Complex Exponential Smoothing

Ivan Svetunkov[a,*], Nikolaos Kourentzes[a]

[a] *Lancaster University Management School*
Department of Management Science, Lancaster, LA1 4YX, UK

Abstract

Exponential smoothing has been one of the most popular forecasting methods for business and industry. Its simplicity and transparency have made it very attractive. Nonetheless, modelling and identifying trends has been met with mixed success, resulting in the development of different modifications of trend models. We present a new approach to time series modelling, using the notion of "information potential" and the theory of functions of complex variables. A new exponential smoothing method that uses this approach is proposed, the "Complex exponential smoothing" (CES). It has an underlying statistical model described in the paper and has several advantages in comparison with the customary exponential smoothing models, that allow CES to model and forecast effectively both trended and level time series, effectively overcoming the model selection problem.

Keywords: Forecasting, exponential smoothing, ETS, model selection, information potential, complex variables

1. Introduction

Exponential smoothing is a very successful group of forecasting methods which is widely used both in theoretical research (see for example Jose and Winkler (2008), Kolassa (2011), Maia and de A.T. de Carvalho (2011), Wang et al. (2012), Athanasopoulos and de Silva (2012), Kourentzes et al. (2014)) and in practise (see different forecasting competitions in Fildes et al.

*Correspondance: I Svetunkov, Department of Management Science, Lancaster University Management School, Lancaster, Lancashire, LA1 4YX, UK.

Email address: `i.svetunkov@lancaster.ac.uk` (Ivan Svetunkov)

(1998), Makridakis and Hibon (2000), Gardner Jr. and Diaz-Saiz (2008), Athanasopoulos et al. (2011)).

The exponential smoothing methods were well known and popular amongst practising forecasters but only in the last decade Single-Source of Errors (SSOE) state-space framework for the exponential smoothing methods (ETS) was proposed (Ord et al. (1997) and Hyndman et al. (2002)). This framework has been widely used since then in different modifications of the exponential smoothing (Gould et al. (2008), De Livera et al. (2011), Koehler et al. (2012), Taylor and Snyder (2012)).

Hyndman et al. (2008a) systematized all the existing exponential smoothing methods and showed that there may exist at least one out of five types of trends (none, additive, damped additive, multiplicative and damped multiplicative), one out of two types of errors (additive and multiplicative) and one out of three types of seasonal components (none, additive and multiplicative) in a time series. The taxonomy proposed by the authors leads to 30 exponential smoothing models that can underlay the different types of time series. Model parameters are optimized using maximum likelihood estimation and for the most exponential smoothing forms analytical variance expressions have been derived. The authors proposed to use information criteria for model selection which should allow choosing the most appropriate exponential smoothing model in each case. While Hyndman et al. (2008a) argued AICc to be the most appropriate information criterion, Billah et al. (2006) demonstrated that there was no significant difference in the forecasting accuracy when using different information criteria.

Furthermore Kolassa (2011) showed that the combination of several exponential smoothing models using Akaike weights produces more accurate forecasts than the forecast produced by a singe model using the same information criteria. Such a combination leads to composite forms of trend and seasonality, which can not be described by any single exponential smoothing form.

In addition to that Kourentzes et al. (2014) showed that the selection of an appropriate ETS components using the standard approach from (Hyndman et al., 2008a) may not lead to accurate forecasts in practice. Even if the correct types of components are known it is often impossible to produce the exponential smoothing model equivalent to the proposed by the authors Multiple Aggregation Prediction Algorithm (MAPA) because of the lack of the flexibility in the former.

Therefore there is a strong evidence that the model selection procedure in the exponential smoothing framework does not guarantee that the appropriate model is chosen. The possible cause of these problems is the general idea that any time series can be decomposed into several clear components. The focus of this article is on non-seasonal data, therefore we will limit any further discussion on the seasonal component. The data generating process for a very general non-seasonal case has the following form:

$$y_t = f(l_{t-1}, b_{t-1}, \epsilon_t), \tag{1}$$

where y_t is the value of the series, l_t is the level component, b_t is the trend component, ϵ_t is the error term and f is some function that allows including these components in either additive or multiplicative form.

However the composite forecasts discussed above hint to the lack of a clear separation between level and trend components. Furthermore the decomposition (1) can be considered arbitrary because depending on the chosen ETS model and its initial values different estimates of the time series components can be obtained. Besides, the trend component is unobservable directly in time series - it can appear only if an appropriate decomposition is done, and sometimes it is hard to distinguish a local-level time series from a trend series.

Consider for example simple exponential smoothing method (SES) proposed by Brown (1956) that has the following form:

$$\hat{y}_t = \alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1}, \tag{2}$$

where $\hat{y}_t$ is the calculated value and α is the smoothing parameter which in theory can lie inside the region (0, 2) (Brenner et al., 1968).

This exponential smoothing method has an underlying statistical model, ETS(A,N,N) which due to Hyndman et al. (2002) has the following state-space form:

$$\begin{cases} \hat{y}_t = l_{t-1} + \epsilon_t \\ l_t = l_{t-1} + \alpha\epsilon_t \end{cases} . \tag{3}$$

When the smoothing parameter in (3) increases the generated time series can reveal features of a trend time series. Furthermore when the smoothing parameter becomes greater than one, the series reveals even clearer global trend (see example in (Hyndman et al., 2008a, p.42)). Selecting the correct model in this situation becomes a challenging task. This is just one example of a data generating process that produces the time series with a complicated structure which is hard to define using the standard approach with time series decomposition (1).

In this paper, we propose a different approach to time series modelling that eliminates the arbitrary distinction between level and trend components, and therefore the model selection procedure. Instead of decomposing the time series into several components two characteristics can be studied: the value of the series y_t and an information potential p_t. The information potential is non-observable component of the series that characterizes the time series and influences the final actual value y_t. We define the information potential as an additional information that potentially can be included in the predicted values, given the observed values, thus the name of the variable.

Joining these two variables in one complex variable $y_t + ip_t$ (Svetunkov, 2012) allows taking both of them into account during the modelling process. Here i is the imaginary unit which satisfies the equation: $i^2 = -1$. The general data generating process using this notation will have the form:

$$y_t + ip_t = f(Q, \epsilon_t + i\xi_t), \tag{4}$$

where Q is a set of some complex variables, chosen for the prediction of y_t+ip_t, ϵ_t is the error term of the real part and ξ_t is so called "information gap" which shows the difference between the information potential and information included in the model.

An important notion that rises from the idea of the information potential is that although the actual value y_t and its prediction are the variables of the main interest in modelling, the unobserved information potential p_t contains additional useful information about the observed series. As such, it is necessary to include it in models, even though it is not the focus of the modelling exercise.

Due to the unobservability of the information potential it should be approximated by some characteristics of the studied time series. In this article it is proposed to approximate the information potential with an error term. To separate it from the real unknown information potential p_t we will use the following disambiguation:

$$\varsigma_t = y_t - \hat{y}_t = \epsilon_t \tag{5}$$

Using these ideas we propose the Complex Exponential Smoothing (CES) based on the data generating process (4), in analogy to the conventional exponential smoothing model.

2. Definition of CES

2.1. CES method and CES model

The complex exponential smoothing method is based on (2) and can be represented with the formula:

$$\hat{y}_{t+1} + i\hat{p}_{t+1} = (\alpha_0 + i\alpha_1)(y_t + i\varsigma_t) + (1 - \alpha_0 + i - i\alpha_1)(\hat{y}_t + i\hat{p}_t), \quad (6)$$

where $\hat{y}_t$ is the calculated value of series, $\hat{p}_t$ is the calculated value of information potential, $\alpha_0 + i\alpha_1$ is complex smoothing parameter and $(1-\alpha_0+i-i\alpha_1)$ is the equivalent complex smoothing parameter.

