INTRODUCTORY ANALYSIS OF GAS CONSUMPTION TIME SERIES IN NON-RESIDENTIAL BUILDINGS FOR PREDICTION PURPOSES USING WAVELET DECOMPOSITION

ALEXANDER HOSOVSKY, JAN PITEL,

JANA MIZAKOVA, KAMIL ZIDEK

Faculty of Manufacturing Technologies with seat in Presov, Technical University in Kosice, Presov, Slovak Republic

Gas consumption prediction in buildings is very important with regard to the improved decision making and better energy utilization rate. Our objective is to perform introductory analysis of gas consumption time series in three different types of non-residential buildings (elementary school, national health institute and railway station) using wavelet transform intended for prediction model identification. We use initial FFT analysis of frequency spectrum using which it was possible to identify interesting frequency components related to specific consumption patterns. In order to find out the optimal level of wavelet decomposition we use entropy-based algorithm applied to maximum level wavelet trees. It was found that gas consumption time series that optimal wavelet decomposition level in elementary school time series was 3 and other two objects was 6. Using sample autocorrelation function plots for obtained components, we were able to the components containing mainly noise which could be removed from prediction model.

KEYWORDS

wavelet analysis, cyclic component, gas consumption, frequency spectrum, temperature

1 INTRODUCTION

The problem of prediction of time series – which is a sequence of time observations taken sequentially [Box 2016] – appears in many areas and is currently of great interest. The prediction of energy use in buildings is of particular interest in effort to improve the energy performance and reducing environmental impact [Zhao 2012]. Wavelets is one of the most powerful signal processing tools allowing for the analysis of signals on several timescales of the local properties of complex signals [Misiti 2007] and have great potential to be applied in the field of time series prediction/forecasting.

A high number of recent works utilizing wavelet analysis and wavelet-hybridized models attest to extreme importance of this technique in the field of time series forecasting regardless the field of application including electrical power engineering [Rana 2016, Li 2015], wind power engineering [Chitsaz 2015, Yu 2017], environmental engineering [Zainuddin 2011], hydrology [Adamowski 2011, Raj 2017] and thermal engineering [Panapakidis 2017].

[Soltani 2002] might be considered one of the seminal works in the area of wavelet analysis-based time series prediction where wavelet decomposition of the original time series is used in order to obtain components that might be better to predict separately. Zhang and other researchers in [Zhang 2017] compared using traditional approach with ARMA/ARIMA models and wavelet-ARMA/ARIMA models for predicting PM_{10} time series and shown that the latter could reduce the forecasting error and realize multi-scale prediction. In contrast to that, a wavelet neural network approach was used in [Adamowski 2011] to predict groundwater level forecasting and performance was compared to that of common artificial network and ARIMA model. This approach was found to be more accurate in predicting monthly average groundwater level. Similarly, in [Chitsaz 2015] a neural network with Morlet wavelets used in activation functions and trained by means of new improved Clonal selection algorithm was utilized for predicting wind power which is important for the operation of wind farms. Wavelet neural network was also used in [Wang 2017] where Variational Mode Decomposition (VDM) method was first used for the decomposition of predicted time series to improve the prediction accuracy and then multi-step ahead forecasting using GA-optimized neural network model was performed with input-output pairs being selected using phase space reconstruction method. Researchers in [Raj 2017] used the combination of standard ANN and wavelet decomposition to predict water table depth with respect to the rainfall in given region. The wavelet decomposition of given time series is usually combined with linear or nonlinear model used for prediction which is quite often of ARIMA type (linear) or ANN type (nonlinear). It is then interesting to compare the performance of these models in given application which was done in [Nury 2017] for temperature time series prediction. Initial time series decomposition using wavelet transform was also used in [Panapakidis 2017] where the obtained components (details and approximation) were predicted using hybrid approach with GA-optimized ANFIS and feedforward neural network. Since detail components at lower scales may contain a significant amount of noise and be thus difficult to predict, it was proposed in [Yu 2017] to use singular spectrum analysis (SSA) to further process highest frequency component. The prediction model itself was based on Elman neural network [Hosovsky 2015] in contrast to commonly used FFNN. In cases where linear model can be used for forecasting, wavelet transform can improve the forecasting accuracy by performing the time series decomposition – such an approach was used in [Choi 2011] where SARIMA model was combined with wavelet transform for sales forecasting and outperformed pure SARIMA model, linear extrapolation with seasonal adjustment as well as evolutionary neural network. In [Maheswaran 2015] the combination of wavelet transform and high order Volterra models was used together with Kalman filter and tested on synthethically generated and actual time series.

It is quite obvious even from this very short review of works related to the use of wavelet transform in time series forecasting that either wavelet decomposition is performed initially followed by the use of mostly nonlinear prediction model (very often in the form of neural network) or it can be directly embedded into the structure of neural network in the form of wavelet activation functions. Our objective is to analyze the gas consumption time series from three non-residential buildings which differ in their consumption profiles with the purpose of using them for prediction purposes. We thus apply the former of the aforementioned approaches where time series are initially decomposed using the wavelet transform. It is expected that

using WT decomposition, the final prediction error can be decreased compared to the approach where prediction model is identified directly from the original time series. In addition to that, it remains open question how to determine the optimal wavelet decomposition level. While it is often done based on some heuristic criteria, we use more rigorous entropy-based (logenergy) criterion which is typically used in wavelet packet decomposition. On the other hand, to determine the optimal level we applied it on full wavelet tree with maximum decomposition level dictated by the number of samples. To estimate the predictability of particular components, sample autocorrelation plots are examined.

