 doi:10.20944/preprints201812.0009.v1

Article

Using the PI ProcessBook to Monitor Activities of Daily Living in Smart Home Care within IoT

Jan Vanus[1,*], Jan Kubicek[1] , Ojan Gorjani[1], Jiri Koziorek[1]

[1] Department of Cybernetics and Biomedical Engineering, Faculty of Electrical Engineering and Computer Science, VSB–Technical University of Ostrava, Ostrava, Czech Republic; jan.vanus@vsb.cz; jan.kubicek@vsb.cz; ojan.majidzadeh.gorjani.st@vsb.cz, jiri.koziorek@vsb.cz

* Correspondence: jan.vanus@vsb.cz; Tel.: +420-59-732-5856

Abstract: This article describes the use of the PI ProcessBook software tool for visualization and indirect monitoring of occupancy of SHC rooms from the measured operational and technical quantities for monitoring of daily living activities for support of independent life of elderly persons. The proposed method for data processing (predicting the CO_2 course using neural networks from the measured temperature indoor T_i (°C), temperature outdoor T_o (°C) and the relative humidity indoor rHi (%)) was implemented, verified and compared in MATLAB SW tool and IBM SPSS SW tool with IoT platform connectivity. Within the proposed method, the Stationary Wavelet Transform de noising algorithm was used to remove the noise of the resulting predicted course. In order to verify the method, two long-term experiments were performed, (specifically from February 8 to February 15, 2015, from June 8 to June 15, 2015) and two short-term experiments (from February 8, 2015 and from June 8, 2015). For the best results of the trained ANN BRM within the prediction of CO_2, the correlation coefficient R for the proposed method was up to 90%. The verification of the proposed method confirmed the possibility to use the presence of persons of the monitored SHC premises for rooms ADL monitoring.

Keywords: Smart Home Care (SHC); monitoring; prediction; trend detection, Artificial Neural Network (ANN), Bayesian Regulation Method (BRM), Wavelet Transformation (WT), SPSS (Statistical Package for the Social Sciences) IBM, IoT (Internet of Things), Activities of Daily Living (ADL).

1. Introduction

At present, there is an increase in the use of modern technologies in the field of building automation (in accordance with European legislation of Directive 2010/31/EU on the energy performance of buildings and in accordance with Directive 2012/27/EU on energy efficiency and changes thereof). In order to provide a user-friendly environment for the management of the operational and technical functions along with providing support for the independent housing of senior citizens and disabled persons in buildings indicated as Smart House Care (SHC), it is necessary to make appropriate visualization of the technological process as required by the users with the possibility of indirect monitoring of the seniors' life activities based on the information obtained from the sensors used for the common management of the operational and technical functions in SHC. The properly devised visualization complements the final correct functionality of the intelligent building. Beaudin describes the use of computational and sensor technology in intelligent buildings with a focus on health monitoring in the households. The work comprises a visualization for displaying health data and a proposal for improving the health and wellbeing of the users [1]. Booysen explores machine-machine communication (M2M) to address the need for autonomous control of remote and distributed mobile systems in intelligent buildings [2]. Basu uses miniature wireless sensors in the

 doi:10.20944/preprints201812.0009.v1

wireless network to track and recognize the behaviour of persons in the house. The visualization includes sensor data from the building to capture the duration of the activities [3]. Fleck describes a system based on intelligent cameras for 24-hour monitoring and supervision of senior citizens. On this occasion, visualization is used to display relevant life information in this intelligent environment, which includes evaluation of the seniors' position and recognition of life activities [4]. The current trend for processing large volumes of measured quantities using Soft Computing Methods (SC) [5-6] is to use the available Big Data Analysis tools within the IoT platform [7-8]. The Internet of Things, shortly known as IoT, can be assumed as an integration layer which creates an interconnection of several physical devices, sensors, actuators and controllers [9]. In simple words, the IoT allows objects other than computers or smartphones to use the Internet for sending and receiving information [10] 9). The number of connected IoT devices is rapidly growing. This growth could be the result of a very wide range of applications, ranging from basic home appliances and security systems to more sophisticated applications. For example, Xu et al. used IoT in order to construct a real-time system monitoring for micro-environment parameters such as temperature, humidity, PM10 and PM2.5 [11]. Q. Min et al. suggested using IoT for monitoring of discrete manufacturing process based on IoT [12]. Wang et al. presented a feasible and reliable plantation monitoring system based on Internet of Things, that combined wireless sensor network, embedded development, GPRS communication technology, web service, and Android mobile platform [13]. Windarto et al. represent an application of IoT by implementing automation of lights and door in a room [14]. Coelho et al. used IoT to collect data from multiple heterogeneous sensors that were providing different types of information a variety of locations in a smart home [15]. Data collection in this kind of examples usually results in big data. The Oxford dictionary defines big data as “Extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions” [16]. One of the possibilities that come with big data is a predictive analysis which can provide predictions about the future or otherwise unknown events. The predictive analysis offers a wide range of applications such as social networking, healthcare, mobility, insurance, finance marketing etc. Nyce suggested to use the predictive analysis for risk identifications and probabilities in order to provide an appropriate insurance rate, his method took advantage of marketing records, underwriting records, and claims records [17]. The predictive analysis includes many different statistical techniques ranging from data mining, predictive modeling to machine learning. Predictive modeling may be applied to many areas such as weather forecasting, business, Bayesian spam filters, advertising and marketing, fraud detection etc. Predictive models are based on variables that are most likely to influence the outcome. [18] These variables are also known as predictors. There are many types of predictive models, such as neural networks and Decision trees. Ahmad et al. compared different methods of predictive modeling for solar thermal energy systems such as random forest, extra trees, and regression trees [19]. The IBM offers a variety of services in terms of predictive analysis such as Watson analytics and SPSS. Additionally, these services can be integrated with IBM cloud (formerly known as Bluemix) with other IBM tools and services, including Artificial Intelligence (AI) and the Internet of things (IoT) [20]. One of the possibilities with IBM services is the Watson platform, which includes services such as Watson Analytics, Watson Machin learning, Watson assistance etc. Watson Analytics is a powerful tool that takes advantages of machine learning and AI technologies combined with natural language querying that is also intuitive and easy to use. The only shortcoming of this platform is a lack of real-time data streaming [21]. There is a wide range of IBM Watson Analytics applications, for example, Nagwanshi, et al. used IBM Watson Analytics processes for distinguishing features of human footprint images [22]. IBM SPSS Statistics comes as a software package which is mainly used for statistical analysis. It is considered as one of the world’s leading statistical software which offers advanced statistical analysis, a vast library of machine-learning algorithms, text analysis, open-source extensibility, integration with big data and seamless deployment into applications. SPSS can be driven externally by a Python or a VB.NET program using supplied "plug-ins" which provides endless opportunities for custom solutions [23]. Arnold investigates the IBM's Watson approach, combines the used techniques with other artificial intelligence methods and develops an interactive concierge service for independent living of elderly

 doi:10.20944/preprints201812.0009.v1

persons in context easy-to-use interaction models based on natural language processing development [24]. Carvalko describes Medical technology verges on incorporating directly into our anatomy processors with the computational power of the Watson IBM computer and Internet-like communications within the eventual merging of synthetic DNA and artificial intelligence that together will bring new diagnostics, medical treatment, and smart nano-prosthetics well within the horizon of the next generation [25]. The goal of Cervenka is to define various ways to apply cognitive systems for unstructured data analysis and management for further use in marketing analytics with IBM Analytics Tools for analyzing unstructured data found on social media [26]. Chen answers the question how cognitive computing can be applied to big data challenges in life sciences research with IBM Watson for training to understand technical, industry-specific content and use advanced reasoning, predictive modeling, and machine learning techniques to advance research faster [27]. Coccoli goes in the direction of developing the smart university model, by using innovative and intelligent services to help to raise a new generation of software engineers but also to promote and disseminate a new way for designing and building innovative applications through hands-on labs on cognitive computing with a software development environment based on the Platform as a Service (PaaS) [28]. Devarakonda developed problem-oriented Electronic Medical Records EMR summarization to address this issue, as a part of a larger effort of adapting IBM Watson to the medical domain with presentation building the next generation EMR, one that is based not on just keeping record but instead on a conceptual understanding of medicine, thereby crossing the threshold from record storage to an intelligent entity for clinical decision making [29]. Guidi describes the main tools available on the market to perform Analytics as a Service (AaaS) using a cloud platform with a use case of IBM Watson Analytics, a cloud system for data analytics, applied to the following research scope: detecting the presence or absence of heart failure disease using nothing more than the electrocardiographic signal, in particular through the analysis of Heart Rate Variability [30]. Kolker describes the PPT-DAM (People-Process-Technology empowered by Data, Analytics, and Metrics) implementation by the Benchmarking Improvement Program at the Seattle Children's Hospital with the results, that cognitive systems in general and IBM Watson in particular, if properly implemented, can bring transformative and sustainable capabilities in healthcare far beyond the current ones [31]. Murtaza discusses an effective strategy to train IBM Watson question answering system for Big Data Analytics where they have observed that if documents are well segmented, contain relevant titles and have consistent formatting, then the recall of the answers can be as high as 95% [32]. AlFaris reviews the smart technologies; the interface and integration of the meters, sensors and monitoring systems with the home energy management system (HEMS) within IoT with the outline that the smart home in practice provides the ability to the house to be net-zero energy building. Especially that it reduces the power demand and improve the energy performance by 37% better than ASHRAE standards for family villas sector [33]. Alirezaie presents a framework called E-care@home, consisting of an IoT infrastructure, which provides information with an unambiguous, shared meaning across IoT devices, end-users, relatives, health and care professionals and organizations and demonstrates the proposed framework using an instantiation of a smart environment that is able to perform context recognition based on the activities and the events occurring in the home [34]. Bassoli introduced a new system architecture suitable for human monitoring based on Wi-Fi connectivity, where the proposed solution lowers costs and implementation burden by using the Internet connection that leans on standard home modem-routers, already present normally in the homes, and reducing the need for range extenders thanks to the long range of the Wi-Fi signal with energy savings of up to 91% [35]. Catherwood presents an advanced Internet of Things point-of-care bio-fluid analyzer; a LoRa/Bluetooth-enabled electronic reader for biomedical strip-based diagnostics system for personalized monitoring, where are solved highlights practical hurdles in establishing an Internet of Medical Things network, assisting informed deployment of similar future systems [36]. The source of CO_2 concentration in residential buildings is primarily the human being [37]. Due to the high cost of SHC sensor equipment, there is an effort to reduce investment costs by using appropriate SC methods, while using available information from existing sensors to ensure management of the operational and technical functions in SHC.

 doi:10.20944/preprints201812.0009.v1

In this article, to monitor the presence of persons in SHC rooms, a method for predicting the CO_2 course using the multi-layered forward ANN using the BRM from the measured indoor temperature values T_i (°C) and the relative humidity indoor rH_i (%), with the subsequent use of trend signal detection based on WTA is revised and verified [38].

The objective of the article is the performance of a new method for prediction CO_2 course from the measured temperature variables T_{indoor} (°C), the relative humidity rH_{indoor} (%) and the temperature $T_{outdoor}$ (°C) within monitoring the presence of persons in an SHC room using the SW Tool PI System (PI—Plant Information enterprise information system) and SPSS IBM IoT platform SW Tool.

The main contribution of the article is to verify the assumption of improving the accuracy of the CO_2 prediction method using ANN BRM with the Wavelet Transformation filter algorithm to remove additive noise from the predicted CO_2 course.

The aim of the first part of the article is the use and processing of information from operationally measured non-electrical quantities determining the indoor environment in the SHC using operational technological units for the determination of the ADL in a real-world SHC environment. To obtain an overview of the ADL of individual rooms of the SHC (time of arrival, time of departure), the indirect measurement of CO_2 concentration (ppm) with operational CO_2 (ppm) sensors is used.

The aim of the second part of the article is to use ANN BRM to predict the measured quantities for the purpose of monitoring the ADL in a real-world SHC environment. It describes the process of using the multilayer forward ANN to predict the course of CO_2 concentration from the measured temperature Ti (°C), relative humidity rH (%) in the interior of the SHC in the selected room R204 and from the outdoor temperature readings To (°C), with the gradient algorithm of error backpropagation using the Bayesian Regulation Algorithm (BRA) prediction. For the classification of prediction quality, a correlation analysis (correlation coefficient R), calculated MSE (Mean Squared Error), Euclidean distance (ED), City Block distance (CB) are used.

The aim of the third part of the article is verification and comparison of WT additive noise cancelation in ANN BRM prediction of the course of CO_2 concentration from the measured temperature Ti (°C), relative humidity rH (%) in the interior of the SHC in the selected room R204 and from the outdoor temperature readings To (°C). For the classification of prediction quality with WT additive noise cancelation, a correlation analysis (correlation coefficient R), calculated MSE (Mean Squared Error), Euclidean distance (ED), City Block distance (CB) are used.

The aim of the fourth part of the article is to use ANN to predict the measured quantities for the purpose of monitoring the ADL in a real-world SHC environment using of SPSS IBM SW Tool within IoT platform implementation. To predict the course of CO_2 concentration from the measured temperature Ti (°C), relative humidity rH (%) in the interior of the SHC in the selected room R204 and from the outdoor temperature readings To (°C), with the ANN Radial Basis Function (RBF) network, which is a feed-forward network that requires supervised learning. For the classification of prediction quality, a correlation analysis (correlation coefficient R), calculated MSE (Mean Squared Error) and Mean Absolute Error (MAE) are used.

2. Decription of used technologies in SHC

SHC is located at VŠB-TU Ostrava in Czech Republic and was created on the basis of a requirement for creating a platform where research and monitoring of current technologies in the field of intelligent buildings could be conducted. The purpose of this platform is, in particular, to verify the efficient energy management and the reduction of the operational costs of the building in real conditions. This is a two-storeyed wooden building with an area of 12.1 x 8.2 m, built in a passive standard. The BACnet technology, which works on the BACnet/DESIGO PX PXC 100-E.D process station, is used for HVAC (Heating, Ventilation and Air Conditioning) management. On the other hand, the KNX technology provides control of lighting, blinds and switching on and off of the socket circuits in the building. The existing security electronic system (EZS) is connected to the KNX technology via binary inputs [24]. To visualize the operational and technical functions in SHC, the Desigo Insight software tool, which contains a database for storing measured values from individual technologies, is used in this building. This software also provides a visualization section for

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

displaying control and regulation of individual functions in the building. However, the Desigo Insight software tool is not suitable for the actual processing of the measured data. Therefore, the OPC BACnet server was used to implement connectivity between the PI System SW tool and the Desigo Insight software tool (Fig. 1). The PI System includes SW tools such as PI ProcessBook for user-friendly data readout with the ability to create an application for visualization and monitoring of SHC residents activities (Fig. 2, Fig. 3) [39-40].

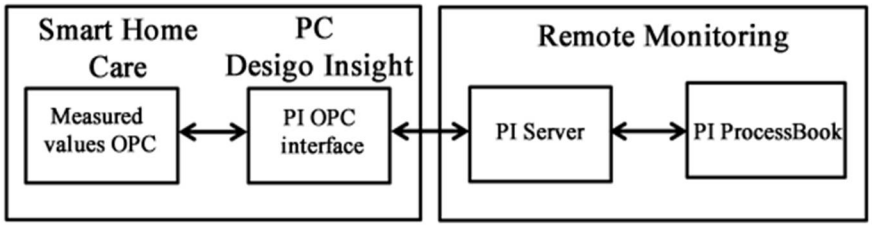


Figure 1. A block diagram of data transfer from the PI OPC interface to the PI Server and PI ProcessBook.

