Article citation info:
VALIS D, FORBELSKÁ M, VINTR Z. Forecasting study of mains reliability based on sparse field data and perspective state space models. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (2): 179–191, http://dx.doi.org/10.17531/ein.2020.2.1.

David VALIS
Marie FORBELSKÁ
Zdeněk VINTR

FORECASTING STUDY OF MAINS RELIABILITY BASED ON SPARSE FIELD DATA AND PERSPECTIVE STATE SPACE MODELS

PROGNOZOWANIE NIEZAWODNOŚCI ELEMENTÓW SIECI WODOCIĄGOWEJ NA PODSTAWIE RZADKICH DANYCH TERENOWYCH I MODELI PRZESTRZENI STANÓW

The elements of critical infrastructure have to meet demanding dependability, safety and security requirements. The article deals with the prognosis of water mains reliability while using sparse irregular filed data. The data are sparse because the only thing we know is the number of mains failures during a given month. Since it is possible to transform the data into a typical reliability measure (rate of failure occurrence – ROCOF), we can examine the course of this measure development in time. In order to model and predict the ROCOF development, we suggest novel single and multiple error state space models. The results can be used for i) optimizing mains operation and maintenance, ii) estimating life cycle cost, and iii) planning crisis management.

***Keywords**: mains, critical infrastructure, reliability prognosis, sparse data, state space models.*

Elementy infrastruktury krytycznej muszą spełniać wysokie wymagania w zakresie niezawodności, bezpieczeństwa i ochrony. Artykuł dotyczy prognozowania niezawodności sieci wodociągowej przy wykorzystaniu nieregularnie rejestrowanych rzadkich danych. Wykorzystane w pracy dane są rzadkie, ponieważ dostarczają jedynie informacji na temat liczby uszkodzeń wodociągu w danym miesiącu. Przekształcenie tych danych w typową miarę niezawodności (wskaźnik występowania uszkodzeń – ROCOF), pozwala zbadać przebieg rozwoju tej miary w czasie. Rozwój ROCOF można modelować i przewidywać za pomocą zaproponowanych w pracy innowacyjnych modeli przestrzeni stanów uwzględniających pojedynczy błąd lub wiele błędów. Uzyskane wyniki można wykorzystać do i) optymalizacji pracy i eksploatacji sieci wodociągowej, ii) szacowania kosztów cyklu życia, oraz iii) planowania zarządzania kryzysowego.

***Słowa kluczowe**: sieć wodociągowa, infrastruktura krytyczna, prognoza niezawodności, rzadkie dane, modele przestrzeni stanów.*

1. Introduction

Water mains are an important part of the country critical infrastructure. The level of their dependability, safety and security is very much required and it must be high. Because of different structural designs, materials used, running failures and repairs, the technical condition of mains varies. Moreover, the system is subject to extreme stress owing to weather conditions. Monitoring the condition of this device is therefore very difficult. Both direct and indirect diagnostics is very problematic and does not provide satisfactory results. The access to the field data of this device is also difficult, and is not that common. Despite all these limitations mentioned, we are able to examine and predict the water mains reliability. For this purpose we use the sparse data set, and in order to model the data, we apply novel single and multiple error state space models.

1.1. State of the art

This part is devoted to the literature research focusing mainly on technical applications containing i) irregular sparse data, ii) single and multiple error state space models, iii) ways of predicting reliability.

Current publications may contain the results which deal with irregular sparse data. As examples, we introduce a few selected works which we found interesting, for different reasons explained later. Antholzer et al [2] propose a solution for accurate image reconstruction in the form of algorithms. They investigate this issue for the sparse data problem in photoacoustic tomography (PAT). They develop a direct and highly efficient reconstruction algorithm based on deep learning. In their approach, image reconstruction is performed with a deep convolutional neural network (CNN), whose weights are adjusted prior to the actual image reconstruction based on a set of training data. The proposed reconstruction approach can be interpreted as a network that uses the PAT filtered back projection algorithm for the

first layer, followed by the U-net architecture for the remaining layers. This work is inspirational but data set analyzed is not very large. Guo et al [17] bear in mind that fault diagnosis bearings in wind turbine and the drivetrain is very important to reduce the maintenance cost of the wind turbine and improve economic efficiency. However, the traditional diagnosis methods have difficulty in extracting the impulsive components from the vibration signal of the wind turbine because of heavy background noise and harmonic interference. In their paper, they propose a novel method based on data-driven multiscale dictionary construction. Firstly, they achieve the useful atom through training the K-means singular value decomposition (K-SVD) model with a standard signal. Secondly, they deform the chosen atom into different shapes and construct the final dictionary. Thirdly, the constructed dictionary is used to sparsely represent the vibration signal, and orthogonal matching pursuit (OMP) is performed to extract the impulsive component. The proposed method is robust to harmonic interference and heavy background noise. Moreover, the effectiveness of the proposed method is validated by numerical simulation and two experimental cases including the bearing fault of the wind turbine generator in the field test. This approach is interesting but the analyzed diagnostic signal is vibration, which includes some noise and is deformed by imperfection in terms of signal transfer. Wang and Lu [39] understand that incomplete modal data is mostly measured for experimental and engineering structures. However, its application into structural damage identification suffers from the drawback that the amount of the incomplete modal data is often insufficient, rendering the identification very sensitive to the measurement noise. Aiming to overcome this drawback, this paper proposes a new damage identification approach that combines the incomplete modal data with the sparse regularization. The realization of the proposed damage identification approach is mainly threefold: (a) The first is the establishment of a new goal function which is decoupled with respect to the damage parameters. To this end, the decomposition of the stiffness matrix so that the damage parameters are contained in a diagonal matrix must be introduced. (b) The second is the application of the alternating minimization approach to get the solution of the new goal function. (c) The third is the development of a novel and simple threshold setting method to properly determine the sparse regularization parameter. The feature of the proposed damage identification approach lies in that the sensitivity analysis is not involved and the exact orders of the modal data are not demanded. Numerical and experimental examples are conducted to verify the proposed damage identification approach. This outcome is also interesting, however we miss practical implementation of the proposed method. Experimental data are even more sparse. Sousa and Wang [35] present an application on bridges monitoring system. It uses on-structure sensors that are able to acquire signals sensitive to traffic load events, which can be used as an indirect indicator of the load magnitude. In this paper, sparse representation algorithms have been innovatively applied to the bridge weight in motion monitoring system data compression. A comparative study is performed based on measurements collected from a real bridge, by exploring different methods including Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and two dictionary learning methods, i.e. Compressive Sensing (CS) and K-means Singular Value Decomposition (K-SVD). This outcome is interesting but again, vibration assessment causes some noise in the final system state determination. Inspirational examples on specific diagnostic data forms analysis can be found in [9, 13, 14, 25, 28, 32].

In the next part of our literature review we focus on applying state-space time series models related to technical systems and reliability. The state-space models have been used by numerous applications based on the data analysis. In the text below there are examples of the outputs we found particularly inspiring. Dos Santos et al [10] propose new reliability models whose likelihood consists of decomposition of data information in stages or times, thus leading to latent state parameters. Alternative versions of some well-known models such as piecewise exponential, proportional hazards, and software reliability models are shown to be included in our unifying framework. In general, latent parameters of many reliability models are high dimensional, and their inference requires approximating methods such as Markov chain Monte Carlo (MCMC) or Laplace. Latent states in their models are related across stages through a non-Gaussian state-space framework. This feature makes the models mathematically tractable and allows for the exact computation of the marginal likelihood function, despite the non-Gaussianity of the state. The proposed non-Gaussian evolution models circumvent the need for approximations, which are required in similar likelihood-based approaches. This model however, has limited practical applicability. Li et al [29] present method of reliability prediction based on state space model. Firstly, signals about machine working conditions are collected based on-line monitoring technology. Secondly, wavelet packet energy parameters are determined based on the monitored signals. Frequency band energy is regarded as characteristic parameter. Then, the degradation characteristics of signal to noise ratio is improved by moving average filtering processing. In the end, SSM is established to predict degradation characteristics of probability density distribution, and the degree of reliability is determined. Milling cutter is used to demonstrate the rationality and effectiveness of this method. Simpler version of state-space model is applied here. Distefano et al [8] present outcomes of their article where the main goal is to demonstrate how state-space based techniques can satisfy demands on reliability and availability assessment. For this purpose some examples of specific dynamic reliability behaviors, such as common cause failure and load sharing, are considered applying state-space based techniques to study the corresponding reliability models. Different repair policies in availability contexts are also explored. Both Markovian and non-Markovian models are studied via phase type expansion and renewal theory in order to adequately represent and evaluate the considered dynamic reliability aspects in case of generally distributed lifetimes and times to repair. Although there are a great number of state-space models applications used not only in a technical area, it is difficult to find the specific application working with a single or multiple error state-space model. The authors offering interesting results in [11, 19].