2   USED METHODS

According to [Tangirala 2015], continuous wavelet transform (CWT) of signal $x(t)$ is defined as

$$Wx(\tau, s) = \frac{\langle x, \psi_{t,s} \rangle}{\|\psi_{t,s}\|_2^2} = \int_{-\infty}^{+\infty} x(t), \psi_{t,s}^*(t)dt \tag{1}$$

where $\psi_{t,s}$ is wavelet generated by scaling and translation of so-called mother wavelet $\psi(t)$

$$\psi_{t,s} = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-\tau}{s}\right), \quad \tau, s \in \mathbb{R} \quad s \neq 0 \tag{2}$$

for which the following must hold

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \qquad \|\psi(t)\|_2 = 1 \tag{3}$$

Wavelet transform is interesting from filtering perspective due to its adaptable band-pass filter capability, which can be controlled through scaling by shifting the center frequency and the width of pass band. The Eq.1 can then be rewritten to the following form

$$Wx(\tau, s) = x * \bar{\psi}_s(\tau), \ where \ \bar{\psi}_s(\tau) = \frac{1}{\sqrt{s}} \psi^*(\tau) \tag{4}$$

where $*$ - convolution.

If s = 1 as a reference point the filters for $s > 1$ are replaced by a low-pass filter, which may be considered as equivalent to introducing a scaling function $\varphi(t)$

$$|\hat{\varphi}(\omega)|^2 = \int_1^\infty |\hat{\psi}(s\omega)|^2 \frac{ds}{s} \ and \ \lim_{\omega \to 0} |\hat{\varphi}(\omega)|^2 = C_\varphi \tag{5}$$

Using this scaling function, the following transform can be defined

$$Sx(\tau, s) = \langle x(t), \varphi_{\tau,s}(t) \rangle = x * \bar{\varphi}_s(\tau) \tag{6}$$

Then, the signal can be re-written into the form which is a sum of an approximation and details at scale s_0:

$$x(t) = \frac{1}{C_\psi s_0} Sx(.,s) * \varphi_{s0}(t) + \frac{1}{C_\psi} \int_0^{s_0} Wx(.,s) * \psi_s(t) \frac{ds}{s^2} \tag{7}$$

Here a discrete version of wavelet transform (DWT) is used, which is a special case of continuous wavelet transform where evaluation is done only at certain scales and translations, $s = 2^j, j \in \mathbb{Z}$ and $\tau = m2^j, m \in \mathbb{Z}$ [Tang 2009].

It can be defined in the following manner

$$Wf(m,j) = \int_{-\infty}^{+\infty} f(t)\psi_{m2^j,2^j}(t)dt \tag{8}$$

with

$$\psi_{m2^j,2^j}(t) = \frac{1}{2^{j/2}} \psi\left(\frac{t-m2^j}{2^j}\right) \tag{9}$$

The signal decomposition given in Eq.7 can now be expressed for a discrete case as follows:

$$x(t) = \sum_m a_{m,j_0}\varphi_{m,j_0}(t) + \sum_{j \geq j_0}\sum_m d_{m,j_0}\psi_{m,j_0}(t) =$$
$$= A_{j_0} + \sum_{j=1}^{j_0} D_j \tag{10}$$

where a_{m,j_0} - approximation coefficients at scale j_0 and $d_{m,j_0}, j = 1, \ldots, j_0$- detail coefficients at level j_0 and all finer scales, A_{j_0}- approximation component of signal $x(t)$ and D_{j_0}- detail components of signal $x(t)$ at level j_0.

In practical cases the values of both types of coefficients (approximation and detail) are computed with the help of impulse responses denoted with $\{g_l[.]\}$ and $\{g_h[.]\}$ corresponding to low- and high-pass filters, respectively. It is important that for computation of coefficient values (whether approximation or detail), the values of approximation details of one scale above are used :

$$a_{j+1}[m] = \sum_{n=-\infty}^{\infty} g_l[n - 2m]a_j[n] = (a_j * \bar{g}_l)[2m] \tag{11}$$
$$d_{j+1}[m] = \sum_{n=-\infty}^{\infty} g_h[n - 2m]a_j[n] = (d_j * \bar{g}_h)[2m] \tag{12}$$

with $\bar{g}_l = g[-n]$ denoting the IR sequence reflection and * - convolution. Then, actually four distinctive objects can be identified [Misiti 2007] :
I. Detail coefficients $d_{m,j_0} = \int x(t)\varphi_{m,j_0}(t)dt$
II. Detail signals $D_j(t) = \sum_{k \in \mathbb{Z}} a_{m,j_0} \psi_{m,j_0}(t)$
III. Approximation coefficients $a_{m,j_0} = \int x(t)\varphi_{m,j_0}(t)dt$
IV. Approximation signals $A_j(t) = \sum_{k \in \mathbb{Z}} a_{m,j_0} \varphi_{m,j_0}(t)$
In (discrete) wavelet transform (DWT), only approximation components of the signal are further decomposed into approximation and detail part while the latter are not analyzed in any way forming partial decomposition tree as shown in Fig.1. In wavelet packet transform the concept of full signal decomposition is utilized, where each of the detail components is decomposed in the same way as approximation components in DWT. This allows for a much richer analysis of given signal yet at the cost of increased computational complexity which is on the order of $O(N\log_2 N)$ in contrast to $O(N)$ for DWT [Sundarjan 2015].

If we assume the following, $w^0(t) = \phi(t), w^1(t) = \psi(t)$ the functions with odd superscripts $2j+1$ are generated from $w^j(t)$ using wavelet filter **g** and those with even superscripts (i.e. $2j$) are generated from $w^j(t)$ using the scaling filter **h** [Ruch 2009]. The wavelet packet function can be then defined in the following way:
Suppose that $\phi(t)$ generates an orthogonal multiresolution analysis $\{V_j\}_{j \in \mathbb{Z}}$ with associated wavelet function $\psi(t)$. The wavelet packet functions are defined by $w^0(t) = \phi(t), w^1(t) = \psi(t)$ and for $n = 2,3, \ldots$
$$w^{2n}(t) = \sqrt{2}\sum_{k \in \mathbb{Z}} h_k w^n(2t - k) \tag{13}$$

$$w^{2n+1}(t) = \sqrt{2}\sum_{k \in \mathbb{Z}} g_k w^n(2t - k) \tag{14}$$
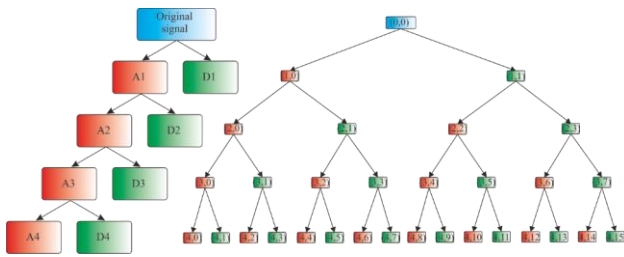
Figure 1. Fourth-level wavelet analysis (left) and wavelet packet analysis (right)