2.1. Visualization for the SHC created in the PI ProcessBook tool

Visualization of the wooden house in the PI ProcessBook tool is divided into several screens, which is possible continuously access from the main screen (Fig. 2). The SHC control technology is integrated on each visualization screen in accordance with the Building Management System (BMS). These screens further comprise buttons for entry into individual rooms. After clicking on the relevant room button, a new window, containing a detailed description of the technology used in the specific room shown in the individual charts, will appear.

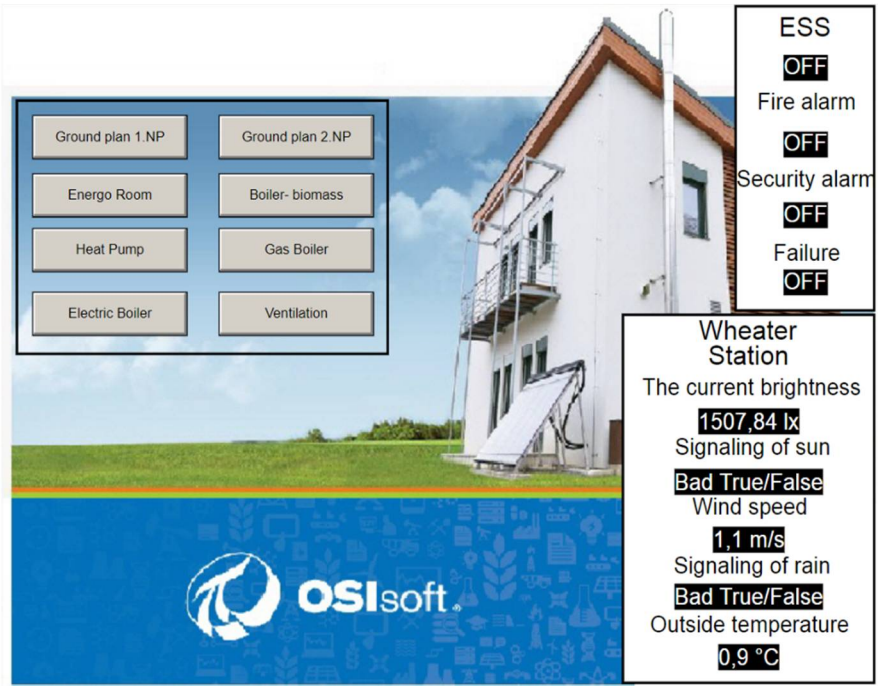


Figure 2. The main visualization screen in SW tool PI ProcessBook.

Each technological element is illustrated in the chart, wherein the individual charts are sorted in accordance with the groups of elements used (Fig. 3). The individual technology units are then displayed on separate screens and can also be viewed from the main screen. This solution was chosen because it is not possible to place all the information about the technologies implemented on one screen so that the screen remained well-arranged. Additional distribution of the technologies into individual screens will allow the user to get a better insight into what elements belong to the individual technologies and which do not anymore. Thanks to this solution selected, orientation in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

the enclosed charts are easier as well as the analysis of the individual quantities and actions in the building.

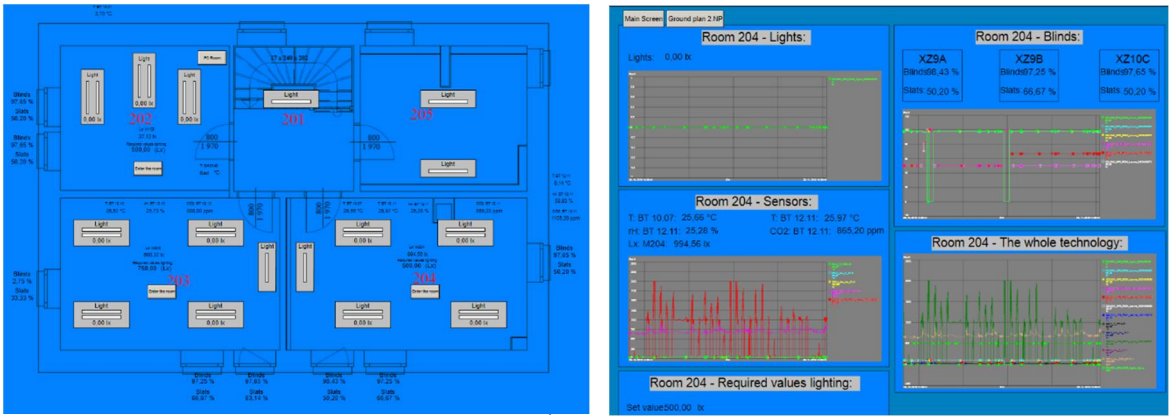


Figure 3. A screen for room 204 with a detailed description of individual technologies and individual charts in SW tool PI ProcessBook.

The visualization, monitoring, and processing of the measured values of non-electric variables, such as measurement of temperature, humidity, CO_2 for monitoring the quality of the indoor environment of the selected room in the building described, are implemented using the PI System software application and MATLAB R2014b SW Tool. Figure 4 shows the process of measuring and processing data in connection with the use of the PI System SW tool (PI ProcessBook) in the SHC.

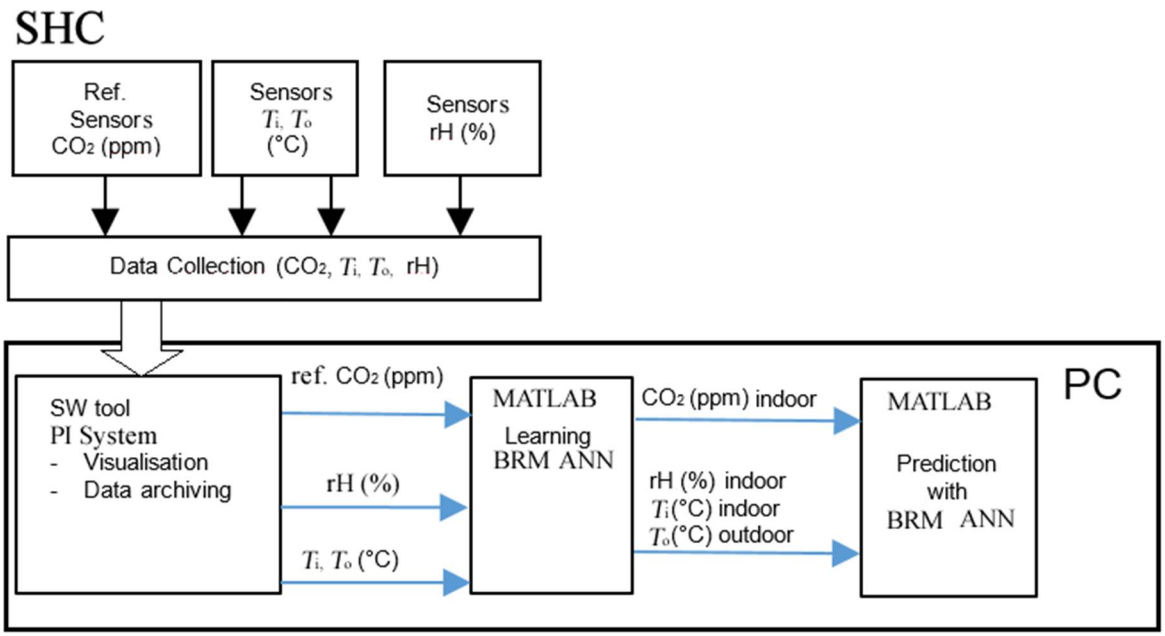


Figure 4. Simplified block scheme of utilized technology in SHC with prediction of CO_2.

3. Proposed Method for building optimized model of CO2 concentration

3.1. Bayesian regularization backpropagation algorithm - Backpropagation algorithm

Back propagation or the error back-propagation algorithm is a gradient algorithm for the adaptation of multi-layer feed forward networks. Organizational dynamics in the error back-propagation is equivalent to the multi-layer Perceptron. The network is formed by fixed acyclic topology with an input and output layer and one or two inner layers of hidden neurons. A universal solution for the number of neurons in the hidden layer so that the computing time is minimized and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

provides the most accurate result, has not yet been found. There is Bayes Regulation Method (BRM) offered for an optimizing training method. The BRM differs from the previous two in that it does not work with the validation set. It can be said about BRM in general that it generalizes small or complex datasets or datasets containing noise rather well. The training ends with the minimizing of weights. In the next step the neural network itself is set up, specifically the number of neurons in the hidden layer. The parameters of the neural network are selected sequentially with 10, 50, 100, 150, 200, 250, 300, 350, 400, 450 and 500 neurons in the hidden layer (Tables 1-8).

For the hidden layer, the hyperbolic tangent function was used and linear functions for the output layer. (Tables 1-8) lists the calculated parameters of MSE, correlation coefficient R, Euclidean distance (ED), City Block distance (CB) in order to compare the quality of learning of BRM. (Tables 1-8) shows that the results gradually improve with an increasing number of hidden neurons. Although this behaviour does not represent a rule in the field of neural networks (it often happens that the setup of algorithms and data and the results are not unconditionally in direct proportion), ANN (BRM) using measurement data exhibit behaviour naturally expected for the task. The lowest values of MSE and at the same time the largest values of correlation coefficients R are shown in BRM. The disadvantage for BRM is the relatively long time in learning ANN. The designed application for monitoring technological processes in the SHC can be used for obtaining information about the current state of the measured variables from common operating sensors (temperature T, rH, CO_2) for further practical applications in the context of operational and technical functions in the SHC.

The text further describes the procedure for determining the appropriate method of predicting the course of CO_2 concentration from the measured values taken by the indoor temperature sensor T_i (°C), (QPA 2062) in an IAB office room (range 0 to 50°C/-35 to 35°C, accuracy ± 1K) and relative humidity rH (%), (QPA 2062) (range 0 and 100%, accuracy ± 5%), outdoor temperature T_o (°C), (AP 257/22), (range: - 30 ... + 80°C, resolution: 0.1°C) using a gradient algorithm of backpropagation for adapting the multilayered forward ANN (Fig. 5) using the BRM.

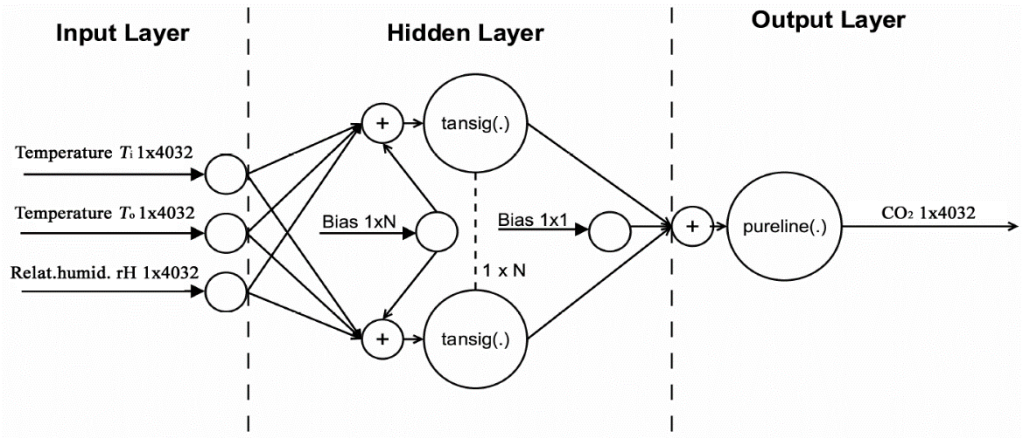


Figure 5. The architecture of the designed ANN.

3.2. Implementation of predictive analysis using IBM SPSS modeler

IBM SPSS offers two type of neural networks: Multilayer Perceptron (MLP) and Radial Basis Function (RBF), where MLP is categorized as a feed-forward, supervised learning network that can support up to two hidden layers. Basically, MLP is a function of one or more predictors that provides minimized prediction error for the targets. Additionally, there is an option of choosing between automatic or manual setting for the number of units [41]. The radial basis function (RBF) network is a feed-forward network that requires supervised learning. Unlike multilayer perceptron's (MLP), this network consists of only one hidden layer. Overall, there are three layers in the RBF network, IBM SPSS algorithm guide describes mathematical models of these layers [42] as following:

Input layer: $J_0 = P\ units,\ a_{0:1}, \dots, a_{0:J_0}$ with $a_{0:j} = X_j$

RBF layer: $j_1\ units, units, a_{1:1}, \dots a_{1:j_1};$ with $a_{1:j} = \emptyset_j(X)$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

$$\emptyset_j(X) = \frac{e^{(-\sum_{p=1}^{P}\frac{1}{2\sigma_{jp}^2}(x_p-\mu_{jp})^2}}{\sum_{j=1}^{J_1} e^{(-\sum_{p=1}^{P}\frac{1}{2\sigma_{jp}^2}(x_p-\mu_{jp})^2}} \quad (1)$$

Output layer: $j_2 = R\ units,\ a_{I:1}, \dots a_{I:j_2}$ with $a_{I:r} = \omega_{I:r} + \sum_{j=1}^{j1} \omega_{rj}\emptyset_j(X)$

Where:

$X^{(m)}$: Input vector

I: Number of layers (for RBM = 2)

Ji: Number of units in layer i

$\emptyset_j(X^{(m)})$: jth unit for input $X^{(m)}, j = 1, \dots, j_1$.

μ_j: Center of $\emptyset_j$

σ_j: Width of $\emptyset_j$

$a_{i:j}^m$: Unit j of layer i

ω_{rj}: weight connecting rth output unit and jth hidden unit of RBF layer

The training of RBF can be divided in two stages. The first stage, determines the basis function by clustering methods and the second stage determines the weights given to the basis function. The SPSS measures the accuracy of neural networks by calculating the percentage of the records for which the predicted value matches the observed value. For the continues values the accuracy is calculated by 1 minus the average of the absolute values of the predicted values minus the observed values over the maximum predicted value minus the minimum predicted value (the following formula) [43].

$$\text{Accuracy} = \frac{1}{n}\sum_{m=1}^{M}\left(1 - \frac{\left|y_r{}^{(m)} - \hat{y}_r{}^{(m)}\right|}{max_m(y_r{}^{(m)}) - min_m(y_r{}^{(m)})}\right), \quad (2)$$

3.3. Signal trend detection based on the Wavelet transformation

3.2.1. Materials and Methods

In this section, we introduce a method for the CO_2 concentration course noise cancelation based on the Wavelet transformation. In the signal processing, we assume that each signal *y(t)* is composed from two essential parts, namely they are the signal trend *T(t)* and a component having a stochastic character *X(t)* which is perceived as the signal noise and details. Based on this definition, we can use the following signal formulation (3):

$$y(t) = T(t) + X(t), \quad (3)$$

The major problem, when the signal trend is being extracted, is the noise detection. There are many applications of the trend detection including the CO_2 measurement. Such signal may be influenced by the glitches which should be removed to obtain a smooth signal for the further processing. The wavelet analysis represents a transformation of the signal *y(t)* to obtain two types of the coefficients, particularly they are the wavelet and scaling coefficients. These coefficients are completely equivalent with the original CO_2 signal. It is supposed that wavelet coefficients are related to changes along a specific defined scale. The main idea of the signal trend detection is to perform an association of the scaling coefficients with the signal trend *T(x)*, on the other hand the wavelet coefficients are supposed to be associated with the signal noise, which is mainly represented by the glitches, when processing the CO_2 signal. In our analysis we consider an uncorrelated noise, adapting the wavelet estimator in order to works as a kernel estimator. Advantage of such approach is formulating an estimator based on the sampled data irregularity. In this method, we are using the scaling coefficients as estimators of the signal trend. We suppose that the sampled CO_2 observations are represented by $Y(t_n)$, thus the CO_2 estimator is given by the formulation (4):

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

$$\hat{T}(t) = \sum_{n=0}^{N-1} Y(t_n) \int E_J(t,s)ds \tag{4}$$

Integration is done over a set of the intervals $(A_n(s))$, their union forms perform partitioning interval covering all the observations t_n, where $t_n \in A_n$. Consequently, E_J is defined as:

$$E_J(t,s) = 2^{-J} \sum_{k\in\mathbb{Z}} \theta(2^{-J}t - k)\theta(2^{-J}s - k) \tag{5}$$

In this expression $\theta(t)$ represents the scaling function. This function is defined as follows:

$$\theta(t) = \sum_{k\in\mathbb{Z}} c_k \theta(2t - k) \tag{6}$$

The wavelet function is defined by the equation:

$$\psi(t) = (-1)^k c_{1-k} \theta(2t - k) \tag{7}$$

The first crucial task is an appropriate selection of the mother's wavelet for the predicted CO_2 signals filtration. Supposing the Daubechies wavelets can well reflect morphological structure therefore, this family is used for our model. Particularly, in our approach we are using the Daubechies wavelet (Db6), with the D6 scaling function utilizing the orthogonal Daubechies coefficients.