Last but not least, we focus on the results dealing with reliability forecasting/prognosis. Gobbato et al [15] present approach based on fatigue assessment of structural components. The first part of the paper provides an overview and extension of a comprehensive reliability-based fatigue damage prognosis methodology – previously developed by the authors – for recursively predicting and updating the remaining fatigue life of critical structural components and/or subcomponents in aerospace structures. In the second part of the paper, a set of experimental fatigue test data, available in the literature, is used to provide a numerical verification and an experimental validation of the proposed framework at the reliability component level (i.e., single damage mechanism evolving at a single damage location). Gobbato et al [16] present the theoretical basis of a novel and comprehensive probabilistic methodology for predicting the remaining service life of adhesively bonded joints within the structural components of composite aircraft, with emphasis on a composite wing structure. Non-destructive evaluation techniques and recursive Bayesian inference are used to (i) assess the current state of damage of the system and (ii) update the joint probability distribution function (PDF) of the damage extents at various locations. A probabilistic model for future aerodynamic loads and a damage evolution model for the adhesive are then used to stochastically propagate damage through the joints and predict the joint PDF of the damage extents at future times. This information is subsequently used to probabilistically assess the reduced (due to damage) global aeroelastic performance of the wing by computing the PDFs of its flutter velocity and the velocities associated with the limit cycle oscillations of interest. Both of these methods are interest-

Table 1. Extracted sample of field data example – number of failures F_t during a respective month and ROCOF

Year	2000		2001		2002		2003	
Month	F_t	ROCOF	F_t	ROCOF	F_t	ROCOF	F_t	ROCOF
January	6	0.19354839	10	0.32258065	6	0.19354839	13	0.41935484
February	12	0.4137931	11	0.39285714	11	0.39285714	7	0.25
March	27	0.87096774	14	0.4516129	14	0.4516129	8	0.25806452
April	23	0.76666667	10	0.33333333	4	0.13333333	13	0.43333333
May	18	0.58064516	12	0.38709677	4	0.12903226	4	0.12903226
June	8	0.26666667	9	0.3	3	0.1	2	0.06666667
July	12	0.38709677	16	0.51612903	7	0.22580645	4	0.12903226
August	8	0.25806452	12	0.38709677	10	0.32258065	8	0.25806452
September	12	0.4	10	0.33333333	7	0.23333333	11	0.36666667
October	17	0.5483871	10	0.32258065	6	0.19354839	12	0.38709677
November	24	0.8	5	0.16666667	6	0.2	7	0.23333333
December	17	0.5483871	7	0.22580645	14	0.4516129	12	0.38709677
Total in year	184		126		92		101	

ing however small testing data set for approach validity confirmation has been presented. Dindarloo [7] presents alternatives to traditional reliability assessment approaches. Both the autoregressive integrated moving average (ARIMA or the Box-Jenkins technique) and artificial neural networks (ANNs) are viable alternatives to the traditional reliability analysis methods (e.g., Weibull analysis, Poisson processes, non-homogeneous Poisson processes, and Markov methods). Time series analysis of the times between failures via ARIMA or ANNs does not have the limitations of the traditional methods such as requirements/assumptions of a priori postulation and/or statistically independent and identically distributed observations for TBFs. The reliability of an load-haul-dump unit was investigated by analysis of time between failures. Seasonal autoregressive integrated moving average (SARIMA) was employed for both modeling and forecasting the failures. The results were compared with a genetic algorithm-based (ANNs) model. In this approach we can see interesting practical application. Kontrec et al [26] propose an approach that supports decision making process in planning and controlling of spare parts in aircraft maintenance systems. Reliability characteristics of aircraft consumable parts were analyzed in order to substantiate this approach. Moreover, the proposed reliability model was used to evaluate characteristics of subassemblies and/or assemblies these parts belong to. Finally, an innovative approach for determining the total amount of parts required in inventory and underage costs, based on observing the total unit time as a stochastic process, is presented herein. This application of stochastic process is also partially inspirational. Other encouraging results which deal with assessing systems reliability and include real field data but not only on water mains, are those introduced in [27,30,36-38].

1.2. Motivation

The main motivation of this article is to show that despite having only restricted information in the data, it is still possible to carry out certain estimates and prognoses of reliability measures. What is more, the data are very austere, and from the mathematics point of view, also very sparse. The only thing available is the number of failures which occurred during single months, the data information value is therefore very low. The assumptions about the future system behaviour are then very uncertain because of numerous influences, e.g. seasonal influences, water mains material, etc. That is the reason why we would like to introduce a suitable mathematical approach based on a few newly proposed single and multiple error state-space models. Using these models, we have the ambition not only to estimate the course of ROCOF, but also predict the system behaviour in the future. This information can significantly contribute to i) the optimization of this part of a critical infrastructure, ii) the planning of water mains maintenance, iii) the support of crisis management and emergency planning, iv) the rationalization of life cycle costs (LCC).

2. Field data analyzed

The analysed data are dichotomous and quantitative. They are the real field data collected in a broad area from the operators of the mains distribution system. In the area covered by the monitoring there are more than 5 million inhabitants supplied by water. The data cover the period longer than 17 years. Table 1 shows the example of such a data segment. We introduce the example of a complete form of the recorded data when having only the number of failures during single months. Although it would be possible to work with quarters which could filter out some data mistakes, we prefer to work with the number of failures during a month. The work with this variable would be possible but not entirely favourable. This value is therefore always transferred into the rate of failures occurrence (ROCOF), since it is necessary to consider and filter out the different number of days during a month. However, we presume that it will be possible to trace certain seasonal influences affecting the data.

The total course of ROCOF during a observed time elapse is shown in Fig. 1. The course is accompanied by trend smoothing non-parametric curves: lowess and cubic spline.

Next, in Fig. 2. we applied box plots to show a preliminary analysis of a time-series based on robust estimates of yearly and monthly ROCOF levels (thick lines in the boxes illustrate a median value) and on robust estimates of ROCOF variability during single years or months (see the box containing 75 % of the values of a given period). The left graph (panel) shows that the failure frequency decreases over

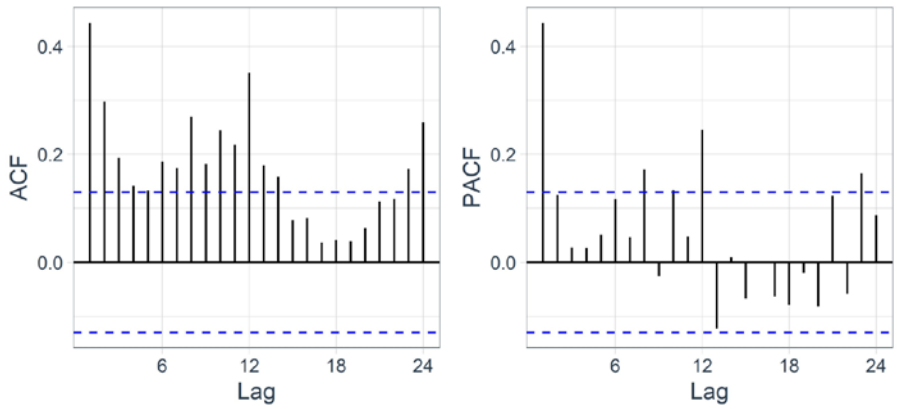

Fig. 1. ROCOF – monthly of observed series: accompanied with lowess trend line and its 95% confidence intervals (left) and with cubic spline and its 95% confidence intervals (right)

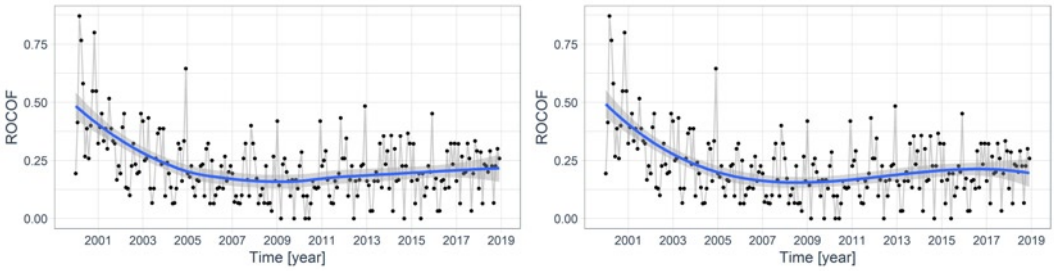

Fig. 2. Visualisation of robust estimates of ROCOF levels and of ROCOF variabilities for each year (left panel) and months (right panel)