A signal containing $N = 2^L$ samples can be decomposed in γ different ways, where γ is the number of binary subtrees of a complete binary tree the depth of which is L. Then the following holds : $\gamma \geq 2^{N/2}$ [Misiti 2016]. Since this is obviously a huge number for typical time series (in our case the time series have a length in excess of 1000, giving $\gamma > 2^{500}$) entropy-based criteria are usually applied to find optimal decomposition. If s is the signal and s_i is i-th coefficient in orthonormal basis [Misiti 2016], the entropy can be defined as

$$E(s) = \sum_i E(s_i) \qquad (15)$$

so that it is an additive cost function with $E(0)=0$. We use logenergy criterion where

$$E(s_i) = \log(s_i^2) \quad \text{and} \quad E(s) = \sum_i \log(s_i^2) \qquad (16)$$

Given time series are actually decomposed to the form given by (10) using Eqs. (11) and (12). Then WPD method in Eqs.(13) and (14) is used to obtain a wavelet tree, in which an optimal level is determined using Eq.16.

3 ANALYZED OBJECTS AND USED DATA

Three different non-residential objects were selected for further analysis using wavelet transform. These objects are shown in Fig.2 and include: elementary school in Krompachy, National Institute of Rheumatic Diseases in Piestany and railway station in Tatranska Strba. Completely different purpose of each of the buildings meant that relatively distinctive gas consumption patterns could be expected. Each of the objects was subject to modernization of its boiler rooms with the installment of new Hoval UltraGas condensing boilers (Fig.3)


Figure 2. Analyzed objects – elementary school in Krompachy (left), National Institute of Rheumatic Diseases in Piestany (middle) and railway station in Tatranska Strba [DevelArt.sk 2018]
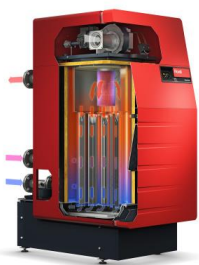

Figure 3. Hoval UltraGas condensing boiler used in each of the examined objects [Hoval.sk 2018]

The data of daily gas consumption for analysis were obtained during three-years period lasting from September 1, 2014 until September 1, 2017. Each data sample thus corresponds to the total gas consumption during given day in one of the buildings and one day was also used as the basic time unit in further analysis. The total number of days (and thus also samples) in every time series is 1097. These datasets can be quickly examined by looking at Fig.4, where the raw data (without any preprocessing) for all objects are shown. For prediction purposes it might be very important to take into account also the outdoor temperature which is naturally the single most important variable affecting the actual gas consumption in any of the analyzed buildings [Corny 2017]. Equally diverse as the purpose of objects in question was their geographic distribution within the country which also affected observed consumption profiles. Actual daily mean temperatures can be observed from Fig.5 together with the differences in whole period mean temperature. Only regularity that can be detected by inspecting raw time series of gas consumption is, as expected, a strong seasonal component which correspond with heating and non-heating periods during year. This period typically starts around the beginning of October and ends at the end of April, lasting around 213 days. Some basic data about the datasets can be found in Tab.1. Even though the object in Piestany is located in the warmest region (three years average of daily mean temperature is 5.1 and 6.6 °C higher than in Krompachy and Tatranska Strba respectively), absolute maximum and average values are dominated by the size of the object (Tab.1). The railway station is located in the coldest region (three years average of 7.8 °C) yet its small size gives lowest absolute maximum and whole period mean consumption values (approximately 258 m³ and 89 m³ respectively). Also, in contrast to other two objects, the gas consumption during non-heating period is significantly higher in National Institute in Piestany, caused almost surely by its specific purpose which necessitates non-negligible gas consumption even during this part of the year.
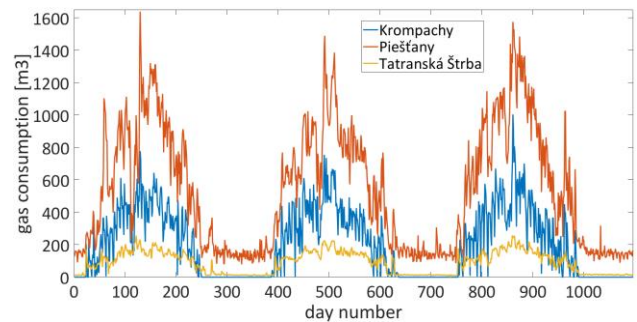
Figure 4. Raw data of gas consumption in three selected types of buildings during three consecutive heating periods (data correspond to a period from September 1, 2014 to September 1, 2017)
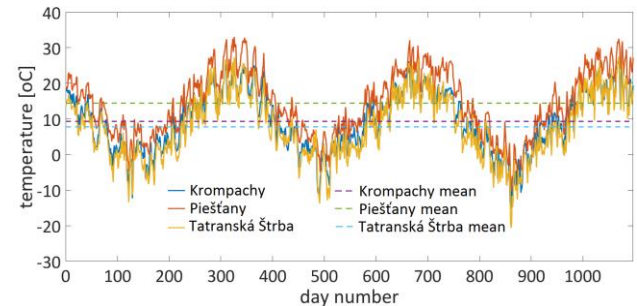
Figure 5. Daily mean and whole period mean temperatures at given locations during observed period (from September 1, 2014 to September 1, 2017)

Table 1. Basic data about the gas consumption and temperature time series for given locations (gas in m³, temperature in ºC)