3.4. Validation ratings used

In order to carry out the objective comparison, the following parameters are considered:

Mean Absolute Error (MAE) represents estimator measuring of difference between two continuous variables. The *MAE* is given by the following expression:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{8}$$

Mean squared error (MSE) represents estimator measuring the average of the error squares between two signals. The *MSE* represents a risk function which corresponds with the expected value of the squared or quadratic error loss. The *MSE* is given by the following expression:

$$MSE(x_1, \hat{x}_2) = \frac{1}{n}\sum_{i=1}^{n} \big(x(i) - \hat{x}(i)\big)^2 \tag{9}$$

Euclidean distance (ED) represents an ordinary straight-line distance between two points lying in the Euclidean space. Based on this distance, the Euclidean space becomes a metric space. The lower Euclidean distance we achieve, the more similar are two signal samples. In our analysis we consider mean of the *ED*. The Euclidean distance is given by the following expression:

$$d(x_1, \hat{x}_2) = \sqrt{(x_1 - \hat{x}_2)^2 + (y_1 - \hat{y}_2)^2} \tag{10}$$

City Block distance (CB) represents a distance between two signals $x_1, \hat{x}_2$ in the space with the Cartesian coordinate system. This parameter can be interpreted as a sum of the lengths of the projections of the line segments between the points onto the coordinate axes. CB distance is defined as follows:

$$d_{cb}(x_1, \hat{x}_2) = \|x_1 - \hat{x}_2\| = \sum_{i=1}^{n} |x_{1i} - \hat{x}_{2i}| \tag{11}$$

Correlation coefficient (*R*) measures a level of the linear dependency between two signals. The more are considered signals linearly dependable, the higher correlation coefficient is. In a comparison with the previous parameters, the correlation coefficient represents a normative parameter. Zero correlation stands for the total dissimilarity between two signals, measured in a sense of their linear dependency. Contrarily, 1, respectively -1 stands for the full positive, respectively negative correlation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

As we have already stated above, in our work we are analyzing samples of day courses and week courses CO_2 prediction. In each measurement, we have a prediction from the neural network with 10, 50, 100, 150, 200, 250, 300, 350, 400, 450 and 500 neurons. Thus, we completely analyze 11 predicted signals for each measurement. These signals are compared against the reference, based on the evaluation parameters stated above. In a term of the Euclidean, City Block distance and MSE lower values indicate a higher agreement between the predicted course CO_2 and reference measurement course CO_2 thus, better result. Contrarily, a higher number of correlation coefficients indicates better results. In the following part of the analysis, we report the results of the quantification comparison. All the tests are done for the Wavelet Db6, with 6-level of decomposition and the Wavelet settings: threshold selection rule - Stein's Unbiased Risk and soft thresholding for selection of the detailed coefficients.

4. The experimental practical part

4.1. First Part ADL information from CO_2 concentration course

This section contains the description of the experiments performed, the setting of the individual parameters, the calculated and measured results. In the first long-term experiment CO_2 (ppm) concentration indoor, temperature T(°C) indoor, temperature T(°C) outdoor and relative humidity rH(%) indoor measurement were performed in a spring month, from June 8 to June 15, 2015. The measured course of CO_2 concentration in Figure 6 indicates the occurrence of persons in the monitored area in SHC room 204 (Fig.3).

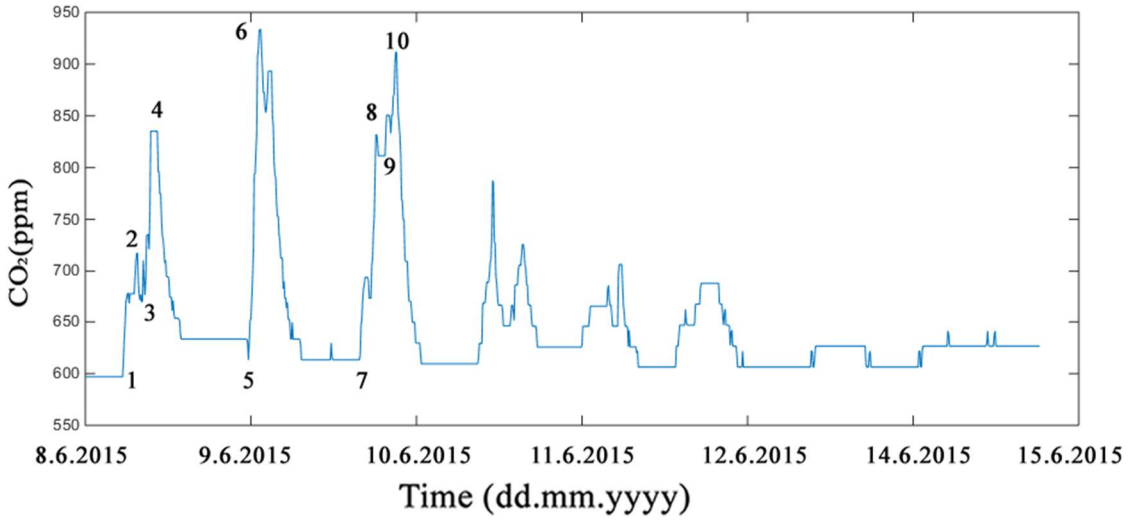


Figure. 6. The reference (measured) course of CO_2 concentration (from June 8 to June 15, 2015)

Legend to Figure 6: 1. arrival (8.6.2015 7:36:00), 2. departure (8.6.2015 10:22:00), Time of Person Presence (TPP) Δt_1 = 3:46:00; 3. arrival (8.6.2015 10:50:00), 4. departure (8.6.2015 13:00:00), TPP Δt_2 = 2:10:00; 5. arrival (9.6.2015 8:40:00), 6.departure (9.6.2015 11:00:00), TPP Δt_3 = 2:20:00; 7.arrival (10.6.2015 7:10:00), 8.departure (10.6.2015 10:20:00), TPP Δt_4 = 3:10:00; 9.arrival (10.6.2015 12:10:00), 10.departure (10.6.2015 14:20:00), TPP Δ t_5 = 2:10:00.

In the second long-term experiment CO_2 (ppm) concentration indoor, temperature T(°C) indoor, temperature T(°C) outdoor and relative humidity rH(%) indoor measurement were performed in a winter month, from February 8 to February 15, 2015. Figure 7 shows the arrival and departure of persons in the monitored area (see the legend) in the selected time interval.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

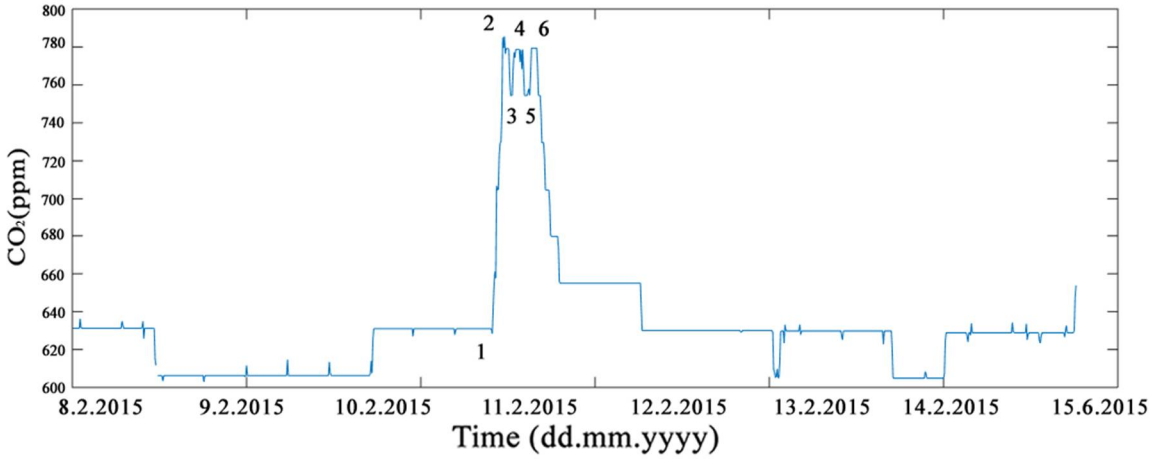


Figure. 7 The reference (measured) course of CO_2 concentration (from February 8 to February 15, 2015)

Legend to Figure 7: 1.arrival (11.2.2015 8:35:00), 2.departure (11.2.2015 10:35:00), TPP Δt_1= 2:00:00; 3.arrival (11.2.2015 12:08:00), 4.departure (11.2.2015 14:08:00), TPP Δt_2 = 2:00:00; 5.arrival (11.2.2015 15:29:00), 6.departure (11.2.2015 16:50:00), TPP Δt_3=1:21:00.

In the third short-term experiment CO_2 (ppm) concentration indoor, temperature T(°C) indoor, temperature T(°C) outdoor and relative humidity rH(%) indoor measurement were performed in a spring month, June 8, 2015. In the fourth short-term experiment CO_2 (ppm) concentration indoor, temperature T(°C) indoor, temperature T(°C) outdoor and relative humidity rH(%) indoor measurement were performed in a winter month, February 8, 2015.

4.2. Second Part ADL information from prediction CO_2 concentration course

The design of the experiments is based on the requirement for verifying the quality of the prediction execution (selected in a time interval of spring (from June 8 to June 15, 2015)) as well as for the winter months (winter (from February 8 to February 15, 2015)). Next, it was necessary to experimentally verify the quality of the prediction by means of cross-validation for the values measured in 8 days. The values measured from June 8 to June 15, 2015, and from February 8 to February 15, 2015, within the individual experiments served for the cross-validation and testing of the trained ANN BRM on the data obtained by measuring from June 8 to June 15, 2015, and from February 8 to February 15, 2015.

The practical implementation and ANN learning process together with the verification of the ANN BRM learned on the test data is carried out in the following steps:

Step 1. Measurement of non-electrical variables T_i, T_0, rH a CO_2 in the SHC room in interval from February 8 to February 15, 2015, from June 8 to June 15, 2015, using operating sensors in a selected SHC room 204 (BACnet technology sensors, Desigo Insight SCADA system, PI System (Fig.1)).

Step 2. Data preprocessing – normalization with the min-max method (MATLAB 2014b).

Step 3. Creating an **ANN BRM** structure (Fig. 5), (MATLAB 2014b).

Step 4. Training and Cross-validation data prediction CO_2 of the results achieved the MSE calculation, and the correlation coefficient R (MATLAB 2014b simulation) for:

a) **Training** ANN (BRM) with data from June 8 to June 15, 2015. The best results of prediction CO_2 were recorded with the **Cross-validation** with data from February 8 to February 15, 2015 for ANN BRM with **50 neurons, MSE=0.002 (ppm), R=0.839 (-), d=0.0677, d_{cb} = 0.0678** (Tab.1).

b) **Training** ANN (BRM) with data from February 8 to February 15, 2015. The best results of prediction CO_2 were recorded with the **Cross-validation** with data from February 8 to February 15, 2015 for ANN BRM with **100 neurons, MSE=0.0034 (ppm), R=0.942 (-), d=0.152, d_{cb} = 0.152** (Tab.3).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

c) **Training** ANN (BRM) with data from June 8, 2015. The best results of prediction CO_2 were recorded with the **Cross-validation** with data from February 8, 2015 for ANN BRM with **100 neurons, MSE=1.617*10^{-5} (ppm), R=0.991 (-),** d**=0.0288,** d_{cb} **= 0.0287** (Tab.5).

d) **Training** ANN (BRM) with data from February 8 to February 15, 2015. The best results of prediction CO_2 were recorded with the **Cross-validation** with data from February 8 to February 15, 2015 for ANN BRM with **250 neurons, MSE=5.349*10^{-5} (ppm), R=0.992 (-),** d**=0.198,** d_{cb} **= 0.198** (Tab.7).

Step 5. Graphic view of the **Cross-validation data prediction** CO_2 results (MATLAB 2014b simulation) for:

a) ANN (BRM) with 50 neurons (Tab.1), Step 4a, (from June 8 to June 15, 2015) with the tested data (from February 8 to February 15, 2015), (Fig. 8a).

b) ANN (BRM) with 100 neurons (Tab.3), Step 4b, (from February 8 to February 15, 2015) with the tested data (from June 8 to June 15, 2015), (Fig. 9a).

c) ANN (BRM) with 100 neurons (Tab.5), Step 4c, (June 8, 2015) with the tested data (February 8, 2015), (Fig. 10a).

d) ANN (BRM) with 250 neurons (Tab.7), Step 4d, (February 8, 2015) with the tested data (June 8, 2015), (Fig. 11a).

Step 6. The use of the Wavelet Transform algorithm (WTA) for noise cancelation in the predicted signal using the ANN BRM (MATLAB 2014b simulation).

a) ANN (BRM) with 50 neurons (Tab.2), (from June 8 to June 15, 2015) with the tested data (from February 8 to February 15, 2015), (Fig. 8b).

b) ANN (BRM) with 450 neurons (Tab.4), (from February 8 to February 15, 2015) with the tested data (from June 8 to June 15, 2015), (Fig. 9b).

c) ANN (BRM) with 450 neurons (Tab.6), (June 8, 2015) with the tested data (February 8, 2015), (Fig. 10b).

d) ANN (BRM) with 250 neurons (Tab.8), (February 8, 2015) with the tested data (June 8, 2015), (Fig. 11b).

Step 7. Comparison of the prediction results of the ANN BRM without filtering and with WTA filtering (Tab. 1 – Tab. 8) in discussion.

Step 8. Training and Cross-validation data prediction CO_2 of the results achieved the MSE calculation, and the correlation coefficient R (implemented in SPSS IBM IoT platform) without filtering WTA for:

a) **Training** ANN (RBF) with data from June 8 to June 14, 2015. The best results of prediction CO_2 were recorded with the **Cross-validation** with data from February 15 to February 21, 2015 (Tab.9).

b) **Training** ANN (RBF) with data from February 15 to February 21, 2015. The best results of prediction CO_2 were recorded with the **Cross-validation** with data from June 8 to February 14, 2015 (Tab.10).

c) **Training** ANN (RBF) with data from February 15, 2015. The best results of prediction CO_2 were recorded with the **Cross-validation** with data from February 16, 2015 (Tab.11).

d) **Training** ANN (RBF) with data from June 8, 2015. The best results of prediction CO_2 were recorded with the **Cross-validation** with data from February 8, 2015 (Tab.12).