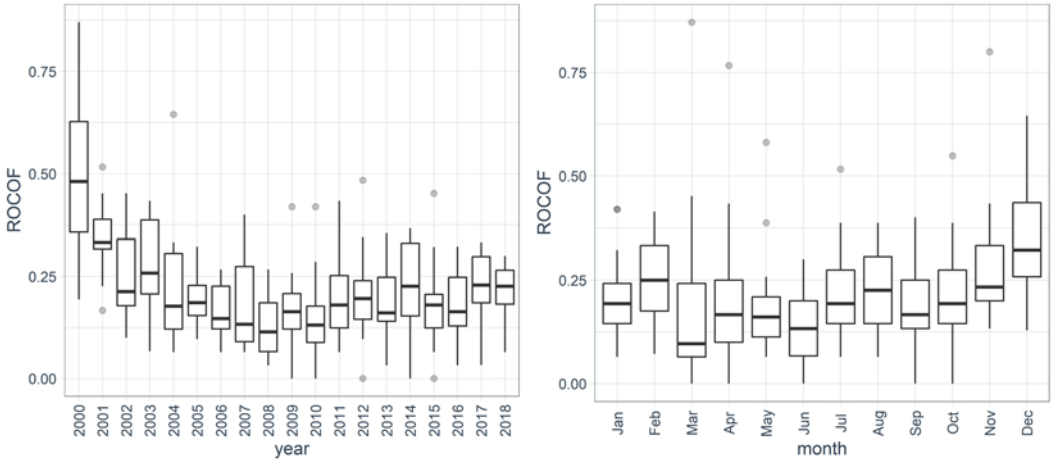

Fig. 3. Course of autocorrelation function during respective quarters (left) and partial autocorrelation function (right)

time – during years, which indicates that the system condition 'gets improved', perhaps due to gradual modernisation of network lines. We can also see that the ROCOF variability fluctuates during single years. The right graph (panel) shows the distribution of ROCOF in respective months and its variability. That is very useful since we work with and study the seasonality.

Since we have the idea of applying the following mathematical tools, we would like to verify the dependence or independence in the data – an autocorrelation function (ACF) and a partial correlation (PACF). Autocorrelation and partial autocorrelation plots are very often used in time series analysis and forecasting. These are plots that graphically summarize the strength of a relationship with an observation in a time series with observations at prior time steps, called lags. The blue dashed lines represent an approximate confidence interval for what is produced by white noise, by default a 95% interval. The ACF identifies the obvious seasonal variation (with high positive autocorrelations at lags 12, 24, …) and shows the slow decay typical for a non-stationary series. The results shown in Fig. 3 demonstrate that the data are independent.

No standard state-space models fit the data distributed in this manner. Therefore, we suggest special forms which capture, describe and fit the behaviour in the data a lot better. They are also able to depict the course of ROCOF in a better and more accurate way.

3. Modelling methodology

When modelling the system behaviour from the reliability point of view, we will follow the approach illustrated in Fig. 4. The first four steps have already been described above, and the remaining steps will be described in view of the applied theory.

Using the recorded data and the calculated ROCOF, we try to find the proper state-space model.

3.1. General theory of state space models

At first, we will describe briefly and generally the basic state-space models theory, later we will give a more detailed description of our newly proposed approaches and modifications of specially designed state-space models to be used for the examined specific technical case.

The state space representation of a linear time series model is given by:

$$y_t = w_t' x_{t-1} + \varepsilon_t \qquad (1)$$

$$x_t = F_t x_{t-1} + \eta_t \qquad (2)$$

for $t = 1, \ldots, n$, where y_t is the observed time series value. w_t (with w_t' its transpose) is assumed to be known vector, F_t a known matrix and x_t is the (possibly unobservable) state vector. Furthermore, ε_t are serially uncorrelated disturbances assumed to have zero mean and variance σ_ε^2, while η_t is a vector of serially uncorrelated disturbances with zero mean and covariance matrix V_t. Usually, ε_t and η_t are assumed to be (multivariate) normally distributed and uncorrelated with each other at all time periods (i.e. $E(\varepsilon_t \eta_s) = 0$ for all $t, s = 1, \ldots, n$). The model is thus characterized by w_t, F_t, σ_ε^2 and V_t. Equation (1) is known as the observation or measurement equation and (2) as the transition or state equation. If w_t, F_t and V_t do not change over time, the model is said to be time-invariant. Next we assume, the underlying state space models are time-invariant.

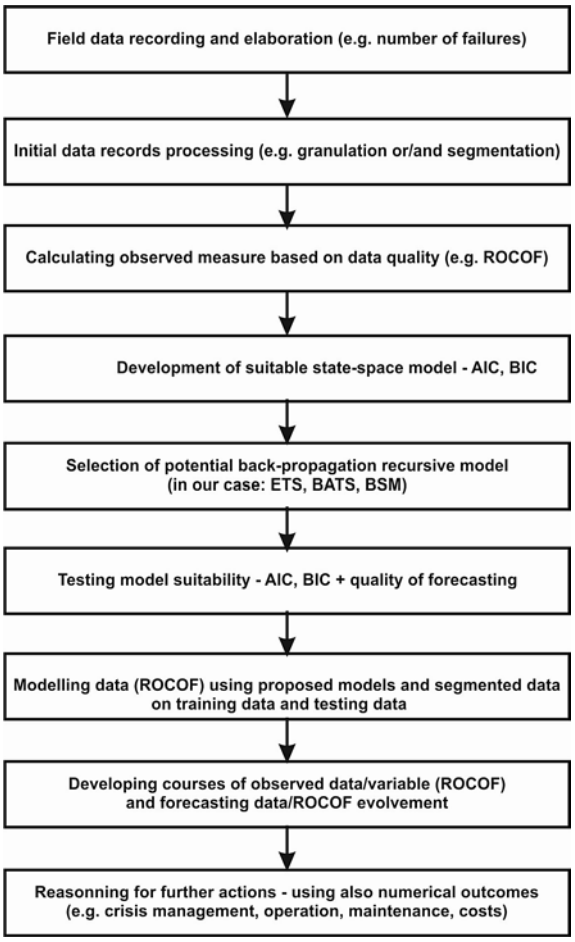

Fig. 4. Methodology of elaboration the recorded data on mains failure

One popular special case of state space models is the class of structural models. For these models, the vector x_t denotes a vector of states corresponding to the trend, cycle and seasonal components (see [18], chapter 2).

A special form of the state space model is the innovations model by Anderson and Moore (1979) [1]:

$$y_t = w_t' x_{t-1} + \varepsilon_t \tag{3}$$

$$x_t = F_t x_{t-1} + g\eta_t \tag{4}$$

where g is a fixed vector. The difference with the general state space model (1)-(2) is that the disturbance terms in the observation and transition equations are now perfectly correlated, implying that there is effectively only a single source of randomness. Since there is only one disturbance term in the model, it is referred to as a single source of error (SSOE) model. Models with more than one disturbance term are known as multiple sources of error (MSOE) models. This model has many common models as special cases, such as multiple regression, exponential smoothing, and ARIMA models. Any ARIMA model can be converted to this form. This model was first related to exponential smoothing in [33].

Since state space models contain both an unobservable state vector and unknown parameters, estimation for state space models has two aspects: i) estimating the unknown parameters, ii) estimating the unobservable state variables.

The first can be done by, for example, maximum likelihood estimation. The second is done by the Kalman filter (see e.g., [18]). The Kalman filter is a recursive procedure for computing the optimal estimator of the state vector at time t, based on the information available at time t (i.e. the observations up to and including y_t). Every time period, when new observations become available, this information is used to update the estimates. For state space models, optimal predictions of future observations and optimal estimates of unobserved components can be made using the Kalman filter. If the model is Gaussian (ε_t,η_t and the initial state are normally distributed), the Kalman filter is optimal in the sense that it yields minimum mean square error estimators (MMSE's).

For both the time-invariant MSOE and SSOE model, the Kalman filter converges to a steady state under some conditions.

In the next part we give a more detailed description of the particular proposed models used for examining our specific technical case.

3.1.1. Single error state space models for exponential smoothing

Exponential smoothing state space methods constitute a broad family of approaches to univariate time series forecasting that have been around for many decades and only in the twenty-first century placed into a systematic framework. The definitive book on the subject is [20]. In general, innovation ETS models are defined according to three model structure parameters: (E) error type, (T) trend type, and (S) seasonality type [20,24]. Each of the parameters can be an N (none), A (additive), or M (multiplicative) state. The trend component also can be d (damped) type.

ETS(A,Ad,A) (seasonal exponential smoothing with damped trends)

Consider the single error state space model given by:

observed series: $y_t = l_t + s_{t-m} + \varepsilon_t \quad \varepsilon_t \sim {}^{\circ}NID\left(0,\sigma_\varepsilon^2\right)$

latent level: $l_t = l_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t \quad 0 < \phi < 1$ – damping parameter

latent trend/drift: $b_t = \phi b_{t-1} + \beta\varepsilon_t$

latent seasonal: $s_t = s_{t-m} + \gamma\varepsilon_t \quad 0 \le \alpha,\beta,\gamma \le 1$ - smoothing parameters

A deterministic representation of the seasonal components can be obtained by setting the smoothing parameters equal to zero.