Analyzing (6) reveals that CES method resembles well known SES method (2): substituting all the real variables in (2) by complex variables leads to the formula (6). Further study of the properties of CES reveals some more features similar to SES. Thus any complex-valued function can be represented as a system of two real-valued functions, so we can represent CES in the following form:

$$\begin{cases} \hat{y}_{t+1} = (\alpha_0 y_t + (1-\alpha_0)\hat{y}_t) - (\alpha_1 \varsigma_t + (1-\alpha_1)\hat{p}_t) \\ \hat{p}_{t+1} = (\alpha_1 y_t + (1-\alpha_1)\hat{y}_t) + (\alpha_0 \varsigma_t + (1-\alpha_0)\hat{p}_t) \end{cases} \quad (7)$$

The system (7) allows to understand the underlying mechanism in CES better. It can be seen that the final forecast in CES consists of two parts: one is produced by SES method and the second, so called "information potential part" also employs the SES mechanism. It is obvious that CES is a non-linear method in it's nature as both first and second equations in (7) are connected with each other and change simultaneously depending on the complex smoothing parameter value. This complicated connection allows forecasting a bigger variety of time series than the conventional exponential smoothing methods, such as SES, Holt and other forecasting methods.

The underlying statistical model of CES can be derived using the idea that any complex variable can be represented in a vector and in matrix forms. The resulting model can be written in the following state-space form (see Appendix Appendix A):

$$\begin{cases} y_t = l_{t-1} + \epsilon_t \\ l_t = l_{t-1} - (1-\alpha_1)c_{t-1} - \alpha_1\varsigma_t + \alpha_0\epsilon_t \\ c_t = l_{t-1} + (1-\alpha_0)c_{t-1} + \alpha_0\varsigma_t + \alpha_1\epsilon_t \end{cases} \quad (8)$$

where l_t is the level component, c_t is the information potential component on observation t and $\epsilon_t \sim N(0, \sigma^2)$.

The state-space form (8) is not typical for standard exponential smoothing models but it has some similarities with trended ETS models. The model (8) can be used as a general form for any type of the information potential. It implies that dependencies in time series have a non-linear structure and no explicit trend component is present in the time series as this model does not need to artificially break the series into level and trend, as ETS does. This idea still allows to rewrite (8) in a shorter, more generic way, resembling the general SSOE state-space framework:

$$\begin{cases} y_t = w' x_{t-1} + \epsilon_t \\ x_t = F x_{t-1} + q \varsigma_t + g \epsilon_t \end{cases}, \tag{9}$$

where $x_t = \begin{pmatrix} l_t \\ c_t \end{pmatrix}$ is state vector, $F = \begin{pmatrix} 1 & -(1-\alpha_1) \\ 1 & 1-\alpha_0 \end{pmatrix}$ is transition matrix, $g = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}$ is persistence matrix, $q = \begin{pmatrix} -\alpha_1 \\ \alpha_0 \end{pmatrix}$ is information potential persistence matrix and $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

The state-space form (8) permits extending the CES in a similar ways that ETS has been extended to include additional states for seasonality or exogenous variables. The form (9) has a main difference compared with the conventional ETS models: it includes information potential term which in the situation of (5) acts as an additional error term which leads to a persistence vector having value of $g+q$. Furthermore the transition matrix in (9) includes smoothing parameters which is an exceptional feature for ETS models. This form allows understanding CES clearer as it uses notations of the generic SSOE state-space exponential smoothing but it uses the completely different approach in time series modelling.

2.2. Properties of CES

CES as a forecasting method has some interesting properties that should be studied. Substituting the calculated values in the right side of equation (6) gives the following recursive form of CES:

$$\begin{aligned} \hat{y}_{t+1} + i\hat{p}_{t+1} = \; & (\alpha_0 + i\alpha_1)(y_t + i\varsigma_t) + \\ & (\alpha_0 + i\alpha_1)(1 - \alpha_0 + i - i\alpha_1)(y_{t-1} + i\varsigma_{t-1}) + \\ & (1 - \alpha_0 + i - i\alpha_1)^2(\hat{y}_{t-1} + i\hat{p}_{t-1}) \end{aligned} \tag{10}$$

Repeating this procedure and assuming that time series has infinite length the following recursive form of CES can be obtained:

$$\hat{y}_{t+1} + i\hat{p}_{t+1} = (\alpha_0 + i\alpha_1)\sum_{j=0}^{T}(1 - \alpha_0 + i - i\alpha_1)^j(y_{t-j} + i\varsigma_{t-j}), \qquad (11)$$

where $T \to \infty$.

The right part of (11) contains all the observations weighed with complex weights forming geometric progression with complex common ratio:

$$(1 - \alpha_0 + i - i\alpha_1) \qquad (12)$$

Using this finding the bounds of complex smoothing parameter can be derived: the series of geometric progression in (11) will converge to some complex number only if the absolute value of (12) is less than one:

$$v = \sqrt{(1 - \alpha_0)^2 + (1 - \alpha_1)^2} < 1 \qquad (13)$$

The condition (13) is crucial for the preservation of the main exponential smoothing principle: the older observations should have smaller weights in the final forecast compared to the newer observaions. If this condition is violated the older observation would influence the forecast more than the newer observations. This condition is called stability condition (Hyndman et al., 2008b) and in the case of CES can be represented graphically on the plane as a circle with a centre with coordinates (1, 1); see Figure 1.

It is well known that any complex variable can be represented in algebraic, exponential and trigonometric forms. Using this property of complex variables the weights distribution in time can be studied in more details. The recursive form of CES (11) can be rewritten in trigonometric form the following way:

$$\hat{y}_{t+1} + i\hat{p}_{t+1} = R\sum_{j=0}^{T} v^j(\cos(\varphi + j\gamma) + i\sin(\varphi + j\gamma))(y_{t-j} + i\varsigma_{t-j}) \qquad (14)$$

Where $R = \sqrt{\alpha_0^2 + \alpha_1^2}$, v is taken from (13), $\varphi = \arctan\frac{\alpha_1}{\alpha_0} + 2\pi k$ and $\gamma = \arctan\frac{1-\alpha_1}{1-\alpha_0} + 2\pi k$, $k \in Z$. As there is no use in having $k \neq 0$ we will make an assumption that $k = 0$ in the calculation of all the polar angles in this research.

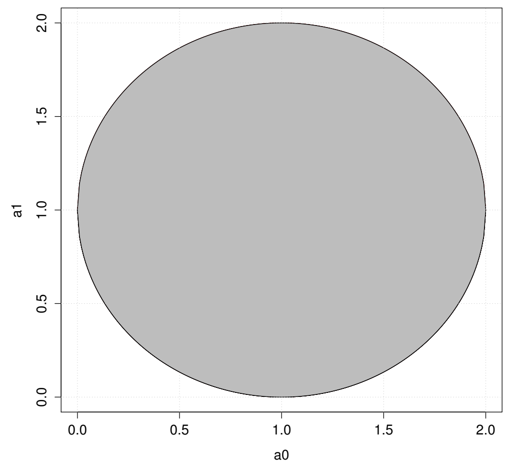


Figure 1: Bounds of complex smoothing parameter.

Multiplying complex variables in the sum of (14) will show how the real and imaginary parts of the forecast are calculated which in its turn can be presented in the following system:

$$\begin{cases} \hat{y}_{t+1} = R\sum_{j=0}^{T} v^j \cos(\varphi + j\gamma) y_{t-j} - R\sum_{j=0}^{T} v^j \sin(\varphi + j\gamma) \varsigma_{t-j} \\ \hat{p}_{t+1} = R\sum_{j=0}^{T} v^j \sin(\varphi + j\gamma) y_{t-j} + R\sum_{j=0}^{T} v^j \cos(\varphi + j\gamma) \varsigma_{t-j} \end{cases} \quad (15)$$

It can already be seen in (15) that the forecast in CES depends on the previous values of the series and the information potential, which are weighted in time using trigonometric functions. This means that weights in (15) can be distributed differently depending on the value of complex smoothing parameter. For example when $\alpha_0 + i\alpha_1 = 1.9 + 0.6i$ weights are distributed as shown on Figures 2a and 2b: they oscillate and diverge to zero slowly but nevertheless the absolute value of these weights decreases in time due to (13). This weights distribution should result in forecasts based on the long term time series characteristics. The other, more classical, example is shown on Figures 2c and 2d. The complex smoothing parameter used on these graphs is equal to $0.5 + 1.1i$ which results in the rapid exponential decline of weights in time. This weights distribution should result in forecasts using short-term characteristics of time series.