Step 9. Evaluation of the results achieved.

The actual experiments were carried out as follows:

4.2.1. The long-term experiment - 8 days. Training ANN (BRM) with data from June 8 to June 15, 2015, was implemented on the data (from February 8 to February 15, 2015), (Tab. 1, Tab.2).

Training ANN, BRM (Step 4a) for neuron counts in an interval of 10-500, measuring the time t(s) of learning, the MSE calculation, and the correlation coefficient R for data from June 8 to June 15, 2015, (Tab. 1). The total number of samples was 1152. Training - 70%, 806 samples, which are presented to the network during the training, and the network is adjusted according to its error). Validation - 15%, 173 samples, which were used to measure the network generalization, and to halt

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

the training when the generalization stopped improving. Testing – 15%, 173 samples, which have no effect on the training and, therefore, they provide an independent measure of the network performance during and after the training. The prediction of CO_2 concentration using the ANN BRM was implemented by, firstly, using the measured values of rH (%) indoors, temperature indoors T_{in} (°C), temperature outdoors T_{out} (°C) on the cross-validation (Step 4a) used the training data (rH (%) indoors, temperature indoors T_{in} (°C), temperature outdoors T_{out} (°C)) that was measured from February 8 to February 15, 2015. Table 1 shows the calculated values of R and MSE parameters for the number of neurons (10 to 500) set within the testing of the ANN BRM (from June 8 to June 15, 2015) for predicting (Step 6a) the course of CO_2 using the ANN BRM **without WT** on the data measured in the interval from February 8 to February 15, 2015. The best calculated minimum value of MSE parameter (MSE = **0.002**) and maximum value of R correlation coefficient (R = **0.839**), for predicted courses of CO_2 are for trained ANN BRM with **50** neurons (Cross Validation). Table 2 shows the calculated values of R and MSE parameters for the number of neurons (10 to 500) set within the testing of the ANN BRM (from June 8 to June 15, 2015) for predicting (Step 6a) the course of CO_2 using the ANN BRM **with WT** (step 4) on the data measured in the interval from February 8 to February 15, 2015. The best calculated minimum value of MSE parameter (MSE = **0.0019**) and maximum value of R correlation coefficient (R = **0.854**), for predicted courses of CO_2 are for trained ANN BRM with **50** neurons (Cross Validation).

Table 1. Comparison of prediction quality ANN (BRM) [8.6.2015–15.6.2015 (data normalized)], with the Cross validation tested data from the interval [8.2.2015–15.2.2015].

Number of neurons	*MSE* (ppm)	*R(-)*	*d(-)*	d_{cb}(-)
10	0.0023	0.819	0.0678	0.0678
50	0.0020	0.839	0.0677	0.0678
100	0.0023	0.818	0.0667	0.0677
150	0.0022	0.822	0.0670	0.0671
200	0.0023	0.827	0.0679	0.0678
250	0.0023	0.819	0.0680	0.0680
300	0.0022	0.818	0.0676	0.0676
350	0.0022	0.823	0.0663	0.0664
400	0.0022	0.826	0.0675	0.0675
450	0.0023	0.816	0.0674	0.0676
500	0.0022	0.826	0.0677	0.0677

Figure 8a illustrates the reference measured course of CO_2 concentration from February 8 to February 15, 2015, and the predicted course of CO_2 concentration using the ANN BRM learned (Step 5a) on the data (from June 8 to June 15, 2015) and predicted within the cross-validation (Step 4a), (with the data from February 8 to February 15, 2015) using the ANN BRM (number of neurons 50), (Tab.1). The red circles indicate the predicted CO_2 values that were incorrectly calculated within the prediction error using the ANN BRM. Figure 8b illustrates the reference measured course of CO_2 concentration from February 8 to February 15, 2015. Figure 8b further illustrates the predicted filtered course of CO_2 concentration using the WT (the ANN BRM learned on the data (from June 8 to June 15, 2015) and predicted within the cross-validation (Step 4a) with the data from February 8 to February 15, 2015, using the ANN BRM (the number of neurons - 50), (Tab.2)).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

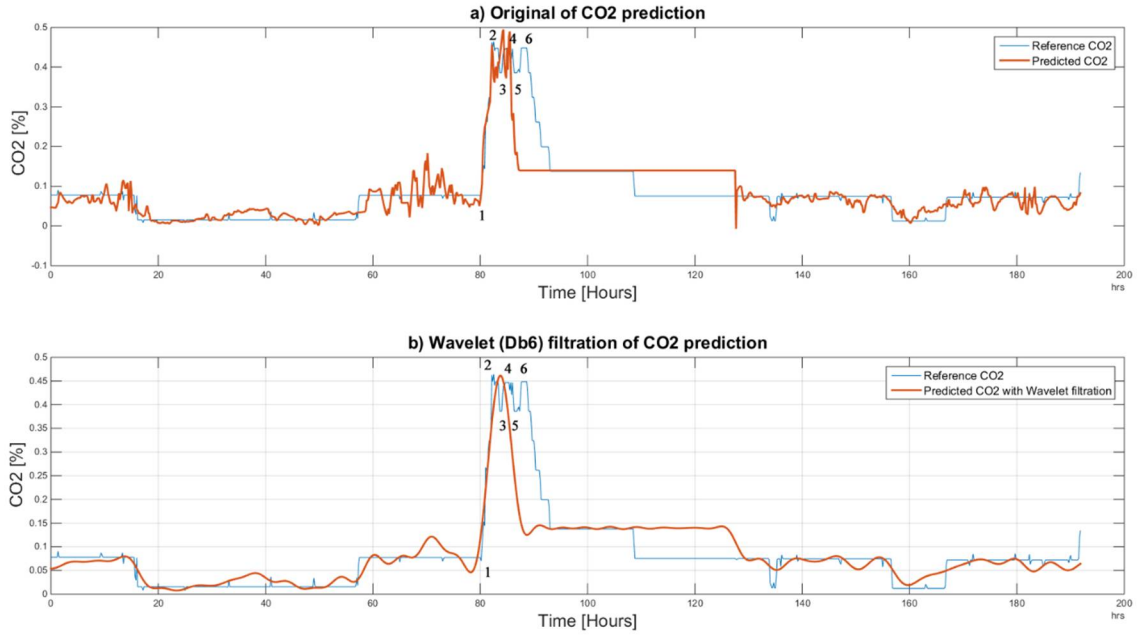


Figure 8a. The reference (measured course of CO_2 concentration (from February 8 to February 15, 2015)) and the predicted course of CO_2 concentration. (The ANN BRM learned (from June 8 to June 15, 2015), the predicted ANN with the data from February 8 to February 15, 2015 – the number of neurons - 50, (Tab. 1).

Figure 8b. The reference (measured course of CO_2 concentration (from February 8 to February 15, 2015)) and predicted filtered course of CO_2 concentration using the WT (the ANN BRM learned on the data (from June 8 to June 15, 2015) and predicted within the cross-validation with the data from February 1 to February 28, 2015, using the ANN BRM (the number of neurons - 50), (Tab.2).

Legend to Figures 8a, 8b: 1. arrival (11.2.2015 8:35:00), 2. departure (11.2.2015 10:35:00), TPP Δt_1= 2:00:00; 3. arrival (11.2.2015 12:08:00), 4.departure (11.2.2015 14:08:00), TPP Δt_2 = 2:00:00; 5.arrival (11.2.2015 15:29:00), 6.departure (11.2.2015 16:50:00), TPP Δt_3=1:21:00.

Table 2. Comparison of prediction quality ANN (BRM) [8.6.2015–15.6.2015 (data normalized)], with the **WT noise cancelling** of Cross validation tested data from the interval [8.2.2015–15.2.2015].

Number of neurons	*MSE* (ppm)	*R(%)*	*d(-)*	d_{cb}(-)
10	0.0021	0.836	0.0663	0.0663
50	0.0019	0.854	0.0664	0.0664
100	0.0021	0.835	0.0653	0.0651
150	0.0021	0.833	0.0659	0.0659
200	0.0020	0.841	0.0666	0.0665
250	0.0021	0.834	0.0668	0.0668
300	0.0021	0.836	0.0661	0.0660
350	0.0021	0.837	0.0650	0.0650
400	0.0021	0.838	0.0660	0.0660
450	0.0021	0.832	0.0661	0.0662
500	0.0020	0.841	0.0663	0.0662

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

4.2.2. The long-term experiment - 8 days. Training the ANN BRM was implemented on the data (from February 8 to February 15, 2015), (Tab.3, Tab.4).

Training ANN, BRM (Step 4b) for neuron counts in an interval of 10-500, measuring the time t(s) of learning, the MSE calculation, and the correlation coefficient R for the data from February 8 to February 15, 2015, (Fig. 9a), (Tab. 3). The total number of samples was 1152. Training - 70%, 806 samples, which are presented to the network during the training, and the network is adjusted according to its error). Validation - 15%, 173 samples, which were used to measure the network generalization, and to halt the training when the generalization stopped improving. Testing – 15%, 173 samples, which have no effect on the training and, therefore, they provide an independent measure of the network performance during and after the training. The prediction of CO_2 concentration **without WT** using the ANN BRM was implemented by the cross-validation (Step 4b) used the training data (rH (%) indoors, temperature indoors T_{in} (°C), temperature outdoors T_{out} (°C)) that was measured from June 8 to June 15, 2015, (Tab. 3). The minimum value of MSE parameter is MSE = **0.0034** and maximum value of correlation coefficient R = **0.942** were calculated for ANN BRM where the configured number of neurons was **100** (Tab. 3). The prediction of CO_2 concentration **with WT** using the ANN BRM was implemented by the cross-validation (Step 5b) used the training data (rH (%) indoors, temperature indoors T_{in} (°C), temperature outdoors T_{out} (°C)) that was measured from June 8 to June 15, 2015, (Tab. 4). The minimum value of MSE parameter is MSE = **0.0024** and maximum value of correlation coefficient R = **0.957** were calculated for ANN BRM where the configured number of neurons was **450** (Tab. 4).

Table 3. Comparison of prediction quality ANN (BRM) [8.2.2015–15.2.2015 (data normalized)], with the Cross validation tested data from the interval [8.6.2015–15.6.2015].

Number of neurons	*MSE* (ppm)	*R(%)*	*d(-)*	d_{cb}(-)
10	0.0062	0.887	0.151	0.152
50	0.0049	0.921	0.154	0.154
100	0.0034	0.942	0.152	0.152
150	0.0034	0.941	0.152	0.152
200	0.0040	0.932	0.154	0.153
250	0.0038	0.933	0.153	0.153
300	0.0049	0.918	0.153	0.153
350	0.0176	0.764	0.156	0.156
400	0.0039	0.933	0.152	0.152
450	0.0038	0.934	0.152	0.152
500	0.0048	0.917	0.153	0.152

Table 4. Comparison of prediction quality ANN (BRM) [8.2.2015–15.2.2015 (data normalized)], with the **WT noise cancelling** of Cross validation tested data from the interval [8.6.2015–15.6.2015].

Number of neurons	*MSE* (ppm)	*R(%)*	*d(-)*	d_{cb}(-)
10	0.0050	0.911	0.148	0.147
50	0.0025	0.957	0.150	0.150
100	0.0026	0.954	0.150	0.149
150	0.0025	0.955	0.150	0.150
200	0.0026	0.954	0.152	0.151
250	0.0025	0.955	0.150	0.150
300	0.0025	0.956	0.151	0.151
350	0.0042	0.926	0.152	0.151
400	0.0024	0.957	0.149	0.149
450	0.0024	0.957	0.149	0.148
500	0.0028	0.951	0.150	0.150

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

Figure 9a illustrates the reference measured course of CO_2 concentration from June 8 to June 15, 2015, and the predicted course of CO_2 concentration using the ANN BRM learned on the data (from February 8 to February 15, 2015) and predicted within the cross-validation (with the data from June 8 to June 15, 2015) using the ANN BRM (number of neurons 100), (Tab. 3).

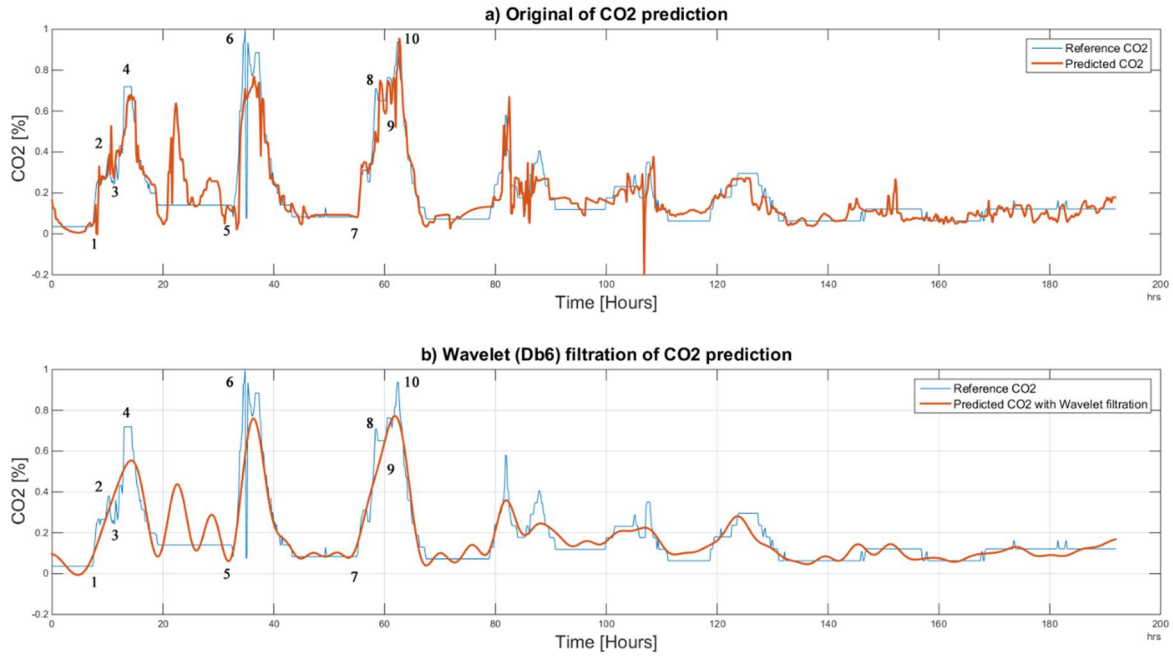


Figure 9. a) The reference measured course of CO_2 concentration from June 8 to June 15, 2015, and the predicted course of CO_2 concentration using the ANN BRM learned on the data (from February 8 to February 15, 2015) and predicted within the cross-validation (with the data from June 8 to June 15, 2015) using the ANN BRM (number of neurons 100), (Tab. 3)

Figure 9. b) The reference measured course of CO_2 concentration from June 8 to June 15, 2015, and the predicted filtered course of CO_2 concentration using the WT (the ANN BRM learned on the data from February 8 to February 15, 2015) and predicted within the cross-validation (with the data from June 8 to June 15, 2015) using the ANN BRM (number of neurons 450), (Tab. 4)

Legend to Figures 9a, 9b: 1. arrival (8.6.2015 7:36:00), 2. departure (8.6.2015 10:22:00), Time of Person Presence (TPP) Δt_1 = 3:46:00; 3. arrival (8.6.2015 10:50:00), 4. departure (8.6.2015 13:00:00), TPP Δt_2 = 2:10:00; 5. arrival (9.6.2015 8:40:00), 6. departure (9.6.2015 11:00:00), TPP Δt_3 = 2:20:00; 7. arrival (10.6.2015 7:10:00), 8. departure (10.6.2015 10:20:00), TPP Δt_4 = 3:10:00; 9. arrival (10.6.2015 12:10:00), 10.departure (10.6.2015 14:20:00), TPP $\Delta\ t_5$ = 2:10:00.