	Krompachy	Piestany	Tatranska Strba
Absolute maximum (gas)	1003.6	1638	257.9
Whole period mean (gas)	199.5	528.5	89.1
Heating periods mean (gas)	333.9	782.1	137.1
Standard deviation in heating periods (gas)	169.1	291.2	43.5
Absolute minimum (temp)	-15.2	-10.4	-20.6
Absolute maximum (temp)	26.2	33.1	27.5
Whole period mean (temp)	9.3	14.4	7.8

Although the main seasonal component corresponding to alternating heating and non-heating periods is clearly visible from time series shown in Fig.4, it is much more difficult to detect other cyclic components that might be present and which are related to typical patterns of human activity in given buildings. Analysis of frequency spectrum of time series may reveal the most important cyclic components, which was performed using FFT. Single-sided amplitude spectra for gas consumption in each of the analysed objects are shown in Fig.6 (Krompachy), Fig.7 (Piestany) and Fig.8 (Tatranska Strba).

## 4   WAVELET ANALYSIS OF GAS CONSUMPTION TIME SERIES

Frequency spectrum analysis of given time series reveal some interesting information about the cyclic components that are present in gas consumption waveforms. These components are affected by many factors but appear to be strongly related to the specific patterns of gas consumption in analyzed types of buildings. We can observe the single-sided frequency spectrum of Krompachy gas consumption time series in Fig.6. We marked the most prominent peaks that might correspond to interesting cyclic components related to typical activities in given type of building as well the number indicating the period of this component (in days). The most prominent of all (in each type of building) is naturally the seasonal component associated with alternating heating and non-heating periods during the year – which is also the only component visible in raw time series shown in Fig.4. The frequency of this component is in each case approximately 0.0027 Hz corresponding to a period of 365.63 days. Compared to objects in Piestany and Tatransk Strba, we can observe the presence of higher number of isolated peaks in frequency spectrum of Krompachy time series which can be attributed to a stronger manifestation of repeating patterns-related consumption profile – of these, biannual period (182.85 days) and weekly period (6.99 days) appear to be the most significant.
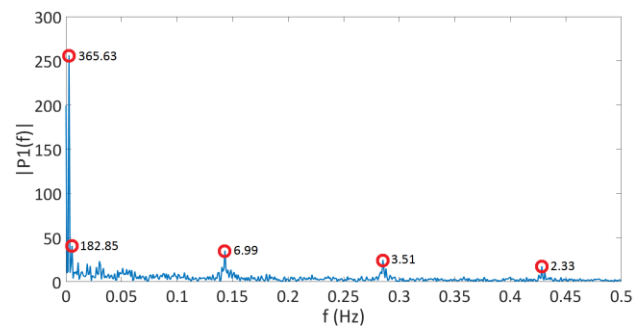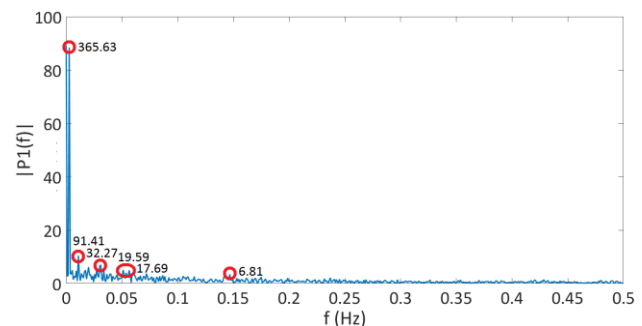

Figure 6. Single-sided amplitude spectrum of gas consumption time series at Krompachy location using FFT

Three of the prominent frequency components are also present in frequency spectrum of gas consumption time series in Piestany object (with period of 365.63, 182.85 and 6.99 days) (Fig.7). The first one naturally corresponds to a seasonal component of alternating heating and non-heating periods and the third one obviously corresponds to weekend-related drops in gas consumption. Biannual period found in both spectra can possibly be associated with the presence of two longer lasting holidays (in summer and in winter).

Apart from the strongest (annual) seasonal component of time series, the spectrum of gas consumption time series in Tatranska Strba (Fig.8) lacks the prominent peak corresponding to biannual period (182.85 days) yet it contains quarterly period (91.41 days) similar to National Institute. Although higher number of peaks in spectrum is indicated in Fig.8 compared to Piestany spectrum, it has to be noted that similar periodic components can be observed in both spectra with slightly smaller prominence in Fig.7 (this was evaluated only subjectively). On the other hand, easily observable is the absence of prominent higher frequency peaks in contrast to Krompachy spectrum.
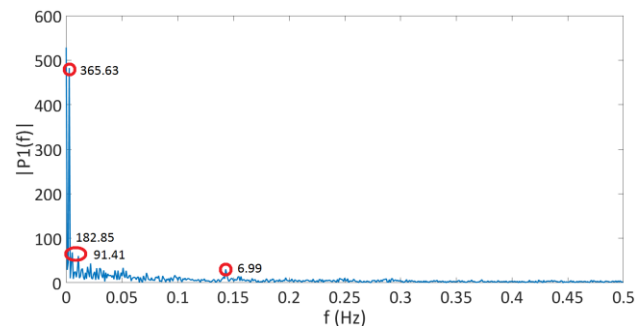
Figure 7. Single-sided amplitude spectrum of gas consumption time series at Piešťany location using FFT

Figure 8. Single-sided amplitude spectrum of gas consumption time series at Tatranská Štrba location using FFT

Although FFT method is valuable for identifiying possible periodic components, its usefulness for time series analysis related to prediction is limited due to the absence of time localization. In this case we can use wavelet analysis which

features time-frequency localization and is thus very advantageous to use for this purpose.

The maximum level of decomposition j_0 from Eq.10 depends on the number samples in time series N as $j_{0max} = \log_2 N$ and for 1097 samples yields $j_{0max} = \log_2 1097 = 10.1 \cong 10$. If time series prediction is the main purpose of wavelet analysis, such a high number of decomposition components (approximation and all details up to this level) is seldom used. Then it remains an open issue how to determine the optimal level of decomposition [Yang 2016]. By examining the spectra in Figs.6, 7 and 8 we could choose the decomposition level to isolate all periodic components that might be interesting in approximation (possibly annual period) and details (biannual or quarterly period, monthly and weekly period). If j_{0max} = 6 was used, the D6 signal would contain components with approximately monthly (32 days) and bimonthly (64 days) periods.