Plots 2a and 2b, showing the decrease of weights on complex plane, demonstrate what is happening with the equivalent complex smoothing parameter when it is exponentiated in the power j with $j = 1, ..., T$: the original vector is rotated on the complex plane. The direction and the degree of this rotation depends on the values of arguments γ and φ, which in their turns depend on the value of complex smoothing parameter $\alpha_0 + i\alpha_1$.

It is obvious that weights distribution depends on the complex smoothing parameter value and in its turn the speed of divergence of these weights also depends on this value. It can be shown that if the absolute value of equivalent complex smoothing parameter v in (13) becomes close to one, the speed of convergence becomes very low which results in flatter weights distribution between the observations and a higher complex weight of the initial value. Vice versa when v becomes close to zero the speed of convergence becomes very high which results in the steep exponential form of weights distribution. As a result only several last observations are used in the final forecast.

Plots on the Figure 2 also demonstrate that due to the complex nature of the weights their sum should be a complex number. For the comparison the sum of weights of ETS(A,N,N) is always equal to one. So to calculate the sum of weights of CES the infinite geometric progression of weights in (11) should be assumed. The sum of this series will be:

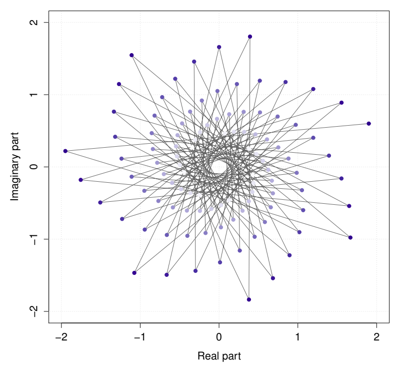
$$S = (\alpha_0 + i\alpha_1) \sum_{j=0}^{\infty} (1 - \alpha_0 + i - i\alpha_1)^j = \frac{\alpha_0^2 - \alpha_1 + \alpha_1^2 + i\alpha_0}{\alpha_0^2 + (1 - \alpha_1)^2} \qquad (16)$$

Analysing (16) it can be seen that the sum of weights depends on the value of the complex smoothing parameter and is usually represented by the complex number. In theory the sum of series can become real number only when $\alpha_0 = 0$ which contradicts with the condition (13). It can also be concluded that the real part of S can be any real number while the imaginary part of S is restricted with positive real numbers only. The value S indicates that CES is not an averaging model in comparison with SES which sometimes is also called "exponentially weighted moving average". S can also be used in the analysis of complex weights distribution in CES and allows comparing different CES estimated for different time series.

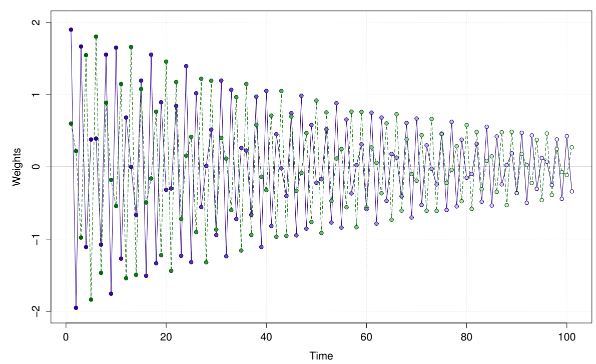
3. Connection with other forecasting models

3.1. Underlying ARIMA

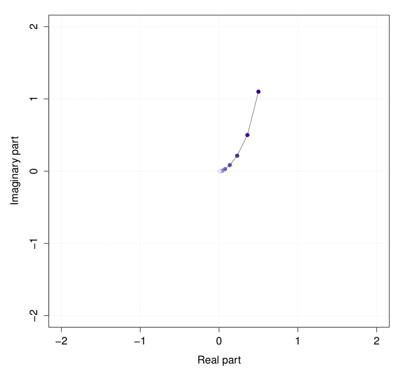
The majority of exponential smoothing models have equivalent underlying ARIMA models. For example, ETS(A,N,N) has underlying ARIMA(0,1,1)



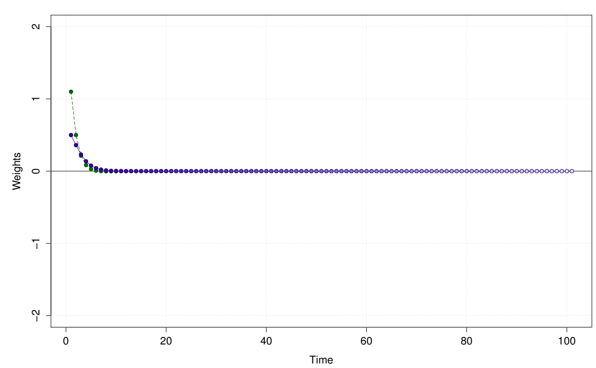
(a) on complex plane



(b) in time



(c) on complex plane



(d) in time

Figure 2: CES weights distribution. 2b and 2d: blue solid lines - real parts of complex weights, green dashed lines - the imaginary parts of complex weights.

model which was shown for example in (Gardner, 1985). CES also has an underlying ARIMA model but it is much more complicated than ARIMA models for the other exponential smoothing models. The analysis of the system (15) shows that the real and the imaginary parts of the forecasted complex variable are independent when they are represented as an infinite series of weighted actual values and information potential. Thus the real part can be analysed separately:

$$\hat{y}_t = R\sum_{j=1}^{T} v^{j-1}\cos(\varphi + (j-1)\gamma)y_{t-j} - R\sum_{j=1}^{T} v^{j-1}\sin(\varphi + (j-1)\gamma)\varsigma_{t-j} \quad (17)$$

The information potential ς_t can be substituted in (17) by ϵ_t due to (5). After other several simple substitutions the following formula of CES can be obtained:

$$y_t - R\sum_{j=1}^{T} v^{j-1}\cos(\varphi+(j-1)\gamma)y_{t-j} = \epsilon_t - R\sum_{j=1}^{T} v^{j-1}\sin(\varphi+(j-1)\gamma)\epsilon_{t-j} \quad (18)$$

Which in it's turn can be rewritten using the backshift operator B:

$$\left(1 - R\sum_{j=1}^{T} v^{j-1}\cos(\varphi + (j-1)\gamma)B^j\right) y_t = \left(1 - R\sum_{j=1}^{T} v^{j-1}\sin(\varphi + (j-1)\gamma)B^j\right)\epsilon_t$$

Substituting terms $Rv^{j-1}\cos(\varphi + (j-1)\gamma) = \phi_j$ and $Rv^{j-1}\sin(\varphi + (j-1)\gamma) = \theta_j$ before the backshift operator the following model can be obtained:

$$\left(1 - \sum_{j=1}^{T} \phi_j B^j\right) y_t = \left(1 - \sum_{j=1}^{T} \theta_j B^j\right)\epsilon_t \quad (19)$$

It is obvious that (19) is an ARMA(T,T) model and the order of this model depends on the complex smoothing parameter value. We have already discussed earlier in this paper that the convergence of weights in (19) depends on the complex smoothing parameter value: if it is close to the centre of the circle in Figure 1 the weights will converge to zero rapidly; if it is close to the bounds, the weights will converge slowly. In terms of (19) this means that the order of ARMA model underlying CES will be very high when the absolute value of the equivalent smoothing parameter (13) is close to 1. And per contra when (13) is close to 0 the order of an underlying ARMA model becomes very low as many coefficients for $j > 1$ in (19) become close to zero.