Figure 9b illustrates the reference measured course of CO_2 concentration from June 8 to June 15, 2015, and the predicted filtered course of CO_2 concentration using the WT (the ANN BRM learned on the data from February 8 to February 15, 2015) and predicted within the cross-validation (with the data from June 8 to June 15, 2015) using the ANN BRM (number of neurons 450), (Tab. 4).

4.2.3. The short-term experiment - 1 day. Training the ANN BRM was implemented on the data (February 8 2015), (Tab. 5, Tab. 6).

Training ANN, BRM (Step 4c) for neuron counts in an interval of 10-800, measuring the time t(s) of learning, the MSE calculation, and the correlation coefficient R for data June 8, 2015, (Fig. 10a), (Tab. 5). The total number of samples – 144. Training - 70%, 100 samples. Validation - 15%, 22 samples. Testing – 15%, 22 samples. The prediction (Step 4c) of CO_2 concentration using the ANN BRM was implemented by, firstly, using the measured values of rH (%) indoors, temperature indoors T_{in} (°C),

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

temperature outdoors T_{out} (°C) on the training data from June 8 2015, (Tab.5). Further, the cross-validation (Step 4c) used the training data (rH (%) indoors, temperature indoors T_{in} (°C), temperature outdoors T_{out} (°C)) that was measured from February 8, 2015, (Tab.5). The minimum value of MSE = **$1.617.10^{-4}$** and maximum value of correlation coefficients R = **0.991** were calculated for ANN BRM where the configured number of neurons was **100** (Tab. 5).

Table 5. Comparison of prediction quality ANN (BRM) [8.6.2015], with the Cross validation tested data from the interval 1 day [8.2.2015].

Number of neurons	*MSE* (ppm)	*R(%)*	*d(-)*	d_{cb}*(-)*
10	$5.838*10^{-5}$	0.961	0.0287	0.0288
50	$1.816*10^{-5}$	0.882	0.309	0.0309
100	$1.617*10^{-5}$	0.991	0.0288	0.0287
150	$5.999*10^{-5}$	0.961	0.0299	0.0299
200	$5.646*10^{-5}$	0.962	0.0306	0.0298
250	$1.838*10^{-4}$	0.883	0.0298	0.0308
300	$6.249*10^{-5}$	0.962	0.0301	0.029
350	$1.926*10^{-4}$	0.881	0.0298	0.0318
400	$6.018*10^{-5}$	0.964	0.0299	0.0299
450	$1.867*10^{-4}$	0.892	0.0301	0.0307
500	$2.319*10^{-5}$	0.984	0.0298	0.0298

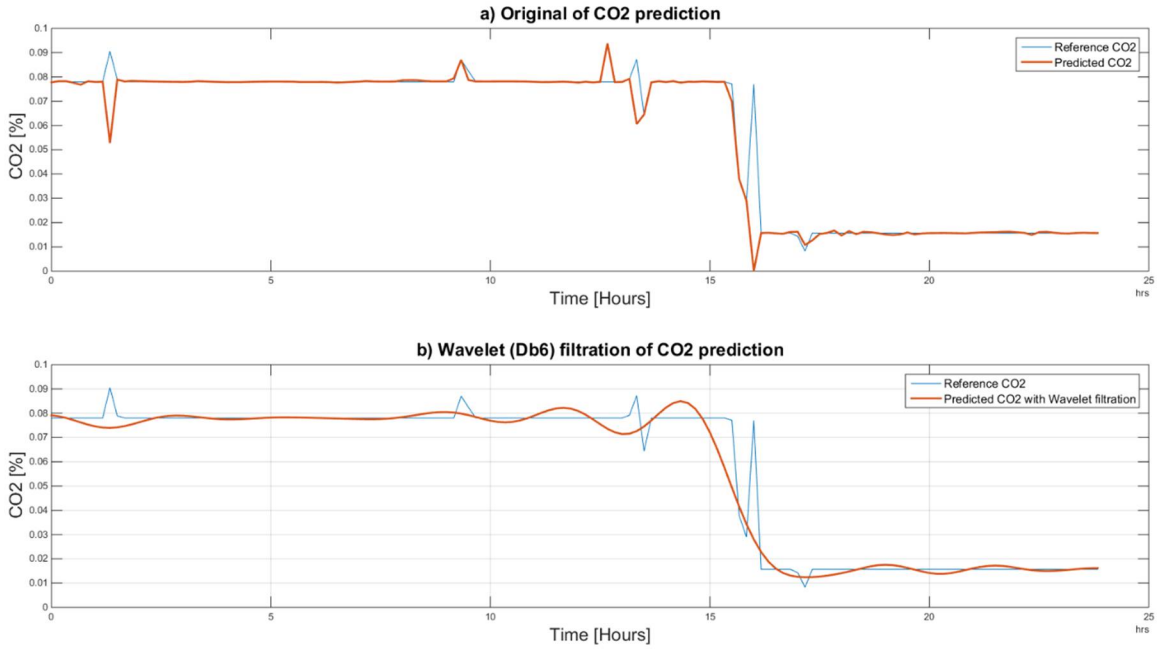
Table 6 shows the calculated values of R and MSE parameters for the number of neurons (10 to 500) set within the testing of the ANN BRM **with WT** (from June 8 2015) for predicting (Step 5c) the course of CO_2 using the ANN BRM **with WT** on the data measured in the interval from February 8, 2015. The best calculated minimum MSE parameter (MSE = **0.0024**) and maximum R correlation coefficient (R = **0.957**), (Table 6) for predicted courses of CO_2 (Fig. 10b) are for trained ANN BRM with **450** neurons (Cross Validation).

Table 6. Comparison of prediction quality ANN (BRM) [8.6.2015], with the **WT noise cancelling** of Cross validation tested data from the interval [8.2.2015].

Number of neurons	*MSE* (ppm)	*R(%)*	*d(-)*	d_{cb}*(-)*
10	0.0050	0.911	0.148	0.147
50	0.0025	0.957	0.150	0.150
100	0.0026	0.954	0.150	0.149
150	0.0025	0.955	0.150	0.150
200	0.0026	0.954	0.152	0.151
250	0.0025	0.955	0.150	0.150
300	0.0025	0.956	0.151	0.151
350	0.0042	0.926	0.152	0.151
400	0.0024	0.957	0.149	0.149
450	0.0024	0.957	0.149	0.148
500	0.0028	0.951	0.150	0.150

Figure 10a illustrates the reference measured course of CO_2 concentration from February 8, 2015, and the predicted course of CO_2 concentration using the ANN BRM learned (Step 4c) on the data (from June, 2015) and predicted within the cross-validation (Step 4c), (with the data from February 8, 2015) using the ANN BRM (number of neurons 450), (Tab.5). The red circles indicate the predicted CO_2 values that were incorrectly calculated within the prediction error using the ANN BRM.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1



585

586 **Figure 10a.** The reference (measured course of CO_2 concentration (from February 8, 2015)) and the
587 predicted course of CO_2 concentration. (The ANN BRM learned (from June 8, 2015), the predicted
588 ANN with the data from February 8, 2015 – the number of neurons - 100, (Tab. 5).

589 **Figure 10b.** The reference (measured course of CO_2 concentration (from February 8, 2015)) and the
590 predicted course of CO_2 concentration **with WT**. (The ANN BRM learned (from June 8, 2015), the
591 predicted ANN with the data from February 8, 2015 – the number of neurons - 450, (Tab. 6).

592 Figure 10b illustrates the reference measured course of CO_2 concentration from February 8, 2015.
593 Figure 10b further illustrates the predicted filtered course of CO_2 concentration using the WT
594 filtration (the ANN BRM learned on the data (from June 8, 2015) and predicted within the cross-
595 validation (Step 4c) with the data from February 8, 2015, using the ANN BRM (the number of neurons
596 - 450), (Tab.6)).

597 4.2.4 The short-term experiment - 1 day. Training the ANN BRM was implemented on the data
598 (Jun 8, 2015), (Tab. 7, Tab. 8).

599 Training ANN, BRM (Step 4d) for neuron counts in an interval of 10-500, measuring the time t(s)
600 of learning, the MSE calculation, and the correlation coefficient R for the data from February 8, 2015,
601 (Fig. 11a), (Tab. 7). The total number of samples was 1152. Training - 70%, 806 samples, which are
602 presented to the network during the training, and the network is adjusted according to its error).
603 Validation - 15%, 173 samples, which were used to measure the network generalization, and to
604 halt the training when the generalization stopped improving. Testing – 15%, 173 samples, which have
605 no effect on the training and, therefore, they provide an independent measure of the network
606 performance during and after the training.
607 The prediction of CO_2 concentration **without WT** using the ANN BRM was implemented by the
608 cross-validation (Step 4d) used the training data (rH (%) indoors, temperature indoors T_{in} (°C),
609 temperature outdoors T_{out} (°C)) that was measured from June 8, 2015, (Tab. 7). The minimum value
610 of MSE parameter is **MSE=5.349*10^{-4}** and maximum value of correlation coefficient R = **0.992** were
611 calculated for ANN BRM where the configured number of neurons was **250** (Tab. 7).
612 The prediction of CO_2 concentration **with WT** using the ANN BRM was implemented by the
613 cross-validation (Step 5d) used the training data (rH (%) indoors, temperature indoors T_{in} (°C),
614 temperature outdoors T_{out} (°C)) that was measured from June 8, 2015, (Tab. 8). The minimum value
615 of MSE parameter is **MSE=5.542*10^{-4}** and maximum value of correlation coefficient R = **0.992** were
616 calculated for ANN BRM where the configured number of neurons was **250** (Tab. 8).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

Table 7. Comparison of prediction quality ANN (BRM) [8.2.2015], with the Cross validation tested data from the interval 1 day [8.6.2015].

Number of neurons	*MSE* (ppm)	*R(%)*	*d(-)*	d_{cb}*(-)*
10	$6.345*10^{-4}$	0.991	0.197	0.197
50	0.0020	0.971	0.196	0.198
100	$7.811*10^{-4}$	0.989	0.198	0.196
150	$5.693*10^{-4}$	0.981	0.196	0.198
200	0.0013	0.981	0.198	0.196
250	$5.349*10^{-4}$	0.992	0.198	0.198
300	$7.695*10^{-4}$	0.988	0.197	0.198
350	0.0013	0.981	0.198	0.197
400	0.0012	0.983	0.198	0.198
450	0.0013	0.982	0.199	0,198
500	$6.191*10^{-4}$	0.991	0.197	0.199

Table 8. Comparison of prediction quality ANN (BRM) [8.2.2015], with the **WT noise cancelling** of Cross validation tested data from the interval [8.6.2015].

Number of neurons	*MSE* (ppm)	*R(%)*	*d(-)*	d_{cb}*(-)*
10	$5.468*10^{-4}$	0.992	0.197	0.197
50	$7.554*10^{-4}$	0.989	0.195	0.195
100	$5.664*10^{-4}$	0.992	0.196	0.196
150	$5.114*10^{-4}$	0.992	0.197	0.197
200	$9.316*10^{-4}$	0.987	0.196	0.196
250	$5.542*10^{-4}$	0.992	0.198	0.198
300	$6.473*10^{-4}$	0.991	0.195	0.198
350	$8.914*10^{-4}$	0.987	0.197	0.197
400	$8.236*10^{-4}$	0.988	0.198	0.198
450	$6.661*10^{-4}$	0.991	0.196	0.196
500	$5.871*10^{-4}$	0.991	0.198	0.198

Figure 9a illustrates the reference measured course of CO_2 concentration from June 8 to June 15, 2015, and the predicted course of CO_2 concentration using the ANN BRM learned on the data (from February 8 to February 15, 2015) and predicted within the cross-validation (with the data from June 8 to June 15, 2015) using the ANN BRM (number of neurons 100), (Tab. 3).

Figure 9b illustrates the reference measured course of CO_2 concentration from June 8 to June 15, 2015, and the predicted filtered course of CO_2 concentration using the WT (the ANN BRM learned on the data from February 8 to February 15, 2015) and predicted within the cross-validation (with the data from June 8 to June 15, 2015) using the ANN BRM (number of neurons 450), (Tab. 4).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

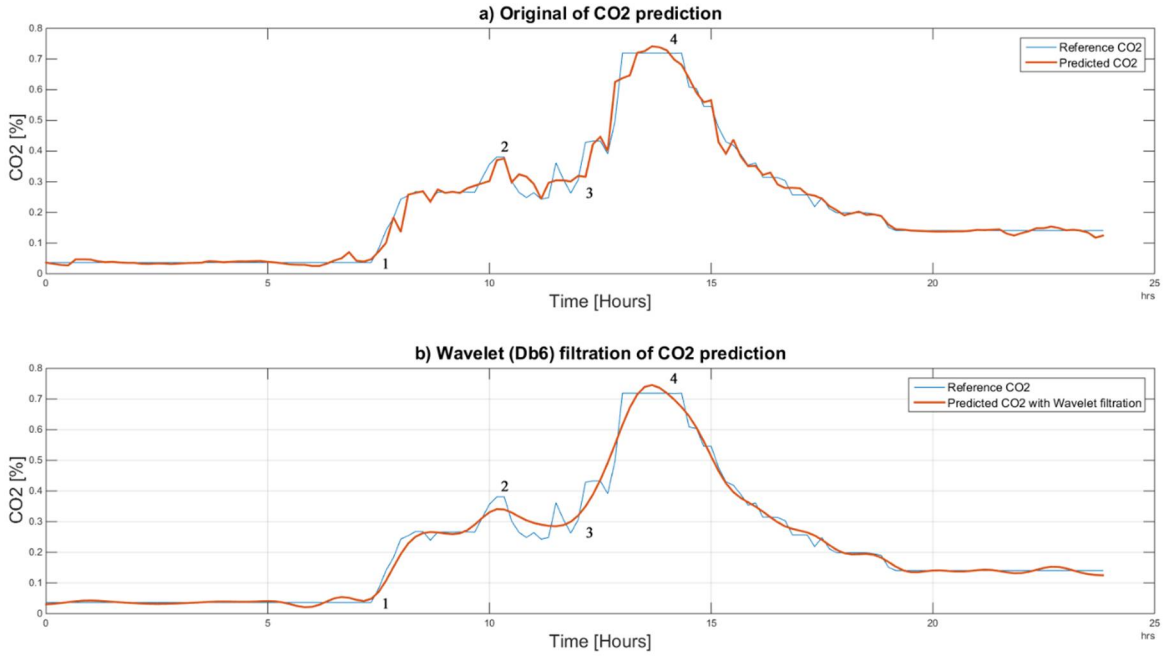


Figure 11. a) The reference measured course of CO_2 concentration from June 8, 2015, and the predicted course of CO_2 concentration using the ANN BRM learned on the data (from February 8, 2015) and predicted within the cross-validation (with the data from June 8, 2015) using the ANN BRM (number of neurons 250), (Tab. 7)

Figure 11. b) The reference measured course of CO_2 concentration from June 8, 2015, and the predicted filtered course of CO_2 concentration using the WT (the ANN BRM learned on the data from February 8, 2015) and predicted within the cross-validation (with the data from June 8, 2015) using the ANN BRM (number of neurons 250), (Tab. 8)

Legend to Figures 11a, 11b: 1 arrival (8:50:00), 2 departure (10:30:00), TPP Δt_1 = 1:40:00; 3 arrival (10:50:00), 4 departure (11:10:00), TPP Δt_2 = 0:20:00; 5 arrival (11:20:00) and 6 departure (11:30:00), TPP Δt_3 = 0:20:00.