Another option is to base the selection of decomposition level on entropy criterion – specifically logenergy criterion show in Eq.16. For wavelet analysis we use db5 wavelet function from Daubechies wavelet family the shape of which, together with scaling function, can be seen in Fig.9. This particular wavelet family is very often applied in the area of time series prediction and it generated smooth approximation and detail signals that were deemed suitable for a given task (of gas consumption prediction). The initial wavelet tree corresponding to a maximum level dictated by the number of samples in time series is shown in Fig.10.
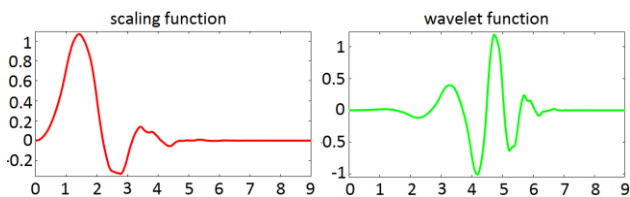

Figure 9. Db5 (Daubechies family) scaling function (left) and wavelet function (right)

The results of optimal selection of wavelet decomposition level can be observed in Fig.10. An initial wavelet tree of depth of 10 was considered for each of the series (shown in figure as first, third and fifth tree from left). Trees were cut at given depth based on the algorithm described in [Misiti 2016]. The entropy value shown in blue corresponds to initial entropy of a signal calculated using the logenergy criterion (Eq.16). It can be seen that optimal levels of decomposition were 6 for both Piestany and Tatranska Strba series and 3 for Krompachy series. Possible explanation for this might be related to the presence of higher number of volatile factors affecting the gas consumption profiles in Piestany and Tatranska Strba objects compared to elementary school in Krompachy.
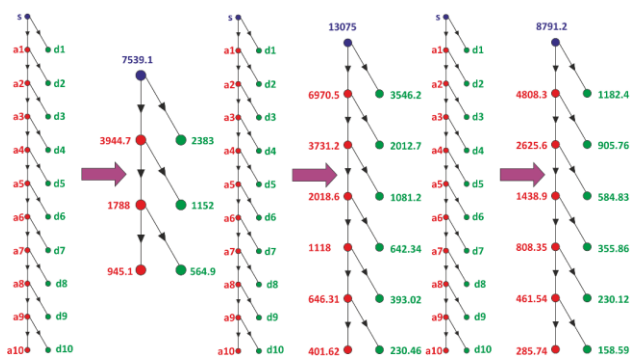

Figure 10. Initial wavelet tree and optimal level wavelet tree based on logenergy entropy criterion for each of the time series (Krompachy left, Piestany middle, Tatranska Strba right)

After making full decomposition with optimal levels, we obtain 3 approximation signals (a1, a2, a3) and 3 detail signals (d1, d2, d3) for Krompachy series and 6 approximation signals (a1, a2, a3, a4, a5, a6) and 6 detail signals (d1, d2, d3, d4, d5, d6) for both Piestany and Tatranska Strba series. The decomposition of original signals is then given as:

$$K = a3+d1+d2+d3$$

$$P = a6+d1+d2+d3+d4+d5+d6$$

$$TS = a6+d1+d2+d3+d4+d5+d6$$

where K – Krompachy time series, P – Piestany time series and TS – Tatranska Strba time series.

The wavelet decomposition at level 3 for Krompachy time series is shown in Fig.11. Performing the decomposition at this level allows us to observe the signals with periods between 1-2 days, 2-4 days, 4-8 days in details and greater than 8 days in approximation signal. In the case of approximation signal, in addition to the most prominent annual seasonal pattern also cycles with approximately monthly period are visible superimposed on the main series. Since the decomposition level was determined based on the entropy, we do not see this component isolated – this would be observable in d5 detail signal. It will be the subject to further research whether the optimal level based on entropy directly transfers to better results in predictions of gas consumption with wavelet decomposition.
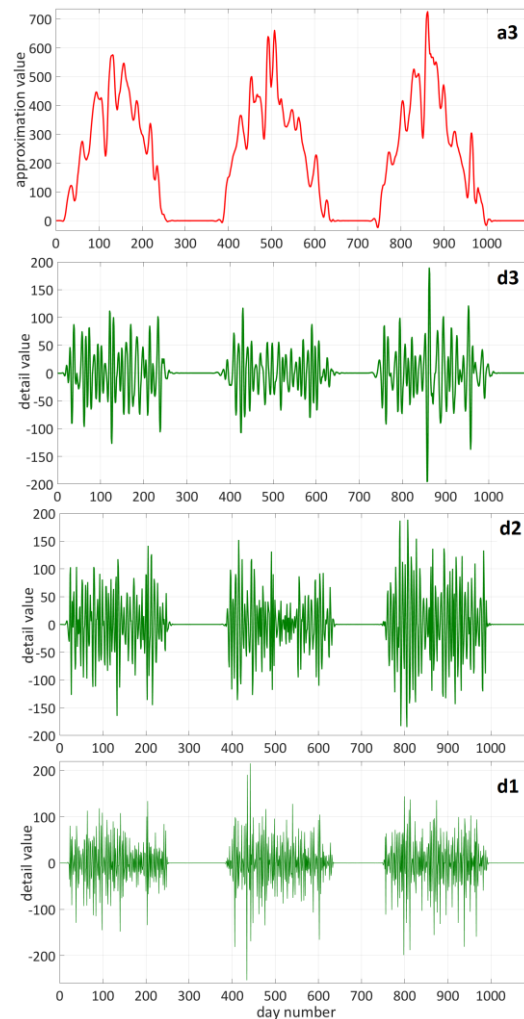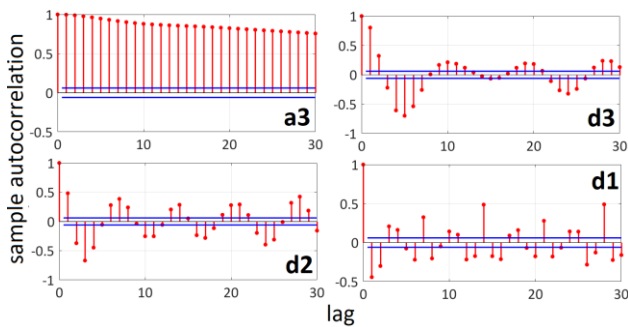