The equation (19) also allows to derive the general stationarity conditions of CES. To determine if CES is stationary with the selected complex smoothing parameter the following characteristic equation for B should be solved:

$$1 - \sum_{j=1}^{T} \phi_j B^j = 0 \tag{20}$$

This characteristic equation will in general have T roots, and if all of them lie outside the unit circle, the model can be considered stationary. When the complex smoothing parameter lies near the centre of the circle (13) the other characteristic equation, with the smaller number of polynomial terms, can be solved instead of (20) but the choice of the order of this smaller equation is subjective. For example when $\alpha_0 + i\alpha_1 = 0.8 + i$ the characteristic equation can be limited with the polynomial of order 4 because the complex weight for $j = 5$ is already very small and is equal to $0.00128 + 0.0016i$. As the result the characteristic equation will have the following form:

$$1 - 0.8B - 0.16B^2 - 0.032B^3 - 0.0064B^4 = 0 \tag{21}$$

Solving it will result in the following roots: 1.001; -5.512; $-0.245-5.315i$; $-0.245+5.315i$ Each of these roots lie outside the unit circle. This means that CES has an underlying stationary ARMA process when $\alpha_0 + i\alpha_1 = 0.8 + i$.

Unfortunately solving characteristic equations when the complex smoothing parameter lies near its bounds (when $v \to 1$) is almost impossible due the high order of underlying ARMA, so this method of stationarity check is not universal for CES.

It should also be noted that when the complex smoothing parameter lies near its bounds the order of underlying ARMA(T,T) becomes so high that the corresponding ARMA can not be estimated directly. Using CES allows to do this estimation in a compact form, where the complex variable $\hat{y}_t + i\hat{p}_t$ holds the information about all the previous values of $y_t + i\varsigma_t$ weighted over time using complex weights.

3.2. Comparison of CES with SES and ETS(A,N,N)

The other interesting feature of CES is that it can be perceived as a more general method for SES and as a more general model for ETS(A,N,N). When $\varsigma_t = 0$ and $\alpha_1 = 1$ the system (7) becomes:

$$\begin{cases} \hat{y}_{t+1} = \alpha_0 y_t + (1 - \alpha_0)\hat{y}_t \\ \hat{p}_{t+1} = y_t + (1 - \alpha_0)\hat{p}_t \end{cases} \tag{22}$$

It is obvious that the first equation in (22) is SES method. It is remarkable that both equations in (22) become independent from each other in these conditions and the second equation in (22) is not needed for the purpose of forecasting.

In our framework the condition $\varsigma_t = 0$ means that the information potential is assumed to be equal to zero, implying that no additional information is used in the time series generation. This condition along with $\alpha_1 = 1$ results in the different CES model derived from (8) which takes form (see Appendix Appendix B):

$$\begin{cases} y_t = l_{t-1} + \epsilon_t \\ l_t = l_{t-1} + \alpha_0 \epsilon_t \\ c_t = \frac{l_{t-1}}{\alpha_0} + \frac{\epsilon_t}{\alpha_0} \end{cases} \tag{23}$$

Now the series is generated only using the level component and no information potential is produced by the model. Furthermore the level and information potential components become independent in the transition equation in (23). As the result the level component is generated the same way as in ETS(A,N,N), so CES model becomes almost equivalent to ETS(A,N,N) model (3), with the only difference in the presence of non-observable information potential component that does not interact with y_t.

When the discussed condition is met the weights in CES are distributed similarly to SES and converge to the complex number (16) which becomes equal to:

$$S = \frac{\alpha_0^2 + i\alpha_0}{\alpha_0^2} = 1 + i\frac{1}{\alpha_0} \tag{24}$$

The value of S in (24) resembles the sum of weights in SES which is well known to be equal to 1.

Lastly the stability region of CES in this situation become equivalent to the stability region of SES:

$$
\begin{aligned}
v &= \sqrt{(1-\alpha_0)^2 + (1-1)^2} < 1 \\
v &= |1-\alpha_0| < 1
\end{aligned} \tag{25}
$$

Which is equivalent to the region $(0, 2)$ for the smoothing parameter discussed earlier.

An additional study of (7) reveals an important information about the complex smoothing parameter and its connection to SES smoothing parameter. If the substitution $\varsigma_t = \epsilon_t$ is made in (7) then after the regrouping of elements the following system will be obtained:

$$
\begin{cases}
\hat{y}_{t+1} = \hat{y}_t + (\alpha_0 - \alpha_1)\epsilon_t - (1-\alpha_1)\hat{p}_t \\
\hat{p}_{t+1} = \hat{y}_t + (\alpha_0 + \alpha_1)\epsilon_t + (1-\alpha_0)\hat{p}_t
\end{cases} \tag{26}
$$

When α_1 is close to one the influence of $\hat{p}_t$ on $\hat{y}_{t+1}$ becomes minimal and the second smoothing parameter α_0 in (26) starts acting similar to the smoothing parameter in SES with a small difference: $\alpha_0 - 1$ in CES becomes close to α in SES. For example the value $\alpha_0 = 1.2$ in CES will correspond to $\alpha = 0.2$ in SES. This also means that when the series is stationary and optimal smoothing parameter in SES should be close to zero the optimal α_0 in CES should be close to one.

3.3. Comparison of CES with ETS(A,A,N)

Another exponential smoothing model that is close in its form to CES is ETS(A,A,N). The general state-space form of ETS(A,A,N) is:

$$
\begin{aligned}
y_t &= l_{t-1} + b_{t-1} + \epsilon_t \\
l_t &= l_{t-1} + b_{t-1} + \alpha\epsilon_t \\
b_t &= b_{t-1} + \beta\epsilon_t
\end{aligned} \tag{27}
$$

Where l_t is level component, b_t is trend component of time series, α and β are smoothing parameters.

To see the similarity with this model, CES state-space model (8) can be split into two systems, where the first system represents the real and the second represents the imaginary part of (8):

$$
\begin{cases}
y_t = l_{0,t-1} - c_{0,t-1} + \epsilon_t \\
l_{0,t} = l_{0,t-1} - c_{0,t-1} + \alpha_0\epsilon_t \\
c_{0,t} = (1-\alpha_1)l_{1,t-1} + (1-\alpha_1)c_{1,t-1} + \alpha_1\varsigma_t
\end{cases} \tag{28}
$$

$$\begin{cases} \hat{p}_t = l_{1,t-1} + c_{1,t-1} \\ l_{1,t} = l_{0,t-1} - c_{0,t-1} + \alpha_1 \epsilon_t \\ c_{1,t} = (1-\alpha_0) l_{1,t-1} + (1-\alpha_0) c_{1,t-1} + \alpha_0 \varsigma_t \end{cases} \quad (29)$$

Where $l_{0,t}$ and $l_{1,t}$ are level estimates, $c_{0,t}$ and $c_{1,t}$ are non-linear trend estimates used in real and imaginary parts of CES respectively.

Comparing first two equations in (27) with first two equations in (28) shows some similarities: both ETS(A,A,N) and CES contain level estimates (l_t and $l_{0,t}$ respectively) and some components that correspond to trend component (b_t and $-c_{0,t}$ respectively). Both these models are based on the level component smoothing using the error term ϵ_t. The differences in the models appear in the third equations in (27) and (28): the trend component is smoothed directly in ETS(A,A,N) while CES smooths it non-linearly, using the second equation (29).

4. Statistical properties

4.1. State-space CES with the specific information potential

Due to the used substitute of the information potential (5) some of the elements in the (8) can be substituted which will lead to the different state-space model underlying CES:

$$\begin{cases} y_t = l_{t-1} + \epsilon_t \\ l_t = l_{t-1} - (1-\alpha_1) c_{t-1} + (\alpha_0 - \alpha_1) \epsilon_t \\ c_t = l_{t-1} + (1-\alpha_0) c_{t-1} + (\alpha_0 + \alpha_1) \epsilon_t \end{cases} \quad (30)$$

The main difference between the general state-space CES (8) and the state-space CES (30) is that due to the information potential substitute (5) the smoothing parameter changes in both equations of the transition equation (30) which leads to additional non-linearity in the model. The new persistence vector allows to easily estimate the conditional variance of the model while the transition matrix can be used in the conditional mean estimation.