BATS model (exponential smoothing state space model with Box-Cox transformation, ARMA errors, trend and seasonal components)

De Livera et al. [6] propose modifications to the linear innovation models in order to include a wide variety of seasonal patterns and solve the problem of correlated errors. To avoid falling into nonlinearity problems, these authors restricted the models to those homoskedastic and the Box-Cox transformation [3] are used when there is some type of specific nonlinearity. The model including the transformation of Box and Cox, ARMA errors (see [4,5,40]) and seasonal patterns can be expressed as follows:

Box-Cox power transformation: $y_t^{(\omega)} = \begin{cases} \dfrac{y_t^{(\omega)} - 1}{\omega} & \omega \neq 0 \\ \ln y_t & \omega = 0 \end{cases}$

transformed observation: $y_t^{(\omega)} = l_t + \phi b_{t-1} + \sum_{i=1}^{M} s_{t-m}^{(i)} + d_t$

latent level: $l_t = l_{t-1} + \phi b_{t-1} + \alpha d_t$

latent trend/drift: $b_t = (1-\phi)b + \phi b_{t-1} + \beta d_t$

latent seasonal: $s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t$

$ARMA(p, q)$ process: $d_t = \sum_{i=1}^{p} \varphi_i d_{t-1} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-1} + \varepsilon_t \; \varepsilon_t \sim NID\left(0, \sigma_\varepsilon^2\right)$

where:

ω	Box-Cox transformation parameter
ϕ	damping parameter of trend (see [12, 34]),
$m_1, \cdots, m_M$	seasonal periods,
b	long-run trend parameter,
b_t	short-run trend parameter in time t,
$\alpha, \beta, \gamma_1^{(i)}, \gamma_2^{(i)}$	smoothing parameters.

These models are called BATS with arguments (ω,ϕ, p, q, m_1, …, m_M). A deterministic representation of the seasonal components can be obtained by setting the smoothing parameters equal to zero.

3.1.2. Multiple error state space model

Among state space models with multiple errors, the following model was chosen as suitable

observed series: $y_t = l_t + s_{1,t} + \varepsilon_t \qquad \varepsilon_t \sim NID\left(0, \sigma_\varepsilon^2\right)$

latent level: $l_{t-1} + b_{t-1} + \eta_t \qquad \eta_t \sim NID\left(0, \sigma_\eta^2\right)$

latent trend/drift: $b_t = b_{t-1} + \xi_t \qquad \xi_t \sim NID\left(0, \sigma_\xi^2\right)$

latent seasonal: $s_{1,t} = -\sum_{j=1}^{s-1} s_{j,t-1} + \nu_t \qquad \nu_t \sim NID\left(0, \sigma_\nu^2\right)$

$$s_{2,t} = s_{1,t-1}$$
$$\vdots$$
$$s_{m-1,t} = s_{m-2,t-1}$$

This model is called BSM (Basic Structural Model). If we omit the drift component b_t in the BSM model, we mark this model as Level-BSM.

3.1.3. Quality of forecasting

Different criteria such as forecast error measurements, the speed of calculation, interpretability and others have been used to assess the quality of forecasting. Forecast error measures or forecast accuracy are the most important in solving practical problems. Typically, the common used forecast error measurements are applied for estimating the quality of forecasting methods and for choosing the best forecasting mechanism in case of multiple objects.

Training and test sets

It is important to evaluate forecast accuracy using genuine forecasts. Consequently, the size of the residuals is not a reliable indication of how large true forecast errors are likely to be. The accuracy of forecasts can only be determined by considering how well a model performs on new data that were not used when fitting the model. When choosing models, it is common practice to separate the available data $Y_n = \{y_1, \ldots, y_n\}$ into two portions, training and test data,

training data: $Y_T = \{y_1, \ldots, y_T\}$, test data: $Y_{test} = \{y_{T+1}, y_{T+2}, \ldots, y_n\}$

where the training data is used to estimate any parameters of a forecasting method and the test data is used to evaluate its accuracy. Because the test data is not used in determining the forecasts, it should provide a reliable indication of how well the model is likely to forecast on new data. This data sub-division into regions can be seen in Fig. 5.

Forecast Accuracy

The forecast accuracy can be evaluated on the test set using residual diagnostics and forecast accuracy measures. A forecast "error" is the difference between an observed value and its forecast. Here "error" does not mean a mistake, it means the unpredictable part of an observation. It can be written as:

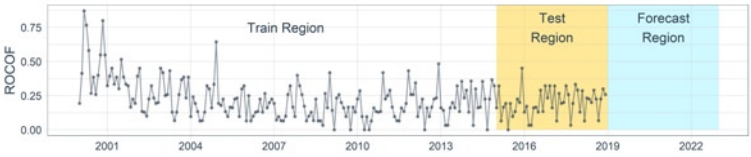

Fig. 5. Example of data sub-division into training, testing and forecasting region

$$e_t = y_t - f_t^{(m)}$$

where y_t is the measured value at time t, $f_t^{(m)}$ is predicted value at time t, obtained from the use of the forecast model m. Note that forecast errors are different from residuals in two ways. First, residuals are calculated on the training set while forecast errors are calculated on the test set. Second, residuals are based on one-step forecasts while forecast errors can involve multi-step forecasts. We can measure forecast accuracy by summarising the forecast errors in different ways – see bellow please.

Scale-dependent errors

The forecast errors are on the same scale as the data. Accuracy measures that are based only on et are therefore scale-dependent and cannot be used to make comparisons between series that involve different units. The two most commonly used scale-dependent measures are based on the absolute errors or squared errors:

Mean absolute error: $MAE = mean(|e_t|)$

Root mean squared error: $RMSE = \sqrt{mean(e_t^2)}$

When comparing forecast methods applied to a single time series, or to several time series with the same units, the MAE is popular as it is easy to both understand and compute. A forecast method that minimises the MAE will lead to forecasts of the median, while minimising the RMSE will lead to forecasts of the mean. Consequently, the RMSE is also widely used, despite being more difficult to interpret.

Percentage Errors

The percentage error is given by:

$$p_t = 100 \frac{e_t}{y_t}$$

Percentage errors have the advantage of being unit-free, and so are frequently used to compare forecast performances between data sets. The most commonly used measure is:

Mean percentage error: $MPE = mean(p_t)$

Mean absolute percentage error: $MAPE = mean(|p_t|)$

Measures based on percentage errors have the disadvantage of being infinite or undefined if $y_t = 0$ for any t in the period of interest, and having extreme values if any y_t is close to zero. Another problem with percentage errors that is often overlooked is that they assume the unit of measurement has a meaningful zero.

Scaled Errors

Scaled errors were proposed by Hyndman & Koehler (2006) [21] as an alternative to using percentage errors when comparing forecast accuracy across series with different units. They proposed scaling the errors based on the training MAE from a simple forecast method.

For a non-seasonal time series, a useful way to define a scaled error uses naïve forecasts:

$$q_j = \frac{e_j}{\frac{1}{T-1}\sum_{t=2}^{T}|y_t - y_{t-1}|}$$

Because the numerator and denominator both involve values on the scale of the original data, q_j is independent of the scale of the data. A scaled error is less than one if it arises from a better forecast than the average naive forecast computed on the training data. Conversely, it is greater than one if the forecast is worse than the average naive forecast computed on the training data. For seasonal time series, a scaled error can be defined using seasonal naive forecasts:

$$q_j = \frac{e_j}{\frac{1}{T-s}\sum_{t=s+1}^{T}|y_t - y_{t-s}|}$$

The mean absolute scaled error is simply

$$MASE = mean(|q_j|).$$

Both evaluation metrics and residuals diagnostics are used. The most common evaluation metrics for forecasting are *RMSE*, which you may have used on regression problems; *MAPE*, as it is scale-independent and represents the ratio of error to actual values as a percent; and *MASE*, which indicates how well the forecast performs compared to a naïve average forecast.

4. Results of ROCOF modelling

In this section we bring results for ROCOF modelling using the respective above proposed new state space models.

4.1. SSOE models – innovations state space models for exponential smoothing

In this paper, the innovation state space models that capture various forms of the exponential smoothing methodology is used first. The method described by Hyndman et al [22] uses a state space framework for the automatic selection of exponential smoothening techniques for forecasting. The framework makes an assessment of best fit – comparing Akaike's Information Criterions (see [23, 24]).

ETS models

Based on the AIC criterion, the most appropriate ETS model for training data Y_T is first found. This model is ETS(A,Ad,A):

$$y_t = [1\,0\,1]\begin{bmatrix} l_{t-1} \\ b_{t-1} \\ s_{t-m} \end{bmatrix} + \varepsilon_t \qquad \varepsilon_t \sim N(0, 0.116^2)$$

$$\begin{bmatrix} l_t \\ b_t \\ s_t \end{bmatrix} = \begin{bmatrix} 1 & 0.972 & 0 \\ 0 & 0.972 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} l_{t-1} \\ b_{t-1} \\ s_{t-m} \end{bmatrix} + \begin{bmatrix} 2.59E-02 \\ 2.91E-03 \\ 1.19E-04 \end{bmatrix}\varepsilon_t$$

Even on the basis of all the data Yn, another type of ETS model was not chosen as optimal.