4.3. Third party - Testing and Quantitative Comparison (WT additive noise canceling)

In our research, we are analyzing signals representing the CO_2 signals. We have a set of the estimated (predicted) signals being compared against the real measured signal CO_2, which is perceived as a reference. In our analysis, we are comparing two samples of day course CO_2 prediction and further two samples of week course CO_2 prediction. Based on the observations, it is apparent that the predicted CO_2 signals do not have a smooth process. They are frequently influenced by rapid oscillations, so-called glitches and also signal fluctuations (Fig. 12). Such signal variations represent the signal noise, impairing the real trend of the CO_2 prediction, which should be reduced. In our analysis, we are using the wavelet filtration to eliminate such signals to obtain the signal trend for further processing. As we have already stated above, we are using the mother's wavelet Db6 for the CO_2 signal trend detection. Firstly, take advantage of the fact that different level of the decomposition allows perceiving more or less the signal details represented by the detailed coefficients. Since we need to perceive the signal trend by eliminating the steep fluctuations, we need to consider an appropriate level of the decomposition. An experimental comparison of individual wavelet settings is reported in Figure 12. Based on the experimental results, we are using 6-level decomposition for the CO_2 signal trend detection. The filtration procedure utilizes the further following settings: threshold selection rule - Stein's Unbiased Risk and soft thresholding for selection of the detailed

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

coefficients.

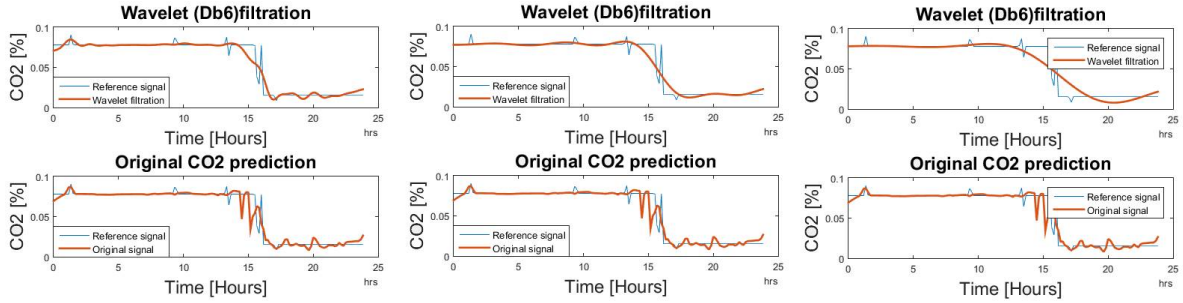


Figure 12. Comparison of original predicted CO_2 signal and Wavelet filtration for: 2-level decomposition (left), 6-level decomposition (middle) and 7-level decomposition (right), (ANN BRM trained with data (8.6.2015), cross validation prediction with data (8.2.2015)).

Wavelet filtration is being used for the extraction of the CO_2 signal trend, simultaneously rapid changes of the signal should be removed. On the Fig. 13 and 14, there is a comparison among the reference signal and predicted signals by the wavelet transformation for the day and week prediction.

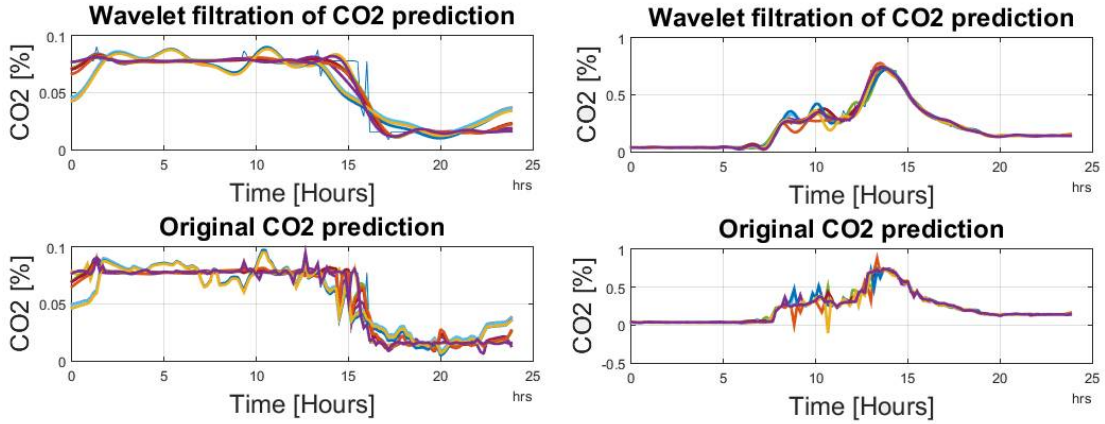
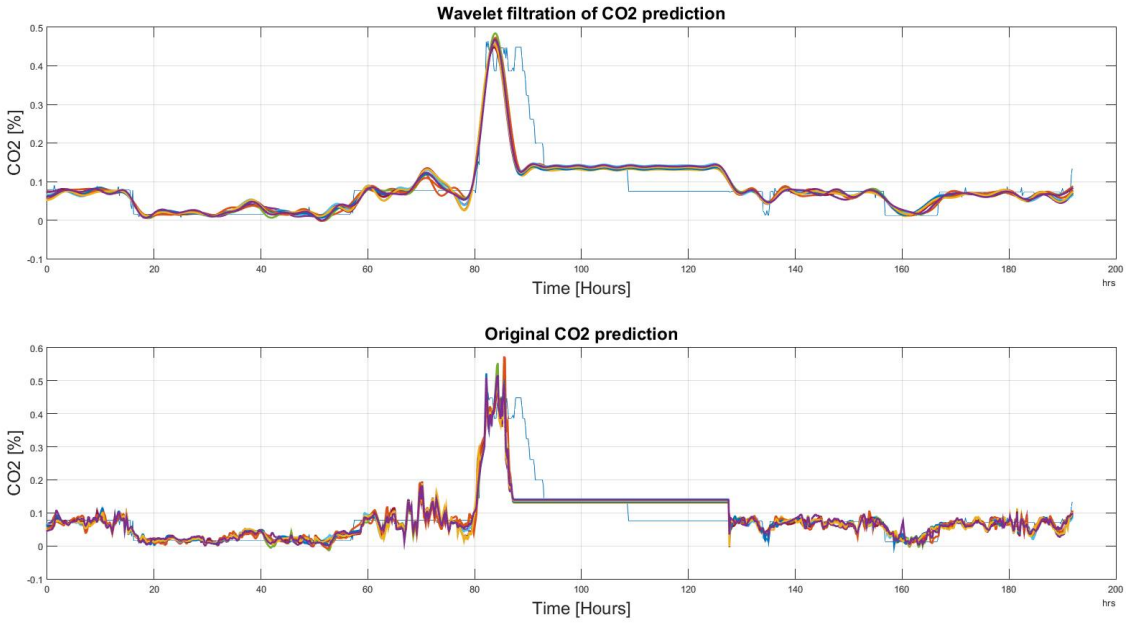


Figure 13. Example of two samples one-day prediction of CO_2 signal. Filtration is done by using the Db6 wavelet with 6-level decomposition), (ANN BRM trained with data (8.6.2015), cross validation prediction with data (8.2.2015)).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

Figure 14. Example of one-week prediction of the CO_2 course. Filtration is done by using the Db6 wavelet with 6-level decomposition, (ANN BRM trained with data (8.6.-15.6.2015), cross validation prediction with data (8.2.-15.2.2015)).

Based on the results, the wavelet filtration is capable of filtering rapid signal changes, whilst preserving the signal trend. In order to justify this situation, we report selected situations showing the glitches deteriorating a smooth signal trend, and a respective wavelet approximation largely reducing such signal parts (Fig.14, 15, 16).

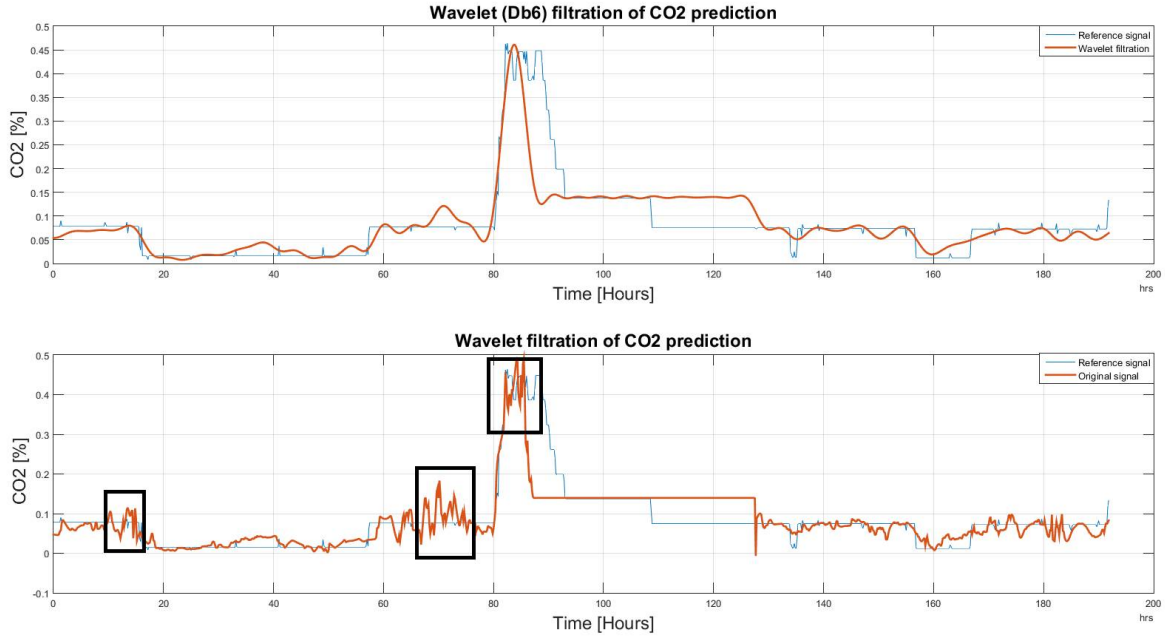


Figure 15. Week prediction of the CO_2 signal containing several glitches and signal spikes, marked as black RoI, (ANN BRM trained with data (8.6.-15.6.2015), cross validation prediction with data (8.2.-15.2.2015)).

As it is obvious, the CO_2 signal contains lots of significant occurrences represented by the glitches and spikes, significantly deteriorating a smoothness of the analyzed signal. Wavelet appears to be a reliable alternative for reduction such signal's parts. On the other hand, we are aware that trend detection, in some cases, reduces the peaks thus, the original signal's amplitude is reduced, such situations are reported on Fig. 15 and 16.

In the last part of our analysis, the objective comparison is carried out. As we have already stated, we are comparing predicted CO_2 signals with signals being filtered out by the wavelet transformation. All the signals are compared against the reference CO_2 signals for day and week prediction.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

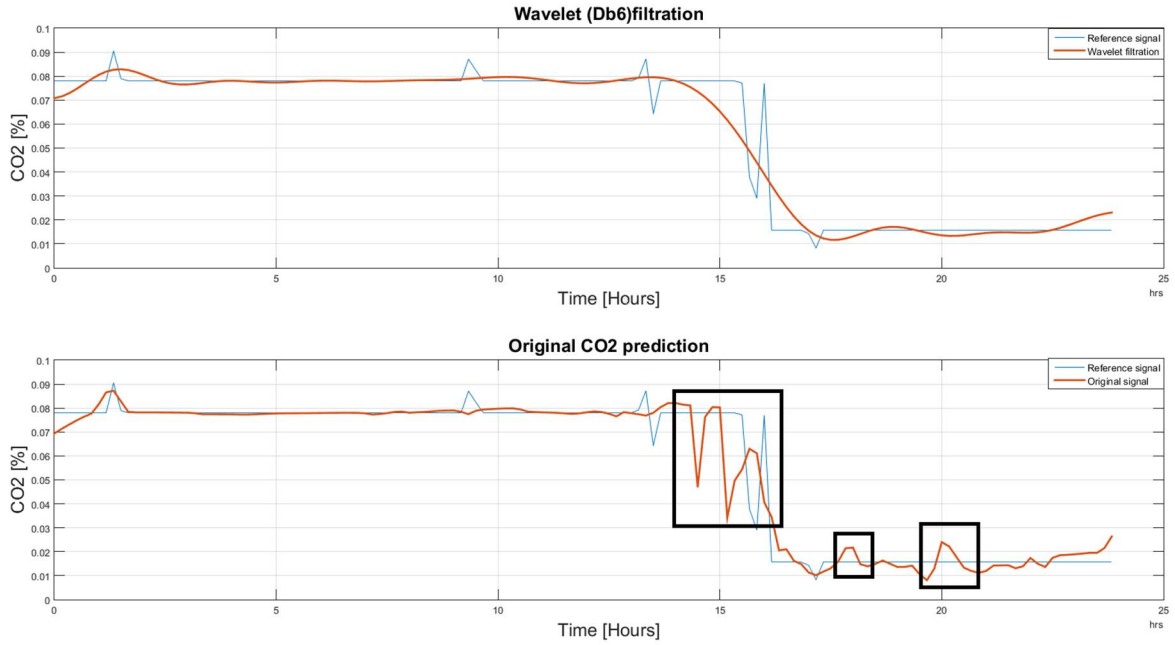


Figure 16. Day prediction of the CO2 signal containing several glitches and signal spikes, marked as black RoI, (ANN BRM trained with data (8.6.2015), cross validation prediction with data (8.2.2015)).

4.4. Fourth part - practical implementation prediction of CO_2 using ANN RBF IBM SPSS (IoT) SW Tool

For practical implementation of ANN RBF to predict the measured quantities for the purpose of monitoring the ADL in a real-world SHC environment was selected SPSS IBM SW Tool within IoT platform implementation. To predict the course of CO_2 concentration from the measured temperature *Ti* (°C), relative humidity rH (%) in the interior in the selected SHC room R204 and from the outdoor temperature readings *To* (°C), with the ANN Radial Basis Function (RBF) network, which is a feed-forward network that requires supervised learning. For the classification of prediction quality, a correlation analysis (correlation coefficient R), calculated MSE (Mean Squared Error) and Mean Absolute Error (MAE) were used. A data type selection operator was used to set default target to CO_2 and default inputs (in this case inputs were used as predictors) to humidity, indoor temperature and outdoor temperature. The transformed data were fed to a neural network for training a ANN model. The developed stream using IBM SPSS is shown in the Figure 17.