4.2. Likelihood function

The state-space form of CES (30) shows that the error term is included in the models additively. As the result the likelihood function for CES is trivial and is similar to the likelihood function of additive exponential smoothing models (Hyndman et al., 2008a, p.68):

$$L(g, x_0, \sigma^2|y) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^T \exp\left(-\frac{1}{2}\sum_{t=1}^{T}\left(\frac{\epsilon_t}{\sigma}\right)^2\right) \tag{31}$$

It was shown in (Hyndman et al., 2008a, p.68) that maximizing the likelihood function (31) is equivalent to minimizing the sum of squared errors:

$$SSE = \sum_{t=1}^{t} \epsilon_t^2 \tag{32}$$

This finding can be used in CES parameters estimation. Obtaining likelihood function (31) for CES also allows calculating information criteria which are based on double negative concentrated log-likelihood function value:

$$-2\log(L(g, x_0|y)) = T\left(\log\left(\frac{2\pi e}{T}\right) + \log\left(\sum_{t=1}^{T}\epsilon_t^2\right)\right) \tag{33}$$

Using (33) CES can be compared with its potential modifications by the means of the information criteria calculation.

4.3. Conditional mean and variance of CES

The conditional mean of CES for h steps ahead with known l_t and c_t can easily be calculated using the state-space (30):

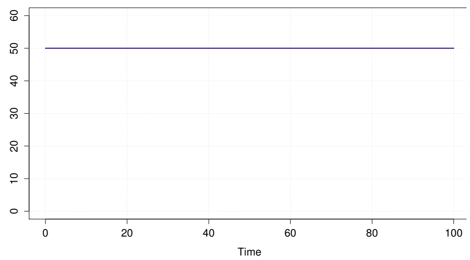
$$E\left(z_{t+h}|x_t\right) = F^{h-1}E(x_t) \tag{34}$$

where $E(z_{t+h}|x_t) = \hat{y}_{t+h} + i\hat{x}_{t+h}$.

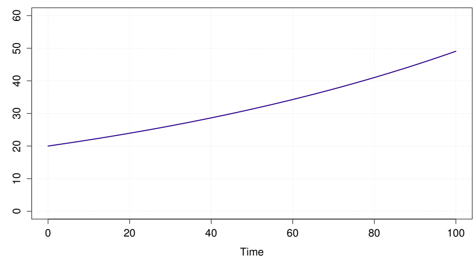
Depending on the complex smoothing parameter value forecasting trajectories can be obtained using (34) and (30). The trajectories will differ depending on the values of l_t and c_t but the complex smoothing parameter has the main importance.

Analysing (34) it can be explicitly stated that when $\alpha_1 = 1$ all the values of the forecast will be equal to the last obtained forecast, which produces the flat line, equivalent to the forecast produced by level ETS. This trajectory is shown on the Figure 3a.

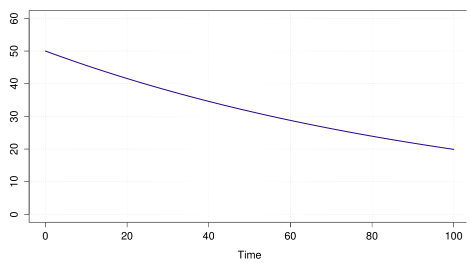
The simulation study done with different values of complex smoothing parameter and with non-negative values of l_t and c_t showed that when $\alpha_1 > 1$ CES produces the trajectory with growth which is shown on Figure 3b. When $\alpha_1 < 1$ and $\alpha_0 < 1$ CES produces the harmonic trajectory and when in



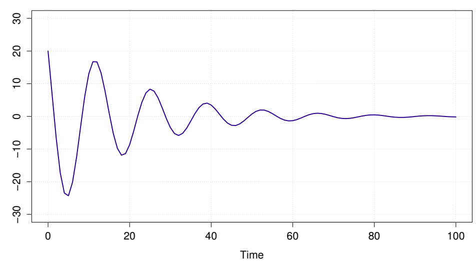
(a) the flat line



(b) exponential growth



(c) exponential decline



(d) harmonic oscilation

Figure 3: Forecasting trajectories.

addition to that $\alpha_0 + \alpha_1 > 1$ the forecasting trajectory will converge to zero (see Figure 3d). When $\alpha_1 < 1$ and $\alpha_0 > 1$ CES produces the exponential decline as shown on Figure 3c.

The conditional mean estimated using (34) consists of two parts: the conditional mean of the series y_t and the conditional mean of the information potential substitute w_t.

The state-space form of CES (30) also allows calculating the conditional variance of CES for h steps ahead with known l_t and c_t which can also be calculated using the general state-space form (9) similar to (Hyndman et al., 2008a, p.96):

$$V\left(z_{t+h}|x_t\right) = \begin{cases} \sigma_\epsilon^2 \left(J_2 + \sum_{j=1}^{h} F^{j-1} g g' (F')^{j-1}\right) & \text{when h} > 1 \\ \sigma_\epsilon^2 & \text{when h=1} \end{cases}, \quad (35)$$

where $J_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

The conditional variance estimated using the formula (35) is in fact a variance-covariance matrix that contains the variance of the series y_t, the variance of the information potential substitute w_t and the covariance between these two variables. However the variance of the actual series is of the main interest, so all the other values can in general be ignored.

4.4. Stationarity and stability conditions for CES

Due to (Hyndman et al., 2008a, p.38) the stationarity condition for the general exponential smoothing in the state-space form (9) is that all the eigenvalues of F should lie inside the unit circle. The majority of ETS models are non-stationary but calculating eigenvalues for F of CES gives the following roots:

$$\lambda = \frac{2 - \alpha_0 \pm \sqrt{\alpha_0^2 + 4\alpha_1 - 4}}{2} \tag{36}$$

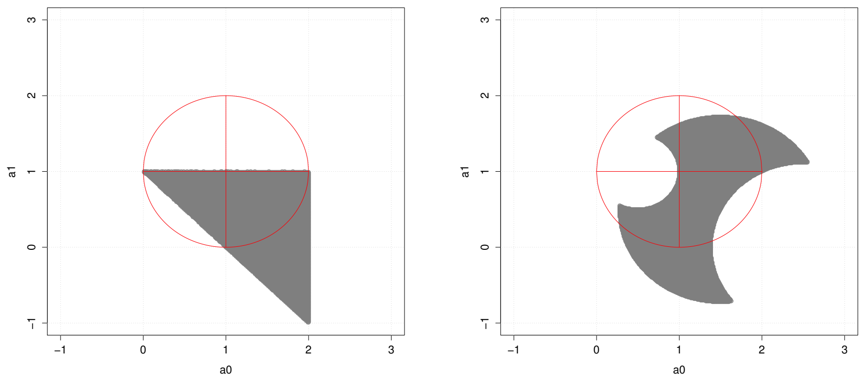
If the absolute values of both roots in (36) is less than 1 then the estimated CES is stationary. The corresponding statinarity region for CES is shown on the Figure 4a. The most important part of this region can be described by two inequalities:

$$\begin{cases} \alpha_1 \leq 1 \\ \alpha_0 + \alpha_1 \geq 1 \end{cases} \tag{37}$$

Having (37) allows to conclude whether the forecasting trajectory of CES will be stationary or not by analysing the estimated complex smoothing parameter.

The other important thing that arises from (30) is stability condition for CES: the general stability condition for CES has already been derived in (13) but due to the differences in state-space forms it should be reestimated. To derive it we firstly need to insert $\epsilon_t = y_t - l_{t-1}$ in the transitional equation in (30). After several manipulations the following system of equations can be obtained:

$$\begin{cases} y_t = l_{t-1} + \epsilon_t \\ \begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 - \alpha_0 + \alpha_1 & -(1-\alpha_1) \\ 1 - \alpha_0 - \alpha_1 & 1 - \alpha_0 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} \alpha_0 - \alpha_1 \\ \alpha_1 + \alpha_0 \end{pmatrix} y_t \end{cases} \tag{38}$$



(a) Stationarity region (b) Stability region

Figure 4: Stability and stationarity regions of CES, derived from state-space form (30). Circle represents the stability condition (13).

