbations using the narmax method. IEEE Transactions on Biomedical Engineering, 68(3), pp.948-958.
[24] Aguirre, L.A., 2019. A Bird's Eye View of Nonlinear System Identification. arXiv preprint arXiv:1907.06803.
[25] Billings, S. and Wei, H.L., 2019, August. NARMAX model as a sparse, interpretable and transparent machine learning approach for big medical and healthcare data analysis. In 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International

Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 2743-2750). IEEE.
[26] Wei, H.L., 2019, June. Sparse, interpretable and transparent predictive model identification for healthcare data analysis. In International Work-Conference on Artificial Neural Networks (pp. 103-114). Springer, Cham.
[27] Chen, S., Billings, S.A. and Luo, W., 1989. Orthogonal least squares methods and their application to non-linear system identification. International Journal of control, 50(5), pp.1873-1896.
[28] Wei, H.L. and Billings, S.A., 2009. Improved model identification for non-linear systems using a random subsampling and multifold modelling (RSMM) approach. International Journal of Control, 82(1), pp.27-42.
[29] Gu, Y. and Wei, H.L., 2018. A robust model structure selection method for small sample size and multiple datasets problems. Information Sciences, 451, pp.195-209.