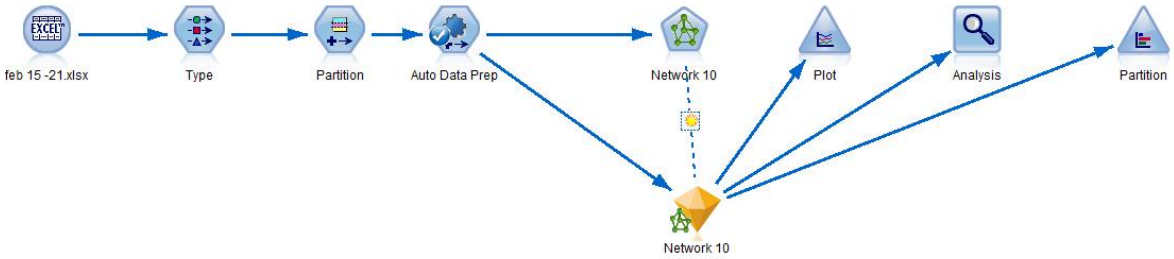


Figure 17. Developed Stream using IMB SPSS modeler.

The ANN used RBF model with various number of neurons (10, 50, 100, 150, 200, 250, 300, 350, 400), (Tab. 9 – Tab. 12). In order to verify the results, it is necessary to cross-validate the models. In the other words using two different set of data for training and validating. In the first case of cross-validation, period of 8.6.2015 to 14.6.2015 was used for training of the neural networks and period of 15.2.2015 to 21.2.2015 was used for cross validation. From Table 9, it can be observed that the measurement number 4 holds the highest linear correlation, but the measurement number 1 with

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

similar linear correlation has significantly lower MAE and MSE. Therefore, it had been selected as the best result for this implementation. The Figure 21 shows that for every rise on the reference signal, the predicted signal rises as well, which could provide an indication of human activity with appropriate filtering. Overall, this validation indicates a poor result.

Table 9. Trained ANN RBF model with data 8.6.2015 to 14.6.2015, within cross-validation with data (15.2.2015 to 21.2.2015).

Order of measurement	*Number of neurons (-)*	*R(-)*	*MAE (ppm)*	*MSE(ppm)*
1	10	0.273	0.064	$9.71*10^{-3}$
2	50	0.041	0.094	$1.21*10^{-2}$
3	100	0.146	0.063	$9.20*10^{-2}$
4	150	0.275	0.285	$1.69*10^{-1}$
5	200	0.102	0.067	$1.20*10^{-2}$
6	250	0.031	0.081	$14.4*10^{-3}$
7	300	0.031	0.081	$14.371*10^{-3}$
8	350	0.031	0.081	$14.371*10^{-3}$
9	400	0.031	0.081	$14.371*10^{-3}$

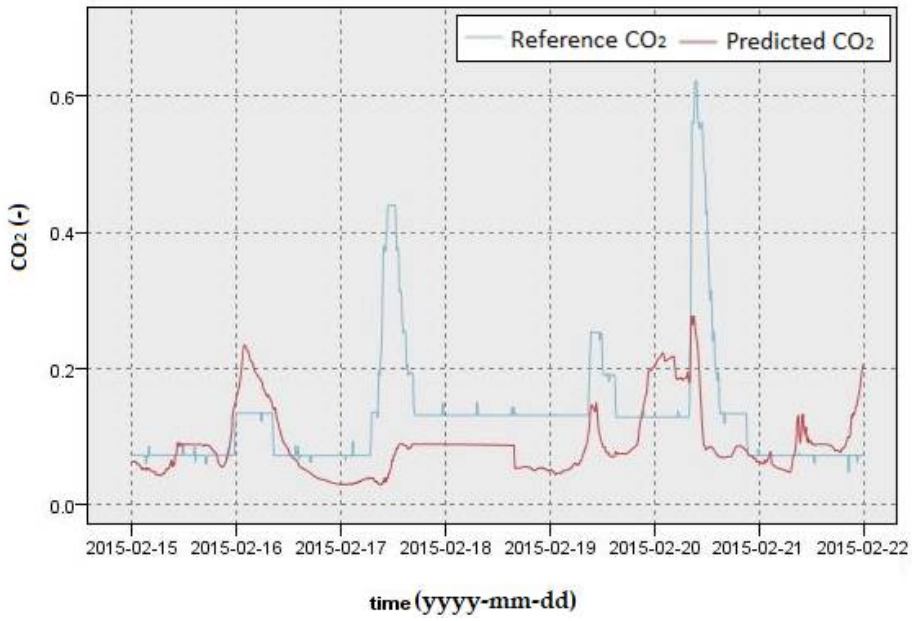


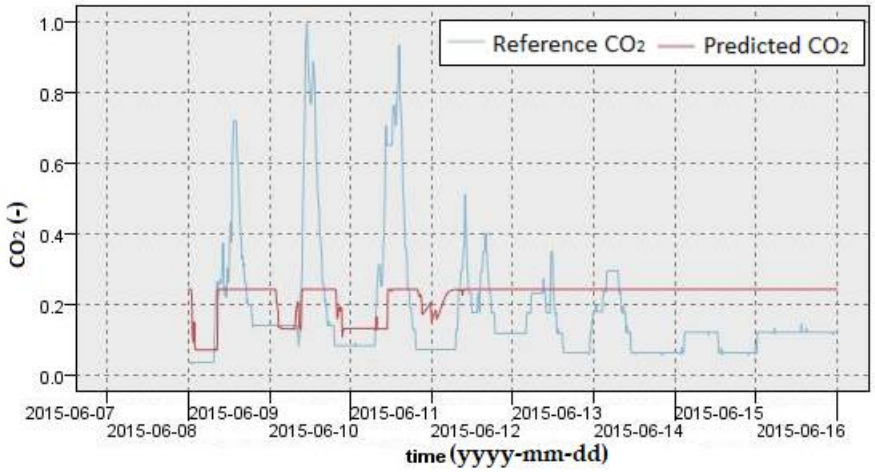
Figure 21. Order of measurement number 1 (Tab. 9), trained ANN RBF model with data 8.6.2015 to 14.6.2015, within cross-validation with data (15.2.2015 to 21.2.2015).

Table 10 represents the ANN RBF model trained with data from 15.2.2015 to 21.2.2015 and cross-validated with data from 8.6.2015 to 14.6.2015. It was shown in trainings stage, the models that are trained using this period of time hold lower accuracy in predictions. In the case of cross-validation, the difference is even more significant (Fig. 22).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

727 **Table 10**. Trained ANN RBF model with data 8.2.2015 to 14.2.2015, within cross-validation with
728 data (15.6.2015 to 21.6.2015).

Order of measurement	*Number of neurons (-)*	*R(-)*	*MAE (ppm)*	*MSE(ppm)*
1	10	0.160	0.168	$3.804*10^{-2}$
2	50	0.168	0.131	$3.20*10^{-2}$
3	100	0.001	0.109	$4.19*10^{-2}$
4	150	0.011	0.109	$3.26*10^{-2}$
5	200	0.031	0.095	$3.46*10^{-2}$
6	250	0.232	0.878	1.01
7	300	0.232	0.878	1.01
8	350	0.232	0.878	1.01
9	400	0.142	0.098	$3.55*10^{-2}$



729

730 **Figure 22.** Order of measurement number 2 (Tab. 10), trained ANN RBF model with data 8.2.2015 to
731 14.2.2015, within cross-validation with data (15.6.2015 to 21.6.2015).

732 A third cross-validation was performed by using 15.2.2015 as training data and 16.2.2015 for
733 validation. These periods of times share similar conditions and smaller data size, therefore an
734 improvement in cross-validation results were expected and achieved. The table 14 indicates that
735 similar to the first implementation, the measurement number 1 holds the best results in terms of linear
736 correlation, mean absolute error and mean square error. Additionally, almost identical (with minor
737 glitches) reference values and predicted values can be observed from Figure 23.

738 **Table 11.** Trained ANN RBF model with data 15.2.2015, within cross-validation with data
739 (16.2.2015).

Order of measurement	*Number of neurons (-)*	*R(-)*	*MAE (ppm)*	*MSE(ppm)*
1	10	0.99	0.002	$1.81*10^{-5}$
2	50	0.621	0.023	$7.43*10^{-4}$
3	100	0.621	0.023	$7.43*10^{-4}$
4	150	0.621	0.023	$7.43*10^{-4}$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

5	200	0.621	0.023	$7.43*10^{-4}$
6	250	0.621	0.023	$7.43*10^{-4}$
7	300	0.621	0.023	$7.43*10^{-4}$
8	350	0.621	0.023	$7.43*10^{-4}$
9	400	0.621	0.023	$7.43*10^{-4}$

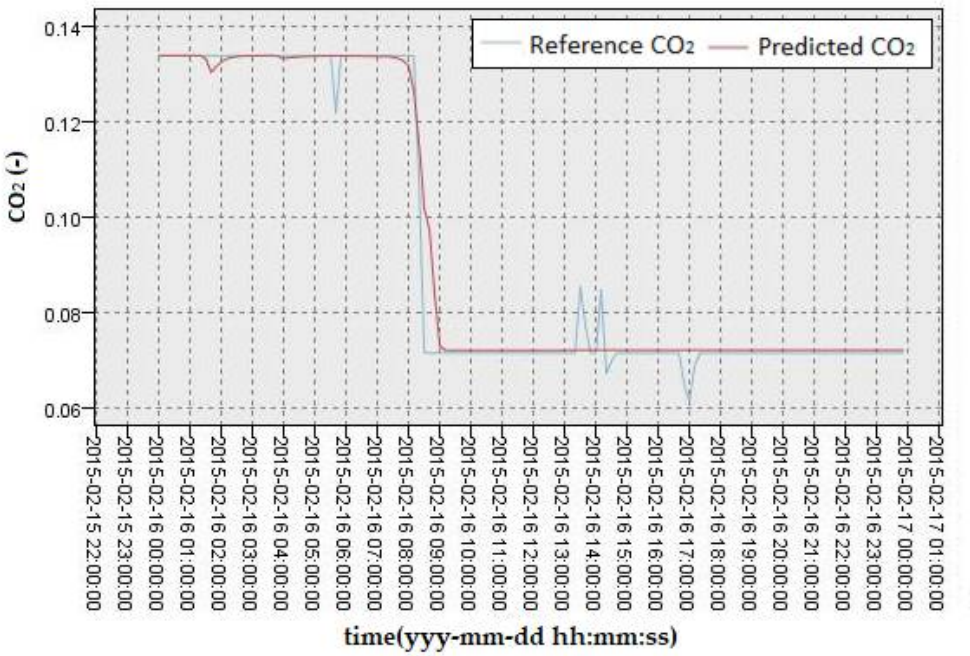


Figure 23. Order of measurement number 1 (Tab. 11), trained ANN RBF model with data 15.2.2015 within cross-validation with data (16.2.2015).

For the last cross-validation, the data from 8.6.2015 was chosen as training period and data from 8.2.2015 for cross-validation. Table 12 shows that except for order of measurement number 9 all other measurement represent very poor result. The prediction result from the measurement number 9 may significantly increase by appropriate filtering method (see figure 24).

Table 12. Trained ANN RBF models with data 8.6.2015, within cross-validation with data (8.2.2015).

Order of measurement	*Number of neurons (-)*	*R(-)*	*MAE (ppm)*	*MSE(ppm)*
1	10	-0.053	0.275	$1.12*10^{-2}$
2	50	-0.564	0.319	$1.73*10^{-2}$
3	100	-0.564	0.319	$1.73*10^{-2}$
4	150	-0.564	0.319	$1.73*10^{-2}$
5	200	-0.564	0.319	$1.73*10^{-2}$
6	250	-0.564	0.319	$1.73*10^{-2}$
7	300	-0.564	0.319	$1.73*10^{-2}$
8	350	-0.564	0.319	$1.73*10^{-2}$
9	400	0.818	0.278	$1.42*10^{-2}$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2018 doi:10.20944/preprints201812.0009.v1

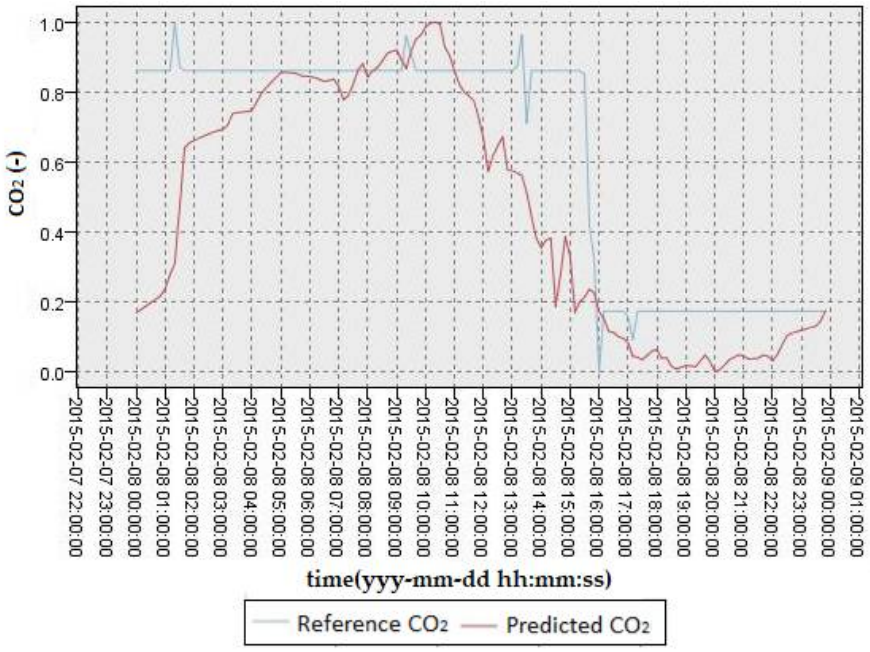


Figure 24. Order of measurement number 9 (Tab. 12), trained ANN RBF model with data 8.6.2015 within cross-validation with data (8.2.2015).

5. Discussion

As it is obvious, the predicted CO_2 signals contain lots of significant occurrences represented by the glitches and spikes, significantly deteriorating a smoothness of the analyzed signals. Such steep fluctuations may have significant impact on the CO_2 accuracy. Wavelet appears to be a reliable alternative for reduction such signal's parts. On the other hand we are aware that trend detection, in some cases, reduce the peaks thus, the original signal's amplitude is reduced, such situations are reported on the Fig. 15 and 16. In our work, we have studied the Daubechies wavelet family. This wavelets, as it is known, are able to well reflect morphological structure of the signals. We are particularly using Db6 wavelet for the trend detection. In the last part of our analysis, the objective comparison is carried out. As we have already stated, we are comparing originally measured CO_2 signals with predicted signals being filtered out by the wavelets. In order to carry out the objective comparison, the following parameters are considered: when considering the MSE, we are getting better results for the wavelet trend detection. It means that we have minimized the difference between the gold standard and filtered signals. The correlation coefficient gives higher values for the predicted CO_2 signals. Anyway, we have achieved just slight differences. The reason might be caused by the fact that the trend detection largely omits higher peaks, therefore, the linear dependence for wavelet filtration is smaller, when compared with the predicted signals. Generally speaking, using the Wavelet filtration lead to more accurate results against the predicted signals, and signals are much more smoothed, not containing steep fluctuations. On the other hand we are aware a certain loss of the amplitude. Therefore, in the future time it would be worth investigating the frequency features of the CO_2 signals in order to objectively determine frequency modifications while filtering by the wavelets.