The matrix $D = \begin{pmatrix} 1-\alpha_0+\alpha_1 & -(1-\alpha_1) \\ 1-\alpha_0-\alpha_1 & 1-\alpha_0 \end{pmatrix}$ is called discount matrix and can be estimated in the general form:

$$D = F - gw' \tag{39}$$

The statistical model is stable if all the eigenvalues of (39) lie inside the unit circle. The eigenvalues can be calculated using the following formula:

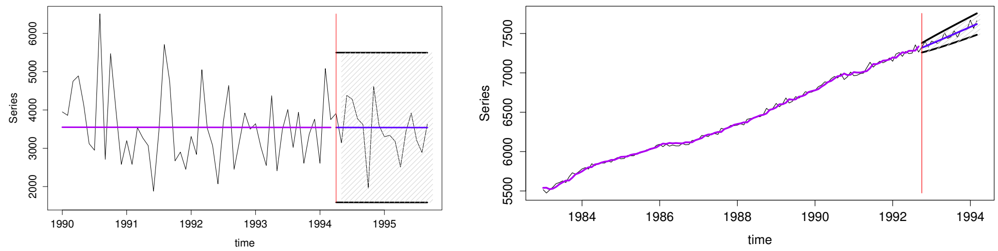
$$\lambda = \frac{2 - 2\alpha_0 + \alpha_1 \pm \sqrt{8\alpha_1 + 4\alpha_0 - 4\alpha_0\alpha_1 - 4 - 3\alpha_1^2}}{2} \tag{40}$$

The stability condition (40) corresponds to the region showed on the Figure 4b. It can be noted that the stability region of CES in the state-space form (30) has some common areas with the region for the general CES (13) though it also has some unique areas which should be taken into account during the parameters estimation.

5. Examples

5.1. Real life time series

We will estimate CES on two time series to see how it works in real life: level series and trend series from M3-Competition. The first series (number



(a) Series number 1664 (b) Series number 2721

Figure 5: Stationary and trended time series. Red line divides the in-sample with the holdout-sample.

1664) is shown on the Figure 5a and the second series (number 2721) is shown on the Figure 5b.

Graphical analysis of the Figure 5a shows that the series is stationary. ADF test for the series 1664 gives the p-value of less than 0.01 (thus the null hypothesis of a unit root is rejected on 0.05 level) while KPSS test results in the p-value of greater than 0.1 (the null hypothesis of level stationarity can not be rejected on 0.05 level). All of these indicate that the series is indeed stationary.

The second time series (Figure 5b) is not stationary. In fact it has a clear trend. ADF test gives the p-value of more than 0.8 while KPSS test shows p-value of less than 0.01. Comparing these values with the critical value of 0.05 allows to conclude that the series is indeed not stationary.

Estimation of CES on the first time series results in the complex smoothing parameter $\alpha_0 + i\alpha_1 = 0.99999 + 0.99996i$. All the roots of the characteristic equation for such a complex smoothing parameter lie outside the unit circle and inequality (37) is satisfied which means that the model produces the stationary trajectory. Furthermore the fact that both real and imaginary parts of the complex smoothing parameter are close to each other means that the error variance will weakly influence the variance of the level component.

Fitted values and both point and 95% interval forecasts produced by this CES are shown on the Figure 5a. The fitted line of CES is a flat line that lies in the middle of the series, the forecast has almost no decline. This indicates that CES was able to identify that the series is stationary. All the observations in the holdout sample lie in the 95% forecasting interval.

CES estimated on the second time series has complex smoothing parameter $\alpha_0 + i\alpha_1 = 1.48187 + 1.00352i$. This means that the forecast of CES on the second time series is influenced by the larger number of observations compared with the CES on the first time series. There are several roots of characteristic equation lying inside the unit circle and the imaginary part of the complex smoothing parameter is greater than one. All of this indicates that the model is non-stationary.

Fitted values and both point and 95% interval forecasts produced by CES on the second time series are shown on the Figure 5b. The fitted line is not as smooth as for the first time series but still goes through the series. All the observations still lie inside the interval which may indicate that the variance was estimated correctly.

These two examples show that CES is capable of identifying if the series is stationary or not and producing the appropriate forecast without the need for the model selection procedure.

5.2. M3 competition results

We have conducted the experiment to see if the forecasting performance of CES is significantly better than the performance of other forecasting methods. All non-seasonal monthly time series from M3-Competition (Makridakis and Hibon, 2000) have been included in the experiment resulting in 814 time series. The following competing forecasting methods have been included in the experiment:

- Naive - Random walk model with Naive method;
- SES - ETS(A,N,N) which corresponds to Simple Exponential Smoothing method;
- AAN - ETS(A,A,N) which corresponds to Holt's method (Holt, 2004);
- MMN - ETS(M,M,N) which underlies Pegel's method (Pegels, 1969);
- AAdN - ETS(A,Ad,N) which underlies additive damped trend method (Gardner and McKenzie, 1985);
- MMdN - ETS(M,Md,N) which corresponds to multiplicative damped trend method (Taylor, 2003);
- ZZN - ETS(Z,Z,N) - the general exponential smoothing model with the model selection procedure proposed by (Hyndman et al., 2002);

Forecasting Method	Mean MASE	Median MASE
Naive	3.216	1.927
SES	3.098	1.681
AAN	2.755	1.602
MMN	3.323	1.703
AAdN	2.726	1.467
MMdN	2.700	1.490
ZZN	2.716	1.466
CES	**2.496**	**1.387**

Table 1: MASE values for different forecasting methods

- CES - Complex Exponential Smoothing.

The size of the holdout sample for monthly data in M3-Competition is 18 observations, so we produced forecasts 18 observations ahead using each of the models and calculated MASEs as proposed in (Hyndman and Koehler, 2006) to compare the forecasting accuracies of the models.

Mean and median MASEs over all the horizons are shown in the Table 1 (with CES values in bold). It can be noted that CES has a lowest mean and median MASE compared to all the other models used on this set of data. Remarkable that it managed not only to outperform ETS(A,N,N) and ETS(A,A,N) models but also ETS(Z,Z,N) with the implemented model selection procedure.

To see why CES is more accurate than the other models mean and median MASEs were also calculated for each of the horizons. It appeared that the difference in accuracies between CES and other models increases with the increase of the forecasting horizon. For example, CES was significantly more accurate than the second best method MMdN for horizon $h = 14, \dots, 18$ but it was not the very best method for the shorter horizon of $h = 1, \dots, 3$. This indicates that CES was able to capture the long-term dependencies in time series compared to the other exponential smoothing models that managed capturing the short-term dependencies.

Conclusions

We presented a new approach to time series analysis in this paper. Instead of taking into account only one real variable, the actual value of series, and decomposing it into several components, we introduced the information potential variable and combined it with the actual series value in one complex variable. Using this approach instead of using the arbitrary decomposition of time series into several components (level, trend, seasonality, error) leads to a new class of models, one of which is the Complex Exponential Smoothing that was studied in this paper.

The proposed CES is a flexible model that is able to distribute weights between different observations in time either exponentially or harmonically. This feature allows CES to capture long-term dependencies and non-linear relations in time series.

It was shown that CES has an underlying ARMA model, the order of which varies depending on the value of complex smoothing parameter. When complex smoothing parameter lies near its bounds, the underlying ARMA becomes of a high order and when the parameter lies near the centre of the bounds, ARMA has a small order. Using this finding the characteristic equation for the corresponding ARMA can be solved for the stationarity condition check for the original CES.

Furthermore the state-space form using Single Source of Error was presented in the paper, which allowed to derive stability and stationarity conditions of CES and the mean and variance of the model. The proposed state-space form also allows including exogenous variables and makes possible the calculation of likelihood function.