Figure 11. Level 3 wavelet decomposition of Krompachy time series (a3 – approximation signal at level 3, d1,d2,d3 – detail signals at levels 1,2 and 3)

In relation to its prediction, it is interesting to observe the autocorrelation function of given time series to see how it is correlated with the delayed copy of itself. Samples corresponding to lags with statistically significant non-zero value contain predictable information about the time series [Tangirala 2015]. Sample autocorrelation functions with 30 lag interval for each component of Krompachy time series after wavelet decomposition are shown in Fig.12. The ACF for approximation signal (a3) shows slow decay which is in accordance with ACF of non-stationary process. Brief inspection of all remaining ACFs for detail signals reveals also very strong presence of seasonal cycles. In case of d3 component the alternating period of 7 and 10 days can be observed in ACF while for d2 component the cycle with weekly period appears to be dominant. Identification of prediction model for these components could be done using an approach similiar to [Krivobokova 2012], where a combination of Bayesian splines was used for trend and seasonal component modeling and ARMA model was used for modeling residuals.


Figure 12. Sample autocorrelation functions for each component of Krompachy time series wavelet decomposition (a3, d1, d2, d3)

In accordance with the results of entropy-based determination of optimal wavelet decomposition, all seven component signals (a6 and d6, d5, d4, d3, d2, d1) are presented in Figs.13-19. Higher level of decomposition allows to obtain signal components with coarser frequency (i.e. longer periods). Approximation signals (a6) now reveal the course of the main seasonal component with roughly yearly period (365 days), which is typically associated with seasonal variations in gas consumption.
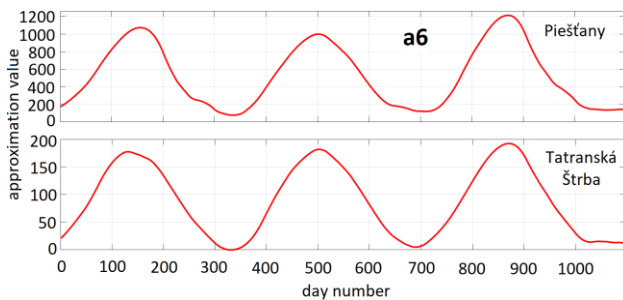

Figure 13. Approximation signals of Piestany and Tatranska Strba time series at level 6

D6 detail signal shown in Fig.14 reveals the components of original time series with time period between 32-64 days. This particular detail signal has more balanced amplitude in Tatranska Strba series compared to Piestany that might be related either to weather differences at given locations or other factors having greater effect on gas consumption profile in National institute than in railway station.
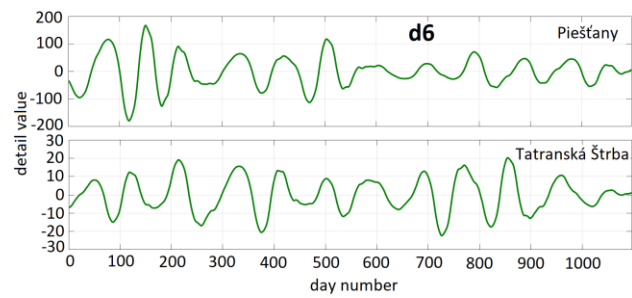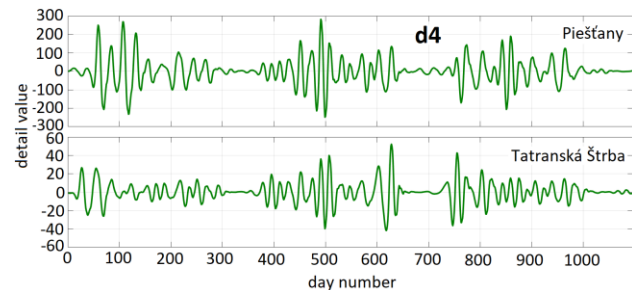

Figure 14. Detail signals of Piestany and Tatranska Strba time series at level 6

Detail signals d5 and d4 (shown in Figs.15, 16) contain components with time periods between 16-32 and 8-16 days respectively. Again, more balanced amplitude of this detail signal can be observed in Tatranska Strba series, with interruptions at non-heating periods in years 2015 and 2016. The variations in amplitude of detail value for particular components are most likely strongly related to outdoor temperatures during period in question throughout given years.
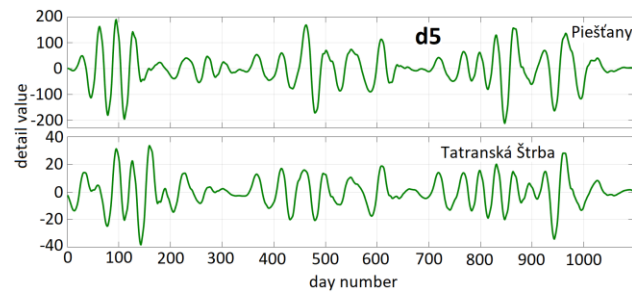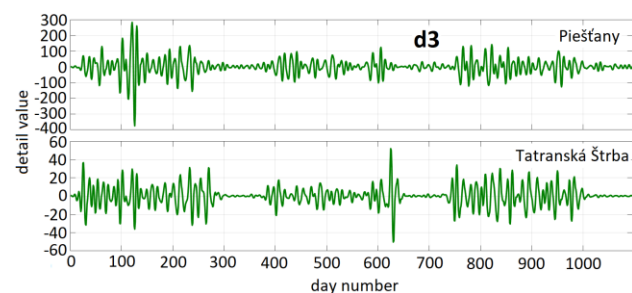

Figure 15. Detail signals of Piestany and Tatranska Strba time series at level 5

Figure 16. Detail signals of Piestany and Tatranska Strba time series at level 4

Detail signals d3, d2 and d1 shown in Fig.17, Fig.18 and Fig.19 contain components with periods between 4-8, 2-4 and 1-2 days respectively. The predictability of these components would be lower in comparison to previous ones due to the higher content of noise.