By reviewing the results from measurement performed using radial basis function networks developed in IBM SPSS modeler, it can be observed that however, the measurement number 1 showed the poorest accuracy in initial training, it provided the most accurate result for case of cross-validation of 8.6.2015 to 14.6.2015 with 15.2.2015 to 21.2.2015 which could lead to possible human activity recognition with proper filtering technique. By swapping the training and validation periods the prediction accuracy significantly the dropped. This drop was a direct result of the lower quality of training data due to the smaller amount of abnormities caused by the presence of humans during

 doi:10.20944/preprints201812.0009.v1

February. In the cross-validation of cross-validation of 15.2.2015 with 16.2.2015, it provided a very accurate result. This improvement of the result is caused by the similarity of the relationship between the data from these two days, therefore, a significant reduction of quality will appear in case of choosing different data form different times of the year. This case was examined in the last cross-validation case for a day from June for training and day from February for validation, which resulted from invalid results for this measurement. By comparing the results from cross-validations performed with measurement 1, it can be observed that for periods of times that share similar conditions and smaller size of data, it performs the best but in other cases, the accuracy significantly drops. This is also could a good indication of overfitting which results in a bad generalization. Other measurements also share similar conditions with measurement number 1 but with more significant inconsistency in the results.

6. Conclusions

The article describes visualization of the operational and technical functions in Smart Home Care using the PI Process Book software tool. In the context of ADL monitoring in the individual rooms, a method for predicting the CO_2 course from the measured temperature and humidity courses using ANN BRM with signal trend detection based on WT for additive noise cancellation from the predicted signal was introduced and verified.

In the first part of the article were used and processed information from operationally measured non-electrical quantities determining the indoor environment in the SHC using operational technological units for the determination of the ADL in a real-world SHC environment for obtain an overview of the ADL of individual rooms of the SHC (time of arrival, time of departure), the indirect measurement of CO_2 concentration (ppm) with operational CO_2 (ppm) sensors was used.

In the second part of the article were used ANN BRM structures to predict the measured quantities for the purpose of monitoring the ADL in a real-world SHC environment. There was described the process of using the multilayer forward ANN to predict the course of CO_2 concentration from the measured temperature Ti (°C), relative humidity rH (%) in the interior of the SHC in the selected room R204 and from the outdoor temperature readings To (°C), with the gradient algorithm of error backpropagation using the BRM prediction. For the classification of prediction quality, a correlation analysis (R), calculated MSE, ED, CB were used.

In the third part of the article were made verifications and comparisons of WT additive noise cancelation in ANN BRM prediction of the course of CO_2 concentration from the measured temperature Ti (°C), relative humidity rH (%) in the interior of the SHC in the selected room R204 and from the outdoor temperature readings To (°C). For the classification of prediction quality with WT additive noise cancelation, a correlation analysis (R), calculated MSE, ED, CB were used. Based on the results, the WT is capable of filtering rapid signal changes, whilst preserving the signal trend.

In the fourth part of the article were ANN RBF structures to predict the measured quantities for the purpose of monitoring the ADL in a real-world SHC environment using of SPSS IBM SW Tool within IoT platform implementation to predict the course of CO_2 concentration from the measured temperature Ti (°C), relative humidity rH (%) in the interior of the SHC in the selected room R204 and from the outdoor temperature readings To (°C) with the ANN RBF structure. For the classification of prediction quality, a correlation analysis (R), calculated MSE and MAE were used. The IBM SPSSS Modeler provides a strong software tool for predictive analytics that includes many predictive models. By observing the result from all cross-validations performed, it is clear that the ANN RBF model used in IBM SPSS Modeler has a poor generalization. In future works with IBM SPSS within IoT, the results form varies types of decision trees will be examined and compared. Additionally, different predictive models can be implemented using IoT IBM Watson analytics, which can provide online analytics with possibilities of near real-time data streaming.

Author Contributions: J.V. methodology; J.V., J.Ku., O.G. software; J.V., J.Ku. validation; J.V., J.Ku., O.G. formal analysis; J.V. investigation; J.V., O.G. resources; J.V. data curation; J.V., K.Ku., O.G. writing—original draft preparation; J.V., O.G. writing—review and editing; J.V. visualization; J.V. supervision; J.Ko. project administration; J.Ko. funding acquisition.

Funding: This research was funded by the Ministry of Education of the Czech Republic (Project No. SP2018/170) and by the European Regional Development Fund in the Research Centre of Advanced Mechatronic Systems project, project number CZ.02.1.01/0.0/0.0/16_019/0000867 within the Operational Programme Research, Development and Education.

Acknowledgments: This article was supported by the Ministry of Education of the Czech Republic (Project No. SP2018/170). This work was supported by the European Regional Development Fund in the Research Centre of Advanced Mechatronic Systems project, project number CZ.02.1.01/0.0/0.0/16_019/0000867 within the Operational Programme Research, Development and Education. The work and the contributions were supported by the project SV4507741/2101, 'Biomedicínské inženýrské systémy XIII'.

Conflicts of Interest: "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results".

References

1. J.S. Beaudin, S.S. Intille and M.E Morris. To Track or Not to Track: User Reactions to Concepts in Longitudinal Health Monitoring. Journal of Medical Internet Research [online]. 2006, 8(4), e29- [cit. 2018-04-04]. DOI: 10.2196/jmir.8.4.e29.
2. M.J. Booysen, Machine-to-Machine (M2M) Communications in Vehicular Networks. KSII Transactions on Internet and Information Systems [online]. 2012, , - [cit. 2017-11-07]. DOI: 10.3837/tiis.2012.02.005.
3. D. Basu, G. Moretti, G. S. Gupta and S. Marsland. Wireless sensor network based smart home: Sensor selection, deployment and monitoring. In: 2013 IEEE Sensors Applications Symposium Proceedings [online]. IEEE, 2013, 2013, s. 49-54 [cit. 2017-11-07]. DOI: 10.1109/SAS.2013.6493555.
4. S. Fleck, and W. Strasser. Smart Camera Based Monitoring System and Its Application to Assisted Living. Proceedings of the IEEE [online]. 2008, 96(10), 1698-1714 [cit. 2017-11-07]. DOI: 10.1109/JPROC.2008.928765.
5. A. Pantazaras, S. E. Lee, M. Santamouris, and J. Yang, "Predicting the CO2 levels in buildings using deterministic and identified models," Energy and Buildings, vol. 127, pp. 774-785, 2016/09/01/ 2016.
6. G. J. Ríos-Moreno, M. Trejo-Perea, R. Castañeda-Miranda, V. M. Hernández-Guzmán, and G. Herrera-Ruiz, "Modelling temperature in intelligent buildings by means of autoregressive models," Automation in Construction, vol. 16, no. 5, pp. 713-722, 2007/08/01/ 2007.
7. M. Aggarwal and M. Madhukar, "IBM's Watson analytics for health care: A miracle made true," in Cloud Computing Systems and Applications in Healthcare , 2016, pp. 117-134.
8. A. Kaur and A. Jasuja, "Health monitoring based on IoT using Raspberry PI," in 2017 International Conference on Computing, Communication and Automation (ICCCA) , 2017, pp. 1335-1340.
9. Petnik, and J. Vanus, "Design of Smart Home Implementation within IoT with Natural Language Interface," *Ifac Papersonline,* vol. 51, no. 6, pp. 174-179, 2018
10. McEwen, A.; Cassimally, H. "Designing the Internet of Things", 1st ed.; John Wiley & Sons Ltd: The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom, 2008; p. 9, ISBN: 978-1-118-43062-0.
11. B.Xu, J.Zheng, Q.Wang "Analysis and Design of Real-Time Micro-Environment Parameter Monitoring System Based on Internet of Things," *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*.
12. Q. Min, Y. F. Ding, T. Xiao *et al.*, "Research of Visualization Monitoring Technology Based on Internet of Things in Discrete Manufacturing Process," *2015 2nd International Symposium on Dependable Computing and Internet of Things (Dcit)*, pp. 128-133, 2015
13. Y. Wang, J. Song, X. Liu et al., "Plantation Monitoring System Based on Internet of Things" 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing pp. 366-369.
14. Y. E. Windarto, and D. Eridani, "Door And Light Control Prototype Using Intel Galileo Based Internet of Things," *2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (Icitacee)*, pp. 176-180, 2017.
15. C. Coelho, D. Coelho, M. Wolf *et al.*, "An IoT Smart Home Architecture for Long-Term Care of People with Special Needs," *2015 Ieee 2nd World Forum on Internet of Things (Wf-Iot)*, pp. 626-627, 2015.

16. Oxford dictionaries, "Definition of big data in English". Available online: https://en.oxforddictionaries.com/definition/big_data (accessed on 25/11/2018)
17. C. Nyce, "Predictive Analytics White paper," 2007. Available online: https://www.the-digital-insurer.com/wp-content/uploads/2013/12/78-Predictive-Modeling-White-Paper.pdf (accessed on 25/11/2018).
18. M. Rouse. "Predictive modeling,". Available online: https://searchenterpriseai.techtarget.com/definition/predictive-modeling (accessed on 25/11/2018).
19. M. W. Ahmad, J. Reynolds, and Y. Rezgui, "Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees," *Journal of Cleaner Production*, vol. 203, pp. 810-821, Dec, 2018.
20. "IBM SPSS software,". Available online: https://www.ibm.com/analytics/spss-statistics-software (accessed on 25/11/2018).
21. "IBM Watson,". Available online: https://www.ibm.com/watson/. (accessed on 25/11/2018).
22. K. K. Nagwanshi, and S. Dubey, "Statistical Feature Analysis of Human Footprint for Personal Identification Using BigML and IBM Watson Analytics," Arabian Journal for Science and Engineering, vol. 43, no. 6, pp. 2703-2712, Jun, 2018
23. T. Perumal, S. K. Datta, C. Bonnet *et al.*, "IoT Device Management Framework for Smart Home Scenarios," *2015 Ieee 4th Global Conference on Consumer Electronics (Gcce)*, pp. 54-55, 2015.
24. O. Arnold, L. Kirsch, A. Schulz, and Ieee, "An Interactive Concierge for Independent Living," (in English), *2014 Ieee 3rd Global Conference on Consumer Electronics (Gcce)*, Proceedings Paper pp. 59-62, 2014.
25. J. R. Carvalko and Ieee, "LAW AND POLICY IN AN ERA OF CYBORG-ASSISTED-LIFE THE IMPLICATIONS OF INTERFACING IN-THE-BODY TECHNOLOGIES TO THE OUTER WORLD," in *2013 Ieee International Symposium on Technology and Society*(IEEE International Symposium on Technology and Society, New York: Ieee, 2013, pp. 204-215.
26. P. Cervenka, I. Hlavaty, A. Miklosik, and J. Lipianska, "Using cognitive systems in marketing analysis," (in English), *Economic Annals-Xxi*, Article vol. 160, no. 7-8, pp. 56-61, Oct 2016.
27. Y. Chen, E. Argentinis, and G. Weber, "IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research," (in English), *Clinical Therapeutics*, Review vol. 38, no. 4, pp. 688-701, Apr 2016.
28. M. Coccoli, P. Maresca, L. Stanganelli, and K. S. I. R. Inc, *Teaching Computer Programming Through Hands-on Labs on Cognitive Computing* (Dms 2016: the 22nd International Conference on Distributed Multimedia Systems). Skokie: Knowledge Systems Institute, 2016, pp. 158-164.
29. M. Devarakonda, D. Zhang, C. H. Tsou, M. Bornea, and Ieee, "Problem-Oriented Patient Record Summary: An Early Report on a Watson Application," (in English), *2014 Ieee 16th International Conference on E-Health Networking, Applications and Services (Healthcom)*, Proceedings Paper pp. 281-286, 2014.
30. G. Guidi, R. Miniati, M. Mazzola, and E. Iadanza, "Case Study: IBM Watson Analytics Cloud Platform as Analytics-as-a-Service System for Heart Failure Early Detection," (in English), *Future Internet*, Article vol. 8, no. 3, p. 16, Sep 2016, Art. no. 32.
31. E. Kolker and V. Ozdemir, "How Healthcare Can Refocus on Its Super-Customers (Patients, n=1) and Customers (Doctors and Nurses) by Leveraging Lessons from Amazon, Uber, and Watson," (in English), *Omics-a Journal of Integrative Biology*, Review vol. 20, no. 6, pp. 329-333, Jun 2016.
32. S. S. Murtaza, P. Lak, A. Bener, and A. Pischdotchian, "How to Effectively Train IBM Watson: Classroom Experience," in *2016 49th Hawaii International Conference on System Sciences*, T. X. Bui and R. H. Sprague, Eds. (Proceedings of the Annual Hawaii International Conference on System Sciences, Los Alamitos: Ieee Computer Soc, 2016, pp. 1663-1670.
33. F. AlFaris, A. Juaidi, and F. Manzano-Agugliaro, "Intelligent homes' technologies to optimize the energy performance for the net zero energy home," (in English), Energy and Buildings, Review vol. 153, pp. 262-274, Oct 2017.
34. M. Alirezaie *et al.*, "An Ontology-based Context-aware System for Smart Homes: E-care@home," (in English), *Sensors*, Article vol. 17, no. 7, p. 23, Jul 2017, Art. no. 1586.
35. M. Bassoli, V. Bianchi, and I. De Munari, "A Plug and Play IoT Wi-Fi Smart Home System for Human Monitoring," (in English), *Electronics*, Article vol. 7, no. 9, p. 13, Sep 2018, Art. no. 200.

939 36. P. A. Catherwood, D. Steele, M. Little, S. McComb, and J. McLaughlin, "A Community-Based IoT
940 Personalized Wireless Healthcare Solution Trial," (in English), *Ieee Journal of Translational Engineering*
941 *in Health and Medicine-Jtehm*, Article vol. 6, p. 13, 2018, Art. no. Unsp 2800313.
942 37. I. Skotnicova, L. Lausova, V. Michalcova, J. Vanus, and Sgem, "TEMPERATURES AND HEAT
943 TRANSFER BENEATH A GROUND FLOOR SLAB IN A PASSIVE HOUSE," in Nano, Bio and Green
944 - Technologies for a Sustainable Future Conference Proceedings, Sgem 2016, Vol Ii(International
945 Multidisciplinary Scientific GeoConference-SGEM, Sofia: Stef92 Technology Ltd, 2016, pp. 269-276.
946 38. J. Vanus et al., "Monitoring of the daily living activities in smart home care," *Human-centric Computing*
947 *and Information Sciences*, Article vol. 7, no. 1, 2017, Art. no. 30.
948 39. J. Vanus et al., "Design of an application for the monitoring and visualization of technological processes
949 with pi system in an intelligent building for mobile devices," in Proceedings of the 9th International
950 Scientific Symposium on Electrical Power Engineering, *ELEKTROENERGETIKA 2017*, 2017, pp. 518-
951 522.
952 40. J. Vanus, Z. Machacek, J. Koziorek, W. Walendziuk, V. Kolar, and Z. Jaron, "Advanced energy
953 management system in Smart Home Care," *International Journal of Applied Electromagnetics and*
954 *Mechanics*, Conference Paper vol. 52, no. 1-2, pp. 517-524, 2016.
955 J. Vanus, R. Martinek, P. Bilik, J. Zidek, P. Dohnalek, and P. Gajdos, "New method for accurate
956 prediction of CO2 in the Smart Home," in Conference Record - *IEEE Instrumentation and Measurement*
957 *Technology Conference*, 2016, vol. 2016-July.
958 41. "IBM SPSS Modeler 16 Algorithms Guide," Available online:
959 ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/16.0/en/AlgorithmsGuide.
960 pdf (accessed on 25/11/2018).
961 42. "IBM SPSS Modeler 17 Algorithms Guide," Available online:
962 ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/17.0/en/AlgorithmsGuide.
963 pdf (accessed on 25/11/2018).
964 43. "IBM SPSS Modeler 18 Algorithms Guide," Available online:
965 ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/18.0/en
966 /AlgorithmsGuide.pdf (accessed on 25/11/2018).