The comparison of CES with ETS(A,N,N) showed that the real part of complex smoothing variable is connected to the smoothing parameter of ETS(A,N,N) while the imaginary part of complex smoothing parameter regulates the direction of the forecast trajectory: if it is greater than one, CES will produce growth, when it it less than one, CES will produce decline. An additional comparison with ETS(A,A,N) shows that CES is a new forecasting model that is based on non-linear trend while ETS(A,A,N) uses linear trend.

We also showed that CES is a flexible model that can produce different types of trajectories in time series and is capable of capturing trends. Using CES makes model selection procedure obsolete as it encompasses both level and multiplicative trend series and also approximates additive trend very well.

Finally, the evaluation that we conducted showed that CES is more accurate than all the included methods, especially ETS(A,N,N), ETS(A,A,N) and automated ETS. This provides evidence that CES can be used instead of these models, capturing both level and trend time series cases. Therefore, we argue that CES effectively overcomes the model selection problem that conventional exponential smoothing and similar models face, that is to distinguish between level and trend series, as well as the nature of the trend of the time series. Evidence from our evaluation suggests that for univariate extrapolative forecasting CES can effectively replace the multiple forms of level and trend exponential smoothing, thus simplifying the forecasting problem. Furthermore, our analysis indicates that a key advantage of CES in comparison to the other models is that it is able to capture long-term dependencies which results in more accurate forecasts for the longer horizons.

In conclusion, the Complex Exponential Smoothing model proposed in this paper has unique and desirable properties for time series forecasting. Using the ideas of complex variables and information potential, CES builds on the established and widely used exponential smoothing ideas to overcome several limitations and modelling challenges of the latter. The current work focused on non-seasonal time series that are important for a wide variety of forecasting applications and needs. However, it did not consider seasonal time series or incorporating exogenous variables in the forecasts, as it was out of the scope of this analysis. However future works should explore extensions of CES in those directions, investigating how the information potential interacts with seasonality and exogenous variables and how to best incorporate these into CES models.

Appendix A. State-space form of CES

First of all any complex variable can be represented as a vector and as a matrix:

$$a + ib = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{A.1}$$

The general CES model (6) can be split into two parts: measurement and transitional equations using (A.1):

$$\begin{cases} \begin{pmatrix} \hat{y}_t \\ \hat{x}_t \end{pmatrix} = \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} \\ \begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} \alpha_0 & -\alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix} \begin{pmatrix} y_t \\ \varsigma_t \end{pmatrix} + \left(\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} \alpha_0 & -\alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix} \right) \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} \end{cases} \tag{A.2}$$

Regrouping the elements of transitional equation in (A.2) the following system of equations can be obtained:

$$\begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} \alpha_0 & -\alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix} \begin{pmatrix} 0 \\ \varsigma_t \end{pmatrix} + \begin{pmatrix} \alpha_0 & -\alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix} \begin{pmatrix} y_t \\ 0 \end{pmatrix} + \\ \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} - \begin{pmatrix} \alpha_0 & -\alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix} \begin{pmatrix} 0 \\ c_{t-1} \end{pmatrix} - \begin{pmatrix} \alpha_0 & -\alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ 0 \end{pmatrix} \tag{A.3}$$

Grouping vectors actual value and level component with complex smoothing parameter and then the level and information potential components leads to:

$$\begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} - \begin{pmatrix} 0 & -\alpha_1 \\ 0 & \alpha_0 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \\ \begin{pmatrix} \alpha_0 & -\alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix} \begin{pmatrix} 0 \\ \varsigma_t \end{pmatrix} + \begin{pmatrix} \alpha_0 & -\alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix} \begin{pmatrix} y_t - l_{t-1} \\ 0 \end{pmatrix} \tag{A.4}$$

The difference between the actual value and the level in (A.4) is the error term: $y_t - l_{t-1} = \epsilon_t$. Using this and after several transformations the following state-space model will be obtained:

$$\begin{cases} \begin{pmatrix} \hat{y}_t \\ \hat{x}_t \end{pmatrix} = \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} \\ \begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 & -(1-\alpha_1) \\ 1 & (1-\alpha_0) \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} -\alpha_1 \\ \alpha_0 \end{pmatrix} \varsigma_t + \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \epsilon_t \end{cases} \tag{A.5}$$

Now if CES should be represented in the state-space form with the SSOE then the measurement equation should also contain the same error term as the transitional equation. Alas imaginary part of the measurement equation in (A.5) is unobservable, independent from the real part and does not contain any information useful for the forecasting. This is why it can be excluded from the final state-space model:

$$\begin{cases} y_t = l_{t-1} + \epsilon_t \\ \begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 & -(1-\alpha_1) \\ 1 & (1-\alpha_0) \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} -\alpha_1 \\ \alpha_0 \end{pmatrix} \varsigma_t + \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \epsilon_t \end{cases} \tag{A.6}$$

The other way of writing down this state-space model is by splitting the level and information potential components into two equations:

$$\begin{cases} y_t = l_{t-1} + \epsilon_t \\ l_t = l_{t-1} - (1-\alpha_1)c_{t-1} - \alpha_1 \varsigma_t + \alpha_0 \epsilon_t \\ c_t = l_{t-1} + (1-\alpha_0)c_{t-1} + \alpha_0 \varsigma_t + \alpha_1 \epsilon_t \end{cases} \tag{A.7}$$

Appendix B. The connection of CES and ETS(A,N,N)

When $\varsigma_t = 0$ and $\alpha_1 = 1$ the information gap will be equal to the negative information potential component $\xi_t = \varsigma_t - c_t = -c_t$ and based on (8) the following state-space model can be obtained:

$$\begin{cases} \begin{pmatrix} y_t \\ 0 \end{pmatrix} = \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ -c_t \end{pmatrix} \\ \begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} \alpha_0 & -1 \\ 1 & \alpha_0 \end{pmatrix} \begin{pmatrix} \epsilon_t \\ -c_t \end{pmatrix} \end{cases} \tag{B.1}$$

The information potential component in the measurement equation in (B.1) now becomes constant: $c_t = c_{t-1}$ - which allows to substitute the component c_{t-1} in the right hand side of the second equation:

$$\begin{cases} \begin{pmatrix} y_t \\ c_t \end{pmatrix} = \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ 0 \end{pmatrix} \\ \begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_t \end{pmatrix} + \begin{pmatrix} \alpha_0 & -1 \\ 1 & \alpha_0 \end{pmatrix} \begin{pmatrix} \epsilon_t \\ -c_t \end{pmatrix} \end{cases} \tag{B.2}$$

And after several substitutions and cancelling outs the following state-space model will be obtained:

$$\begin{cases} \begin{pmatrix} y_t \\ c_t \end{pmatrix} = \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ 0 \end{pmatrix} \\ \begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_t \end{pmatrix} + \begin{pmatrix} \alpha_0 & 0 \\ 1 & \alpha_0 \end{pmatrix} \begin{pmatrix} \epsilon_t \\ -c_t \end{pmatrix} \end{cases} \tag{B.3}$$

The measurement equation in (B.3) can be substituted by the simpler, univariate equation, due to the constancy of the information potential component and the absence of the information potential. Besides dividing level and information potential components into two equations in the transition equation of (B.3) leads to the following state-space model:

$$
\begin{cases} y_t = l_{t-1} + \epsilon_t \\ \\ l_t = l_{t-1} + \alpha_0 \epsilon_t \\ c_t = (1-\alpha_0)c_t + l_{t-1} + \epsilon_t \end{cases} \tag{B.4}
$$

Which after simple substitutions finally leads to the following model:

$$
\begin{cases} y_t = l_{t-1} + \epsilon_t \\ \\ l_t = l_{t-1} + \alpha_0 \epsilon_t \\ c_t = \frac{l_{t-1}}{\alpha_0} + \frac{\epsilon_t}{\alpha_0} \end{cases} \tag{B.5}
$$

George Athanasopoulos and Ashton de Silva. Multivariate exponential smoothing for forecasting tourist arrivals. *Journal of Travel Research*, 51 (5):640–652, 2012. doi: 10.1177/0047287511434115. URL `http://jtr.sagepub.com/content/51/5/640.abstract`.