ETS(A,Ad,A):

$$y_t = [1\,0\,1]\begin{bmatrix} l_{t-1} \\ b_{t-1} \\ s_{t-m} \end{bmatrix} + \varepsilon_t \qquad \varepsilon_t \sim N(0, 0.107^2)$$

$$\begin{bmatrix} l_t \\ b_t \\ s_t \end{bmatrix} = \begin{bmatrix} 1 & 0.964 & 0 \\ 0 & 0.964 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} l_{t-1} \\ b_{t-1} \\ s_{t-m} \end{bmatrix} + \begin{bmatrix} 1.55E-02 \\ 1.39E-03 \\ 1.06E-04 \end{bmatrix}\varepsilon_t$$

Components of this state space model for both training and all data are presented in Fig. 6.

BATS models

Again, using the AIC criterion, we find the most appropriate BATS model for training data, which is model BATS(1, {0,0}, -, {12}) (see [23, 24]). Since $\omega = 1$, the observed data is not transformed. In addition, dumping parameter ϕ is zero, so the model becomes simpler as the drift component falls out. Moreover, the orders of both *AR* and *MA* processes are zero, so d_t is white noise ε_t. Components of this state space model for both training and all data are presented in Fig. 7.

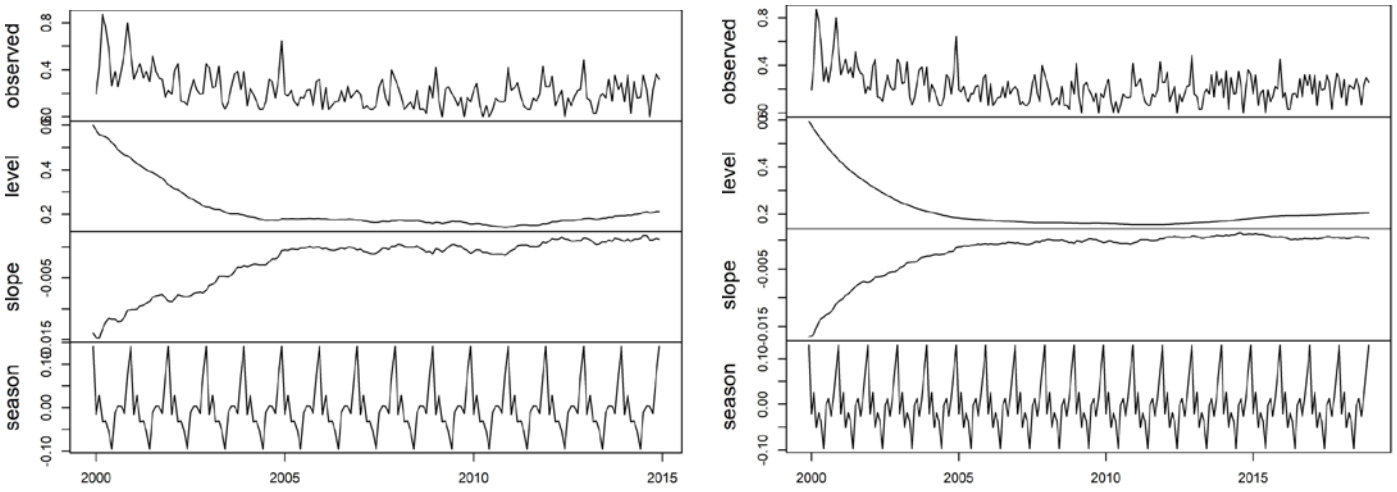

Fig. 6. Components of ETS method (left panel: training data, right panel: all data)

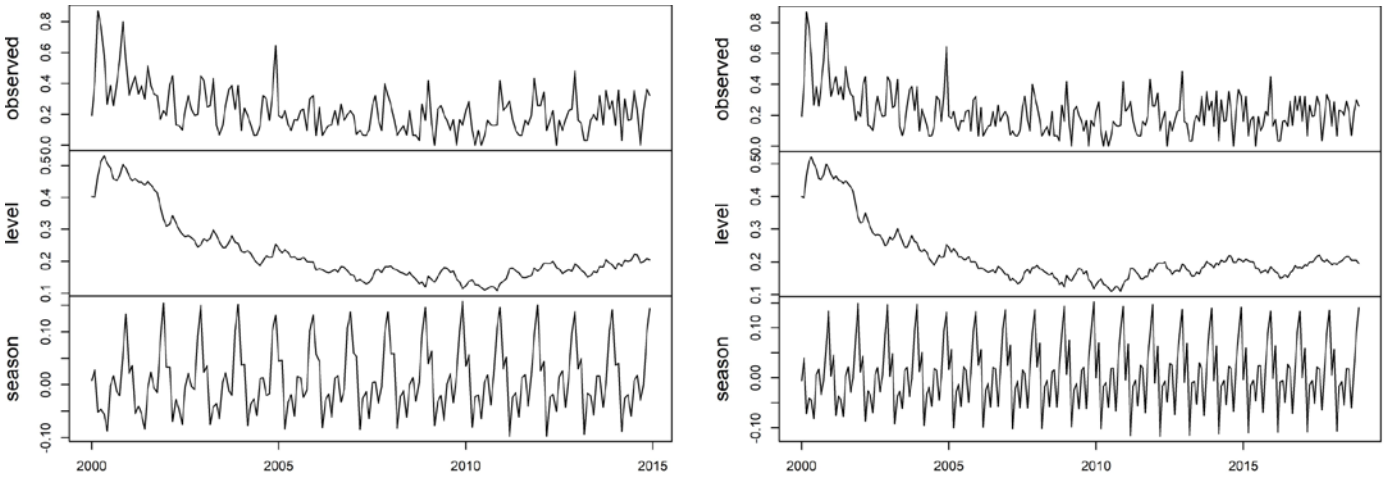

Fig. 7. Components of BATS method (left panel: training data, right panel: all data)

4.2. MSOE models – innovations state space models for exponential smoothing

The MSOE models are based on the state space form, the Kalman filter, and the associated smoother. The likelihood is constructed from the Kalman filter in terms of the one-step-ahead prediction errors and maximized with respect to the hyperparameters by numerical optimization. The score vector of the parameters can be obtained via a smoothing algorithm which is associated with the Kalman filter. Once the hyperparameters have been estimated, the filter is used to produce one-step-ahead prediction residuals which enables us to compute diagnostic statistics for normality, serial correlation, and goodness of fit. The smoother is used to estimate unobserved components, such as trends and seasonals, and to compute diagnostic statistics for detecting outliers and structural breaks.

BSM models

Next, the BSM model was fitted to the training data with the following results:

$$y_t = \left[10 \mid 1\,0\,0\,0\,0\,0\,0\,0\,0\,0\right]\begin{bmatrix}\mu_{t-1}\\ s_{t-1}\end{bmatrix} + \varepsilon_t;\ \varepsilon_t \sim N\left(0,\sigma_\varepsilon^2\right)\ \sigma_\varepsilon = 1.07E-01$$

$$\begin{bmatrix}\mu_t\\ s_t\end{bmatrix} = \begin{bmatrix}F_\mu & 0\\ 0 & F_{seas}\end{bmatrix}\begin{bmatrix}\mu_{t-1}\\ s_{t-1}\end{bmatrix}\ \mu_t \sim N_2\left(0,\sum\mu\right)\ \sum\mu = \begin{bmatrix}\sigma_\eta^2 & 0\\ 0 & \sigma_\xi^2\end{bmatrix}$$

$$\sigma_\eta = 1.76E-05$$

$$\sigma_\xi = 5.43E-04$$

$$s_t \sim N_{11}\left(0,\sum s\right)\ \sum s = \begin{bmatrix}\sigma_\nu^2 & 0\\ 0 & 0\end{bmatrix}\ \sigma_\nu = 1.55E-02$$

where μt = (lt, bt), st = (s1,t, … , s11,t), $F_\mu = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$, $F_{seas} = \begin{bmatrix}-1 & -1 & \cdots & -1\\ 0 & & I_{10} & \end{bmatrix}$

and I10 is 10×10 matrix.