Figure 17. Detail signals of Piestany and Tatranska Strba time series at level 3
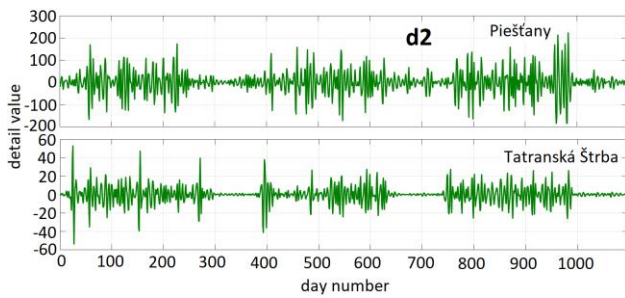
Figure 18. Detail signals of Piestany and Tatranska Strba time series at level 2
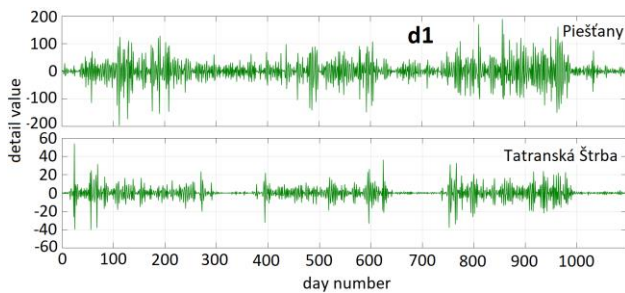

Figure 19. Detail signals of Piestany and Tatranska Strba time series at level 1

More useful information about the possible predictability of given component can be extracted from sample autocorrelation function plots, which are shown (for details) in Fig.20 for lag interval of 30. All of the plots show strong correlation between the successive lag samples with gradually decreasing time period – which is in accordance with increasing frequency of components passed through the wavelet filters. Less correlated lag samples can be observed only for d1 components of both time series where most of the correlation for higher lags is found within (or slightly out of) the confidence bounds. Even though in this case some portion of d1 component is still predictable, the proportion of noise within would make it likely candidate for removal from the final prediction model. Its inclusion or removal would depend on the requirements for the final prediction error (possibly with regard to accuracy/complexity of the resulting model).
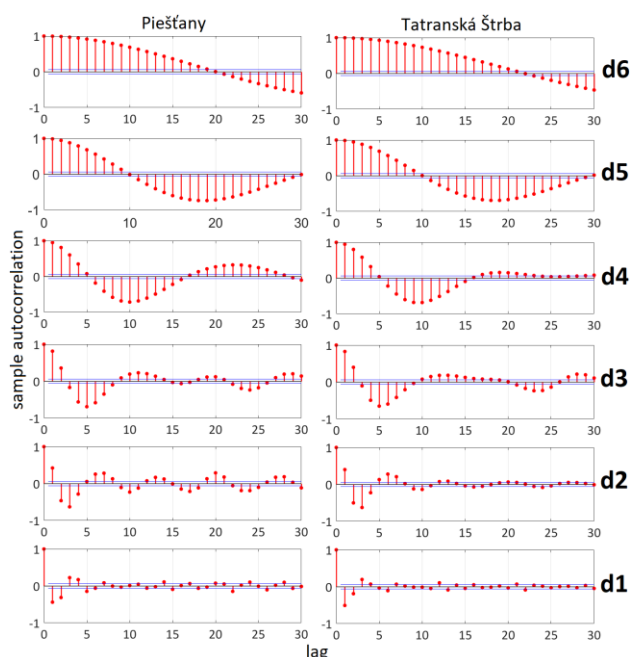

Figure 20. Sample autocorrelation function plots for detail signals of wavelet decomposition for Piestany and Tatranska Strba time series

5 CONCLUSION AND FURTHER RESEARCH

The main objective of the paper was to perform introductory analysis of gas consumption time series in different types of non-residential buildings intended using wavelet transform for prediction purposes. By performing initial FFT analysis of frequency spectrum of gas consumption profiles during three years, it was possible to identify the cycles that might be of interest in derivation of prediction models. In all analyzed time series naturally the most important cycle is annual seasonality which is associated with alternation of heating and non-heating periods. Other cycles corresponding to typical gas consumption patterns were also found – to isolate them as a function of time wavelet transform was used. In order to determine the optimal level of decomposition, logenergy entropy-based criterion was used and level 3 was found optimal for Krompachy time series and level 6 for both Piestany and Tatranska Strba time series. Sample autocorrelation function plots (lag interval 30) for each of the components confirmed strong correlation between the lags (with cyclic character) which implies possibly good predictability with approach typical for seasonal time series (SARIMA models for instance). On the other hand, D1 components of Piestany and Tatranska Strba time series showed only weak correlation, with ACF resembling that of white noise which could point to their possible removal from the components considered for prediction.

Several points will be subject to further research – it remains to be found out whether the optimal level determined based on the entropy criterion also translates to lower prediction errors in final prediction model and also what is the relationship between the particular components (approximation and details) and outdoor temperature in terms of prediction accuracy. This can be included into ARIMAX models which make use of exogenous variable – in this case outdoor temperature.