George Athanasopoulos, Rob J. Hyndman, Haiyan Song, and Doris C. Wu. The tourism forecasting competition. *International Journal of Forecasting*, 27(3):822 – 844, 2011. ISSN 0169-2070. doi: http://dx.doi.org/10.1016/j.ijforecast.2010.04.009. URL `http://www.sciencedirect.com/science/article/pii/S016920701000107X`. Special Section 1: Forecasting with Artificial Neural Networks and Computational Intelligence Special Section 2: Tourism Forecasting.

Baki Billah, Maxwell L. King, Ralph D. Snyder, and Anne B. Koehler. Exponential smoothing model selection for forecasting. *International Journal of Forecasting*, 22(2):239–247, April 2006. ISSN 0169-2070. URL `http://www.sciencedirect.com/science/article/pii/S016920700500107X`.

J. L. Brenner, D. A. D'Esopo, and A. G. Fowler. Difference equations in forecasting formulas. *Management Science*, 15(3):141–159, November 1968. ISSN 00251909. URL `http://www.jstor.org/stable/2628786`.

R.G. Brown. Exponential smoothing for predicting demand, 1956.

Alysha M. De Livera, Rob J. Hyndman, and Ralph D. Snyder. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496):1513–1527,

2011. doi: 10.1198/jasa.2011.tm09771. URL http://dx.doi.org/10.1198/jasa.2011.tm09771.

Robert Fildes, Michle Hibon, Spyros Makridakis, and Nigel Meade. Generalising about univariate forecasting methods: Further empirical evidence. *International Journal of Forecasting*, 14(3):339–358, September 1998. ISSN 0169-2070. URL http://www.sciencedirect.com/science/article/pii/S0169207098000090.

Everette S. Gardner. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1):1–28, 1985. ISSN 1099-131X. doi: 10.1002/for.3980040103. URL http://dx.doi.org/10.1002/for.3980040103.

Jr. Gardner, Everette S. and Ed. McKenzie. Forecasting trends in time series. *Management Science*, 31(10):1237–1246, October 1985. ISSN 00251909. URL http://www.jstor.org/stable/2631713.

Everette S. Gardner Jr. and Joaquin Diaz-Saiz. Exponential smoothing in the telecommunications data. *International Journal of Forecasting*, 24(1):170–174, January 2008. ISSN 0169-2070. URL http://www.sciencedirect.com/science/article/pii/S0169207007000465.

Phillip G. Gould, Anne B. Koehler, J. Keith Ord, Ralph D. Snyder, Rob J. Hyndman, and Farshid Vahid-Araghi. Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research*, 191(1):207–222, November 2008. ISSN 0377-2217. URL http://www.sciencedirect.com/science/article/pii/S0377221707008740.

Charles C. Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1):5–10, January 2004. ISSN 0169-2070. URL http://www.sciencedirect.com/science/article/pii/S0169207003001134.

Rob J. Hyndman and Anne B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, October 2006. ISSN 0169-2070. URL http://www.sciencedirect.com/science/article/pii/S0169207006000239.

Rob J Hyndman, Anne B Koehler, Ralph D Snyder, and Simone Grose. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18(3):439–454, July

2002. ISSN 0169-2070. URL http://www.sciencedirect.com/science/article/pii/S0169207001001108.

Rob J. Hyndman, Anne B. Koehler, J. Keith Ord, and Ralph D. Snyder. *Forecasting With Exponential Smoothing: The State Space Approach*. Springer-Verlag Berlin Heidelberg, 2008a. URL http://link.springer.com/book/10.1007%2F978-3-540-71918-2.

RobJ. Hyndman, Muhammad Akram, and BlythC. Archibald. The admissible parameter space for exponential smoothing models. *Annals of the Institute of Statistical Mathematics*, 60(2):407–426, 2008b. ISSN 0020-3157. doi: 10.1007/s10463-006-0109-x. URL http://dx.doi.org/10.1007/s10463-006-0109-x.

Victor Richmond R. Jose and Robert L. Winkler. Simple robust averages of forecasts: Some empirical results. *International Journal of Forecasting*, 24(1):163–169, January 2008. ISSN 0169-2070. URL http://www.sciencedirect.com/science/article/pii/S0169207007000878.

Anne B. Koehler, Ralph D. Snyder, J. Keith Ord, and Adrian Beaumont. A study of outliers in the exponential smoothing approach to forecasting. *International Journal of Forecasting*, 28(2):477 – 484, 2012. ISSN 0169-2070. doi: http://dx.doi.org/10.1016/j.ijforecast.2011.05.001. URL http://www.sciencedirect.com/science/article/pii/S0169207011000902.

Stephan Kolassa. Combining exponential smoothing forecasts using akaike weights. *International Journal of Forecasting*, 27(2):238 – 251, 2011. ISSN 0169-2070. doi: http://dx.doi.org/10.1016/j.ijforecast.2010.04.006. URL http://www.sciencedirect.com/science/article/pii/S0169207010001032.

Nikolaos Kourentzes, Fotios Petropoulos, and Juan R. Trapero. Improving forecasting by estimating time series structural components across multiple frequencies. *International Journal of Forecasting*, 30(2):291 – 302, 2014. ISSN 0169-2070. doi: http://dx.doi.org/10.1016/j.ijforecast.2013.09.006. URL http://www.sciencedirect.com/science/article/pii/S0169207013001477.

Andr Luis Santiago Maia and Francisco de A.T. de Carvalho. Holts exponential smoothing and neural network models for forecasting interval-valued time series. *International Journal of Forecasting*, 27(3):740 – 759,

2011. ISSN 0169-2070. doi: http://dx.doi.org/10.1016/j.ijforecast.2010.02.012. URL http://www.sciencedirect.com/science/article/pii/S0169207010000506. Special Section 1: Forecasting with Artificial Neural Networks and Computational Intelligence Special Section 2: Tourism Forecasting.

Spyros Makridakis and Michle Hibon. The m3-competition: Results, conclusions and implications. *International Journal of Forecasting*, 16(4):451–476, October 2000. ISSN 0169-2070. URL http://www.sciencedirect.com/science/article/pii/S0169207000000571.

Keith Ord, Anne B. Koehler, and Ralph D. Snyder. Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of the American Statistical Association*, 92(440):1621–1629, December 1997. ISSN 01621459. URL http://www.jstor.org/stable/2965433.

C. Carl Pegels. Exponential forecasting: Some new variations. *Management Science*, 15(5):pp. 311–315, 1969. ISSN 00251909. URL http://www.jstor.org/stable/2628137.

S. Svetunkov. *Complex-Valued Modeling in Economics and Finance*. SpringerLink : Bücher. Springer, 2012. ISBN 9781461458760. URL http://books.google.co.uk/books?id=XNqvi56BT3IC.

James W. Taylor. Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19(4):715–725, October 2003. ISSN 0169-2070. URL http://www.sciencedirect.com/science/article/pii/S0169207003000037.

James W. Taylor and Ralph D. Snyder. Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing. *Omega*, 40(6):748 – 757, 2012. ISSN 0305-0483. doi: http://dx.doi.org/10.1016/j.omega.2010.03.004. URL http://www.sciencedirect.com/science/article/pii/S0305048310000307. Special Issue on Forecasting in Management Science.

Ju-Jie Wang, Jian-Zhou Wang, Zhe-George Zhang, and Shu-Po Guo. Stock index forecasting based on a hybrid model. *Omega*, 40(6):758 – 766, 2012. ISSN 0305-0483. doi: http://dx.doi.org/10.1016/j.omega.2011.07.008. URL http://www.sciencedirect.com/science/article/pii/

S0305048311001435. Special Issue on Forecasting in Management Science.