Next, the BSM model was fitted to the all of the available monthly data with the following results:

$$y_t = \left[10 \mid 1\,0\,0\,0\,0\,0\,0\,0\,0\,0\right]\begin{bmatrix}\mu_{t-1}\\ s_{t-1}\end{bmatrix} + \varepsilon_t;\ \varepsilon_t \sim N\left(0,\sigma_\varepsilon^2\right)\ \sigma_\varepsilon = 9.84E-02$$

$$\begin{bmatrix}\mu_t\\ s_t\end{bmatrix} = \begin{bmatrix}F_\mu & 0\\ 0 & F_{seas}\end{bmatrix}\begin{bmatrix}\mu_{t-1}\\ s_{t-1}\end{bmatrix}\ \mu_t \sim N_2\left(0,\sum\mu\right)\ \sum\mu = \begin{bmatrix}\sigma_\eta^2 & 0\\ 0 & \sigma_\xi^2\end{bmatrix}\ \sigma_\eta = 2.99E-03$$

$$\sigma_\xi = 4.60E-04$$

$$s_t \sim N_{11}\left(0,\sum s\right)\ \sum s = \begin{bmatrix}\sigma_\nu^2 & 0\\ 0 & 0\end{bmatrix}\ \sigma_\nu = 1.50E-02$$

Components of this state space model for both training and all data are presented in Fig. 8.

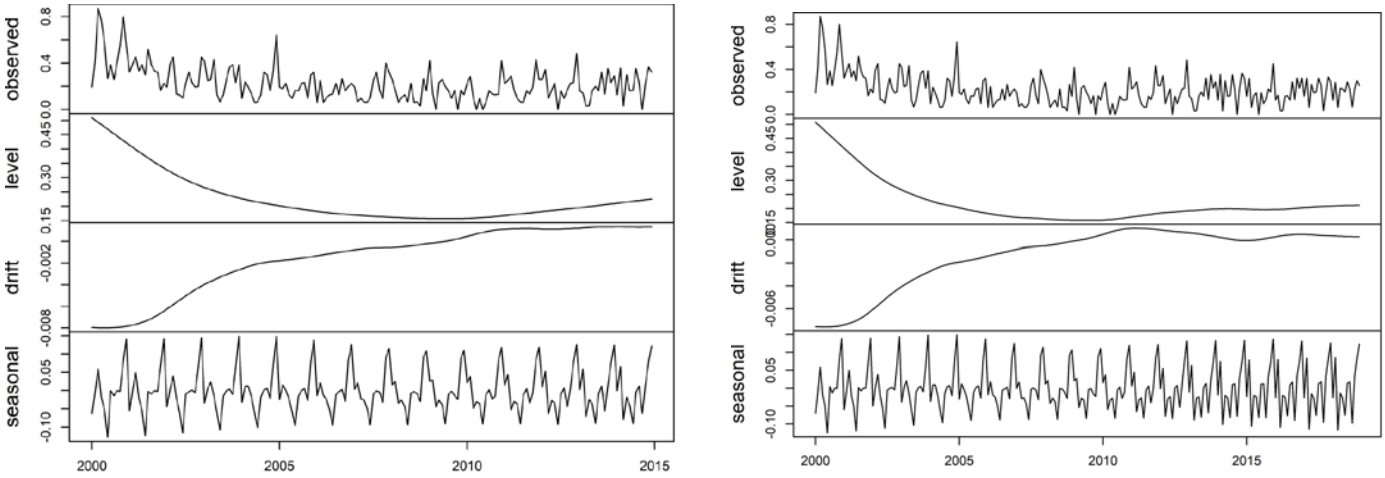

Fig. 8. Components of BSM method (left panel: training data, right panel: all data)

Level-BSM models

Next, the Level-BSM model was fitted to the training data with the following results:

$$y_t = [10|1\,0\,0\,0\,0\,0\,0\,0\,0\,0]\begin{bmatrix} l_{t-1} \\ s_{t-1} \end{bmatrix} + \varepsilon_t;\ \varepsilon_t \sim N(0,\sigma_\varepsilon^2)\ \ \sigma_\varepsilon = 1.04E-01$$

$$\begin{bmatrix} l_t \\ s_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & F_{seas} \end{bmatrix}\begin{bmatrix} l_{t-1} \\ s_{t-1} \end{bmatrix}\ \ l_t \sim N(0,\sigma_\eta^2)\ \ \sigma_\eta = 1.71E-02$$

$$s_t \sim N_{11}(0,\sum s)\ \ \sum s = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & 0 \end{bmatrix}\ \ \sigma_v = 1.61E-02$$

where $s_t = (s_{1,t}, \ldots, s_{11,t})$, $F_{seas} = \begin{bmatrix} -1 & -1 & \cdots & -1 \\ 0 & & I_{10} & \end{bmatrix}$ and I_{10} is 10×10 matrix.

Next, the Level-BSM model was fitted to the all of the available monthly data with the following results:

$$y_t = [10|1\,0\,0\,0\,0\,0\,0\,0\,0\,0]\begin{bmatrix} l_{t-1} \\ s_{t-1} \end{bmatrix} + \varepsilon_t;\ \ \varepsilon_t \sim N(0,\sigma_\varepsilon^2)\ \ \sigma_\varepsilon = 9.55E-02$$

$$\begin{bmatrix} l_t \\ s_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & F_{seas} \end{bmatrix}\begin{bmatrix} l_{t-1} \\ s_{t-1} \end{bmatrix}\ l_t \sim N(0,\sigma_\eta^2)\ \ \sigma_\eta = 1.56E-02$$

$$s_t \sim N_{11}(0,\sum s)\ \ \sum s = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & 0 \end{bmatrix}\ \ \sigma_v = 1.55E-02$$

Components of this state space model for both training and all data are presented in Fig. 9.

In all the graphs in Figures 6 – 9 we might observe very clearly specific development of the "level" value course and the affected ROCOF behaviour changes during the "seasonal" course.

5. Discussion to ROCOF modelling and performance measures

This and following part 5.1 are devoted to graphic and numerical results which help us to i) find the model which could fit a given type of the analyzed data best, ii) get an idea of the courses of single model types for both the training data and the testing data especially in the forecasting region, iii) numerically compare the results of single models for the courses – fitting, level, drift, forecast a 95% prediction intervals in the forecasting region for training and all data.

Based on the outcomes presented further there can be seen small but still existing difference in the applied models. Although one may say that divergence in these models is not significant it is not absolutely truth. Both of these models have their mathematical principles, therefore advantages and practical applicability especially in terms of these field data forms.

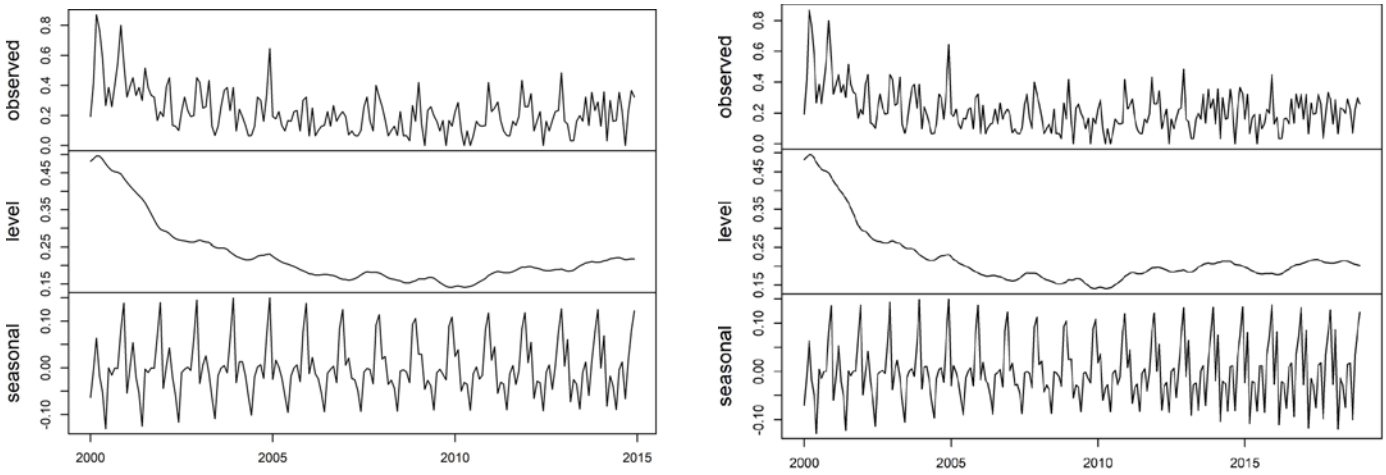

Fig. 9. Components of Level-BSM method (left panel: training data, right panel: all data)

Table 2. Forecast accuracy on test data with rating

Model	Type	RMSE	rRMSE	MAE	rMAE	MASE	rMASE
ets	SSOE	0.0886	3	0.0728	3	0.6480	3
bats	SSOE	0.0751	2	0.0637	2	0.5673	2
bsm	MSOE	0.0904	4	0.0742	4	0.6604	4
level-bsm	MSOE	0.0715	1	0.0595	1	0.5302	1

Table 3. Forecast accuracy on training data with rating

Model	Type	AIC	rAIC	BIC	rBIC	RMSE	rRMSE	MAE	rMAE	MASE	rMASE
ets	SSOE	178.3885	2	235.8617	2	0.1107	2	0.0842	2	0.7494	3
bats	SSOE	173.3949	1	179.7809	1	0.1110	3	0.0862	3	0.7433	2
bsm	MSOE	306.1965	3	318.9683	3	0.0980	1	0.0753	1	0.6706	1
level-bsm	MSOE	328.2960	4	337.8749	4	0.1223	4	0.0937	4	0.8342	4