REFERENCES

[Adamowski 2011] Adamowski, J., Chan, H.F. A wavelet neural network conjunction model for groundwater level forecasting. Journal of Hydrology, September 2011, Vol.407, 2011, pp. 28-40. ISSN 0022-1694

[Box 2016] Box, G.E.P. et al. Time Series Analysis – Forecasting and Control. Hoboken: John Wiley & Sons, 2016. ISBN 978-1-118-67502-1

[Chitsaz 2015] Chitsaz, H. et al. Wind power forecast using wavelet neural network trained by improved Clonal selection algorithm. Energy Conversion and Management, January 2015, Vol.89, pp. 588-598. ISSN 0196-8904

[Choi 2011] Choi, T.-M. et al. A hybrid SARIMA wavelet transform method for sales forecasting. Decision Support Systems, April 2011, Vol.51, pp. 130-140. ISSN 0167-9236

[Corný 2017] Corny, I. Overview of Progressive Evaluation Methods for Monitoring of Heat Production and Distribution. Procedia Engineering, Vol. 190, 2017, pp. 619-626. ISSN 1877-7058

[DevelArt.sk 2018] DevelArt.sk Official webpage of Ekolmont s.r.o. company. Available from: https://www.vykurovanie-ekolmont.sk/

[He 2016] He, Z. Wavelet Analysis and Transient Signal Processing Applications for Power Systems. Singapore : John Wiley & Sons, 2016. ISBN 978-1118977033

[Hoval.sk 2018] www.hoval.sk Official webpage of Hoval company. Available from: http://www.hoval.sk

[Hosovsky 2015] Hosovsky, A., Pitel, J., Zidek, K. Enhanced Dynamic Model of Pneumatic Muscle Actuator with Elman Neural Network. Abstract and Applied Analysis, Vol. 2015, pp. 1-16. ISSN 1687-0409

[Krivobokova 2012] Krivobokova, T., Rosales, L.F. Detrending time series with cycle and seasonal components. In: International Workshop on Modern Methods in Financial Statistics, Suzhou, 21-23 May, 2012, 1-20

[Li 2015] Li, S. et al. Short-term load forecasting by wavelet transform and evolutionary extreme learning machine. Electric Power Systems Research, May 2015, Vol.122, pp.96-103, ISSN 0378-7796

[Maheswaran 2015] Maheswaran, R., Khosa, R. Wavelet Volterra Coupled Models for forecasting of nonlinear and non-stationary time series. Neurocomputing, February 2015, Vol.149, pp.1074-1084, ISSN 0925-2312

[Misiti 2007] Misiti, M. et al.(Eds.) Wavelets and their applications. London : ISTE Ltd., 2007. ISBN 978-1905209316

[Misiti 2016] Misiti, M. et al. Wavelet Toolbox – User's Guide. Natick : The MathWorks Inc., 2016.

[Nury 2017] Nury, A.H. et al. Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series in northeastern Bangladesh. Journal of King Saud University – Science, January 2017, Vol.29, pp.47-61, ISSN 1018-3647

[Panapakidis 2017] Panapakidis, I.P., Dagoumas, A.S. Day-ahead natural gas demand forecasting based on the combination of wavelet transform and ANFIS/genetic algorithm/neural network model. Energy, January 2017, Vol.118, pp.231-245, ISSN 0360-5442

[Raj 2017] Raj, A.S. et al. Wavelet based analysis on rainfall and water table depth forecasting using Neural Networks in Kanyakumari district, Tamil Nadu, India. Groundwater for Sustainable Development, June 2017, Vol.5, pp.178-186, ISSN 2352-801X

[Rana 2016] Rana, M., Koprinska, I. Forecasting electricity load with advanced wavelet neural networks. Neurocomputing, March 2016, Vol.182, pp.118-132, ISSN 0925-2312

[Ruch 2009] Ruch, D.K., Van Fleet, P.J. Wavelet Theory – An Elementary Approach with Applications. Hoboken : John Wiley & Sons, 2009. ISBN 978-0-470-38840-2

[Soltani 2002] Soltani, S. On the use of wavelet decomposition for time series prediction. Neurocomputing, October 2002, Vol.48, pp.267-277, ISSN 0925-2312

[Sundararajan 2015] Sundararajan, D. Discrete Wavelet Transform – A Signal Processing Approach. Singapore : John Wiley & Sons, 2015. ISBN 978-1119046066

[Tang 2009] Tang, Y.Y. Wavelet Theory Approach to Pattern Recognition. Singapore : World Scientific Publishing, 2009. ISBN 978-981-4273-95-4

[Tangirala 2015] Tangirala, A.K. Principles of System Identification. Boca Raton : CRC Press, 2015. ISBN 978-1-4398-9602-0

[Wang 2017] Wang, D. et al. Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction, December 2017, Vol.113, pp. 1345-1358. ISSN 0960-1481

[Yang 2016] Yang, M. et al. Discussion on the Choice of Decomposition Level for Wavelet Based Hydrological Time Series Modeling. Water, May 2016, Vol.197, pp.1-11, ISSN 2073-4441

[Yu 2017] Yu, C. et al. An improved Wavelet Transform using Singular Spectrum Analysis for wind speed forecasting based on Elman Neural Network. Energy Conversion and Management, Vol.148, pp.895-904, ISSN 0196-8904

[Zainuddin 2011] Zainuddin, Z., Pauline, O. Modified wavelet neural network in function approximation and its application in prediction of time-series pollution data. Applied Soft Computing, December 2011, Vol.11, pp.4866-4874, ISSN 1568-4946

[Zhang 2017] Zhang, H. et al. Forecasting of PM_{10} time series using wavelet analysis and wavelet-ARMA model in Taiyuan, China. Journal of the Air & Waste Management Association, February 2017, Vol.67, pp.776-788, ISSN 2162-2906

[Zhao 2012] Zhao, H., Magoules, F. A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, August 2012, Vol.16, pp.3586-3592, ISSN 1364-0321

CONTACT:

doc. Ing. Alexander Hosovsky, PhD.
Department of Industrial Engineering and Informatics
Faculty of Manufacturing Technologies with seat in Presov
Technical university of Kosice
Bayerova 1, 08001 Prešov, Slovakia
alexander.hosovsky@tuke.sk