Table 4. Forecast accuracy on all data with rating

Model	Type	AIC	rAIC	BIC	rBIC	RMSE	rRMSE	MAE	rMAE	MASE	rMASE
ets	SSOE	238.4869	1	300.2151	2	0.1032	2	0.0793	2	0.7613	3
bats	SSOE	243.7871	2	250.6458	1	0.1058	3	0.0814	3	0.6936	2
bsm	MSOE	492.1602	3	505.8776	3	0.0906	1	0.0689	1	0.6609	1
level-bsm	MSOE	514.8491	4	525.1372	4	0.1239	4	0.0956	4	0.9176	4

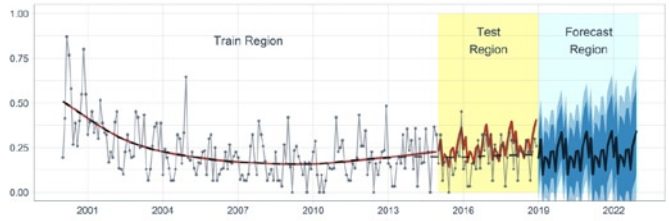

Fig. 10. Forecasting and testing – ETS model

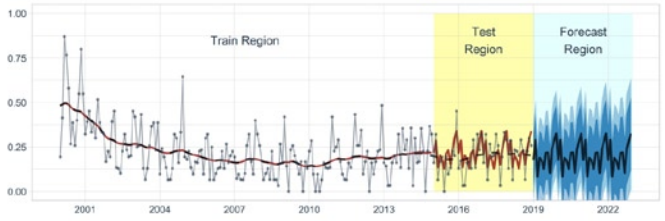

Fig. 11. Forecasting and testing – BATS model

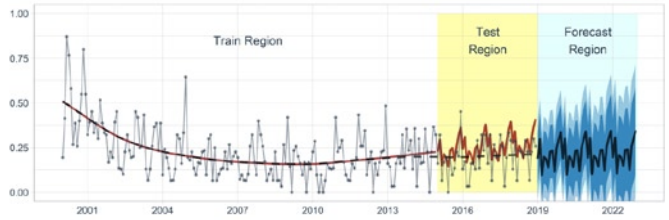

Fig. 12. Forecasting and testing – BSM model

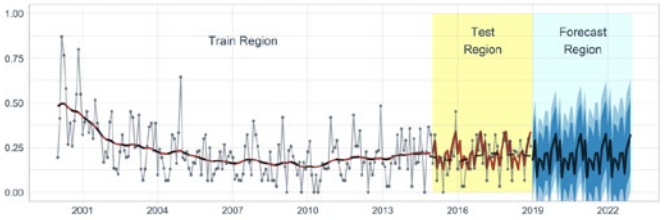

Fig. 13. Forecasting and testing – Level-BSM model

Table 5. Numerical values of ETS model components

Date	Nr. days	Nr. Failures	ROCOF	ETS Fit	ETS Level	ETS Drift	ETS Seasonal	ETS Forecast	ETS PI 95L	ETS PI 95U
1.1.2000	31	6	0.193548	0.570834	0.576737	-1.46E-02	-0.015703446			
1.2.2000	29	12	0.413793	0.590349	0.5579475	-1.47E-02	0.027815288			
1.3.2000	31	27	0.870968	0.511642	0.5529108	-1.33E-02	-0.031934835			
...										
1.10.2018	31	7	0.22580645	0.21918228	0.2049066	4.34E-04	0.014286635			
1.11.2018	30	9	0.3	0.27814934	0.205359	4.49E-04	0.072826575			
1.12.2018	31	8	0.25806452	0.33630317	0.2056702	3.24E-04	0.130503361			
1.1.2019	31	*~6*						0.184536	-0.02582	0.394893
1.2.2019	28	*~7*						0.232937	0.02258	0.443294
1.3.2019	31	*~5*						0.155278	-0.05508	0.365637
...										
1.10.2022	31	*~7*						0.2269827	0.014268325	0.4396971
1.11.2022	30	*~9*						0.2855796	0.072760766	0.4983983
1.12.2022	31	*~11*						0.3433112	0.130386422	0.5562359

Table 6. Numerical values of BSM model components

Date	Nr. days	Nr. Failures	ROCOF	BSM Fit	BSM Level	BSM Drift	BSM Seasonal	BSM Forecast	BSM PI 95L	BSM PI 95U
1.1.2000	31	6	0.193548	0.43750179	0.507319	-7.58E-03	-6.98E-02			
1.2.2000	29	12	0.413793	0.48954893	0.4999693	-7.58E-03	-1.04E-02			
1.3.2000	31	27	0.870968	0.55150313	0.4926839	-7.59E-03	5.88E-02			
...										
1.10.2018	31	7	0.22580645	0.244381	0.21077	2.61E-04	3.36E-02			
1.11.2018	30	9	0.3	0.293988	0.210966	2.59E-04	8.30E-02			
1.12.2018	31	8	0.25806452	0.333904	0.211154	2.59E-04	1.23E-01			
1.1.2019	31	~6						0.18671103	-0.036826694	0.4102488
1.2.2019	28	~8						0.2973334	0.073650738	0.5210161
1.3.2019	31	~3						0.09532622	-0.129205549	0.319858
...										
1.10.2022	31	~8						0.25667698	-0.109152479	0.6225064
1.11.2022	30	~9						0.30634686	-0.064033347	0.6767271
1.12.2022	31	~11						0.34633359	-0.027338535	0.7200057

5.1. Performance model measures

In this part we introduce numerical calculations of single measures which show the fitness of the proposed models. The forecast accuracy for the test data is put in Table 2, the forecast accuracy for the training data is put in Table 3, and the forecast accuracy for all the data is put in Table 4. The outcomes are always accompanied by rating of prioritization.

The modeling outputs along with prognoses are introduced in a graphical form first, see Fig. 10 – 13.

For each of the most suitable models of SSOE – ETS and MSOE – BSM groups we performed simulations within a forecast region in order to determine prediction intervals – PI (95% lower and upper (PI 95L and PI 95U). The graphical outcomes are put in Fig. 14 and 15, and the following numerical outcomes are put in Tables 5 and 6. The predicted numbers of water mains failures according to single models are put in *Italics*.

The introduced results show that graphical and numerical model forms are interesting, however, in their prognoses they complement each other significantly even when it comes to the predicted failure numbers.

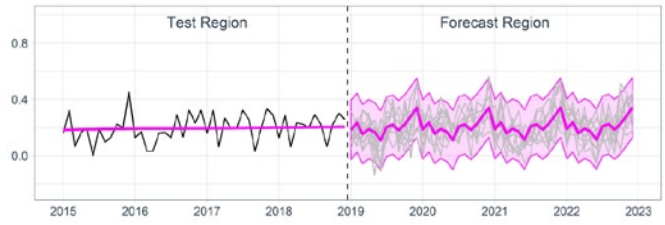

Fig. 14. Simulation inside the forecast region – ETS model – trend with 95% PI

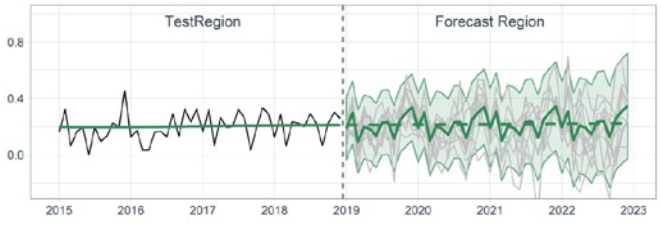

Fig. 15. Simulation inside the forecast region – BSM model – trend with 95% PI

6. Conclusions

In our article we introduce new and promising state-space models which seem to be very useful and suitable every time there is insufficient information on the system failures. The models SSOE and MSOE, or their representatives ETS and BSM, belong to the group of structural models which fit this type of analyzed field data very well, and also are suitable for forecasting failures and reliability measures development.

In our future work we would like to verify in practice, whether the achieved results agree and to what extent with our calculations. Luckily, as early as at the beginning of 2019, the first checks of our results conformed to our calculations significantly.

All the outcomes, both graphical and numerical, were acquired with the help of R Studio [31].

Acknowledgement

This article has been created with support of funds of partial development of UO-FVT-K202-MOBAUT.

References

1. Anderson B D O, Moore J B. Optimal filtering. Prentice Hall, New York: Englewood Cliffs 1979.
2. Antholzer S, Haltmeier M, Schwab J. Deep learning for photoacoustic tomography from sparse data. Inverse Problems in Science and Engineering 2019; 27(7): 987-1005, https://doi.org/10.1080/17415977.2018.1518444.
3. Box George E P, Cox D R. An analysis of transformations. Journal of the Royal Statistical Society 1964; 26(2): 211-252, https://doi.org/10.1111/j.2517-6161.1964.tb00553.x.
4. Box George E P, Jenkins G M. Time Series Analysis: Forecasting and Control. San Francisco: Holden Day 1976.

5. Brockwell P J, Davis R A. Time Series: Theory and Methods, Berlin: Springer 1991. https://doi.org/10.1007/978-1-4419-0320-4
6. De Livera A M, Hyndman R J, Snyder R D. Forecasting time series with complex seasonal patterns using exponential smoothing. Journal of the American Statistical Association 2011; 106(496): 1513-1527, https://doi.org/10.1198/jasa.2011.tm09771.
7. Dindarloo S. Reliability forecasting of a load-haul-dump machine: a comparative study of ARIMA and neural networks. Quality and Reliability Engineering International 2016; 32(4): 1545-1552, https://doi.org/10.1002/qre.1844.
8. Distefano S, Longo F, Trivedi K S. Investigating dynamic reliability and availability through state-space models. Computers & Mathematics with Applications 2012; 64(12): 3701-3716, https://doi.org/10.1016/j.camwa.2012.02.038.
9. Duan RX, Lin Y N, Zeng Y N, Fault diagnosis for complex systems based on reliability analysis and sensors data considering epistemic uncertainty. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 558-566, https://doi.org/10.17531/ein.2018.4.7.
10. dos Santos TR, Gamerman D, Franco G D. Reliability Analysis via Non-Gaussian State-Space Models. IEEE Transactions on Reliability 2017; 66(2): 309-318, https://doi.org/10.1109/TR.2017.2670142.
11. Faes L, Porta A, Javorka M, Nollo G. Efficient computation of multiscale entropy over short biomedical time series based on linear state-space models. Complexity 2017: 1768264, https://doi.org/10.1155/2017/1768264.
12. Gardner J E S, McKenzie E. 1985. Forecasting trends in time series. Management Science 1985; 31(10): 1237-1246, https://doi.org/10.1287/mnsc.31.10.1237.
13. Glowacz A. Fault diagnosis of single-phase induction motor based on acoustic signals. Mechanical Systems and Signal Processing 2019; 117: 65-80, https://doi.org/10.1016/j.ymssp.2018.07.044.
14. Glowacz A. Fault Detection of Electric Impact Drills and Coffee Grinders Using Acoustic Signals. Sensors 2019; 19(2): 269, https://doi.org/10.3390/s19020269.
15. Gobbato M, Kosmatka J B, Conte J P. A recursive Bayesian approach for fatigue damage prognosis: An experimental validation at the reliability component level. Mechanical Systems and Signal Processing 2014; 45(2): 448-467, https://doi.org/10.1016/j.ymssp.2013.10.014.
16. Gobbato M, Conte J P, Kosmatka J B, Farrar C R. A reliability-based framework for fatigue damage prognosis of composite aircraft structures. Probabilistic Engineering Mechanics 2012; 29: 176-188, https://doi.org/10.1016/j.probengmech.2011.11.004.
17. Guo, Y J, Zhao Z B, Sun R B, Chen X F. Data-driven multiscale sparse representation for bearing fault diagnosis in wind turbine. Wind Energy 2019; 22(4): 587-604, https://doi.org/10.1002/we.2309.
18. Harvey A C. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge: University Press 1989, https://doi.org/10.1017/CBO9781107049994.
19. Hu Y W, Liu S J, Lu H T, Zhang H C. Remaining useful life model and assessment of mechanical products: a brief review and a note on the state space model method. Chinese Journal of Mechanical Engineering 2019; 32(1):15, https://doi.org/10.1186/s10033-019-0317-y.
20. Hyndman R J, Koehler A B, Ord J K, Snyder R D. Forecasting with Exponential Smoothing: The State Space Approach. Berlin Germany: Springer 2008, https://doi.org/10.1007/978-3-540-71918-2.
21. Hyndman R J, Koehler A B. Another look at measures of forecast accuracy. International Journal of Forecasting 2006; 22: 679-688, https://doi.org/10.1016/j.ijforecast.2006.03.001.
22. Hyndman R J, Koehler A B, Snyder R D, Grose S. A state space framework for automatic forecasting using exponential smoothing methods. International Journal of Forecasting 2002; 18: 439-454, https://doi.org/10.1016/S0169-2070(01)00110-8.
23. Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E, Yasmeen F (2019). Forecast: Forecasting functions for time series and linear models. R package version 8.9, <URL: http://pkg.robjhyndman.com/forecast>.
24. Hyndman R J, Khandakar Y (2008). Automatic time series forecasting: the forecast package for R. Journal of Statistical Software 2008; 26(3): 1-22, https://doi.org/10.18637/jss.v027.i03.
25. Jimenez Cortadi A, Irigoien I, Boto F, Sierra B, Suarez A, Galar D. A statistical data-based approach to instability detection and wear prediction in radial turning processes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(3): 405-412, https://doi.org/10.17531/ein.2018.3.8.
26. Kontrec N Z, Milovanovic G V, Panic S R, Milosevic H. A reliability-based approach to nonrepairable spare part forecasting in aircraft maintenance system. Mathematical Problems in Engineering 2015: 731437, https://doi.org/10.1155/2015/731437.
27. Koucký M, Vališ D. Reliability of sequential systems with a restricted number of renewals. London: Taylor & Francis ltd. Risk, Reliability and Societal Safety, Vols. 1-3, ESREL 2007 Stavanger; 1845-1851.
28. Kozłowski E, Mazurkiewicz D, Żabinski T, Prucnal S, Sęp J. Assessment model of cutting tool condition for real-time supervision system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(4): 679-685, https://doi.org/10.17531/ein.2019.4.18.
29. Li H K, Zhou S, Kan H L, Cong M. Milling cutter condition reliability prediction based on state space model. Journal of Vibroengineering 2016; 18(7): 4312-4328, https://doi.org/10.21595/jve.2016.16648.
30. Pietrucha-Urbanik K, Studzinski A. Selected issues of costs and failure of pipes in an exemplary water supply system. Rocznik Ochrona Srodowiska 2016; 18: 616-627.
31. R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2015). URL https://www.R-project.org/ (accessed 16.06.03).
32. Rymarczyk T, Klosowski G. Innovative methods of neural reconstruction for tomographic images in maintenance of tank industrial reactors. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 261-267, https://doi.org/10.17531/ein.2019.2.10.
33. Snyder R D. Recursive estimation of dynamic linear models. Journal of the Royal Statistical Society 1985; 47(2): 272-276, https://doi.org/10.1111/j.2517-6161.1985.tb01355.x.
34. Snyder R. Discussion. International Journal of Forecasting 2006; 22(4): 673-676, https://doi.org/10.1016/j.ijforecast.2006.05.002.
35. Sousa H, Wang Y. Sparse representation approach to data compression for strain-based traffic load monitoring: A comparative study. Measurement 2018; 122: 630-637, https://doi.org/10.1016/j.measurement.2017.10.042.
36. Vališ D, Mazurkiewicz D. Application of selected Levy processes for degradation modelling of long range mine belt using real-time data. Archives of Civil and Mechanical Engineering 2018; 18(4): 1430-1440, https://doi.org/10.1016/j.acme.2018.05.006.

37. Vališ D, Žák L. Contribution to prediction of soft and hard failure occurrence in combustion engine using oil tribo data. Engineering Failure Analysis 2017; 82:583-598, https://doi.org/10.1016/j.engfailanal.2017.04.018.
38. Vališ D, Mazurkiewicz D, Forbelská M. Modelling of a Transport Belt Degradation Using State Space Model. In: Proceedings of the 2017 IEEE International Conference on Industrial Engineering & Engineering Management. Singapur: IEEE, 2017: 949-953, https://doi.org/10.1109/IEEM.2017.8290032.
39. Wang L, Lu Z R. Sensitivity-free damage identification based on incomplete modal data, sparse regularization and alternating minimization approach. Mechanical Systems and Signal Processing 2019; 120:43-68, https://doi.org/10.1016/j.ymssp.2018.10.013.
40. Whittle P. Hypothesis Testing in Time Series Analysis. Uppsala: Almquist and Wicksell 1951.

David VALIS
Department of Combat and Special Vehicles
University of Defence
Kounicova str. 65, 662-10 Brno, Czech Republic

Faculty of Transport and Computer Science,
University of Economics and Innovation,
Projektowa 4, 20-209 Lublin, Poland

Marie FORBELSKÁ
Department of Statistics and Operation Analysis
Mendel University in Brno
Zemědělská 1, 61300 Brno, Czech Republic

Zdeněk VINTR
Department of Combat and Special Vehicles
University of Defence
Kounicova str. 65, 662-10 Brno, Czech Republic

E-mails: david.valis@unob.cz, mforbelska@gmail.com,
zdenek.vintr@unob.cz