1 Time series

Forecasting of time series with (S)AR(I)MA models is a well-established concept that has been studied thoroughly for many decades and provides good forecasting accuracy (Arlt et al., 2017; Khandelwal et al., 2015).It has found application in many domains such as economy (MEHR QUELLEN),

The assumption ARIMA models are based on is that the values of a target variable are generated by a linear combination of past values of the same variable and white noise (Khandelwal et al., 2015), thus making it a stochastic process, "i.e. an ordered sequence of random variables" (Andreoni and Postorino, 2006), with data entries at equally distant intervals (Hunt, 2003).
A mathematical assumption underlying times series processes is stationarity. (Weak) Stationarity is given when mean and covariance are independent oft time t, and the relationship between two values at time points t and $t + i$ is the same as the relationship between two values at time points s and $s + i$, i.e. independent of the exact position in the time series, but provided the distance between any two values is the same (Vogel, 2015).
If $\mu(t) \neq \mu$, i.e. the mean is not independent of time, the assumption of stationarity is violated; there is a so-called trend. In the case of a non-stationary time series, stationarity can be obtained by differencing (Andreoni and Postorino, 2006; Hunt, 2003) First-order differencing is denoted as:

$$BY_t = Y_t - Y_{t-1} \tag{1}$$

Autoregressive (AR) processes are processes where a value of a variable at t depends on weighted previous values of the variable itself plus a white noise term $\epsilon \sim WN(0, \sigma_\epsilon^2)$. An *AR(p)* process of order p has the form:

$$Y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \epsilon_t \tag{2}$$

Moving Average (MA) processes of order q are denoted like this:

$$Y_t = \mu + \epsilon_t - \phi_1 \epsilon_{t-1} - \phi_2 \epsilon_{t-2} - ... - +\phi_q \epsilon_{t-q} \tag{3}$$

This means the value of a target value at t depends on a white noise and previous white noise weighted by ϕ. Note that μ may be 0 and MA-processes are stationary.

ERKLÄRE differencing, transformation, seasonality

However, the process underlying time series data may change over time - it is subject to uncertainty (Adhikari, 2015). A time series model may be biased or overfitted as well as its parameters misspecified.

1.1 AR(I)MA models

As the name suggests, AR(I)MA(p,d,q) - auto-regressive integrated moving average - models model time series data with an AR and an MA component, and trend in data through differencing (which is the "I-part").
The parameters p, d, and q respectively denote the order of the AR component, the degree of differencing and the MA component (Zhao et al., 2018). If there is also a seasonal component - thus a SARIMA model is to be fitted - there are additional parameters P, D, Q referring to the seasonal orders (or degrees) of AR, differencing and MA.

The parameters p,q,P and Q may be identified using an Information Criterion suchs as BIC or AIC. (FORMEL falls nicht schon geschehen oder Referenz auf Formel) In order to appropriately model time series data, (Box and Jenkins, 1976) proposed a method to identify suitable parameters - AR, MA and differencing - of an ARIMA model. It consists of the following four steps:
1. Visual identification of model parameter through ACF for MA- and PACF for AR-parameters.

References

Adhikari, R. (2015). A neural network based linear ensemble framework for time series forecasting. *Neurocomputing*, 157:231–242.

Andreoni, A. and Postorino, M. N. (2006). A multivariate arima model to forecast air transport demand.

Arlt, J., Trcka, P., and Arltová, M. (2017). The problem of the sarima model selection for the forecasting purpose. *Statistika*, 97:25–32.

Box, G. E. P. and Jenkins, G. M. (1976). *Time series analysis: Forecasting and control.* Holden-Day series in time series analysis. Holden-Day, San Francisco, Calif., rev. ed. edition.

Hunt, Ü. (2003). Forecasting of railway freight volume: Approach of estonian railway to arise efficiency. *Transport*, 18.

Khandelwal, I., Adhikari, R., and Verma, G. (2015). Time series forecasting using hybrid arima and ann models based on dwt decomposition. *Procedia Computer Science*, 48:173–179.

Vogel, J. (2015). *Prognose von Zeitreihen: Eine Einführung für Wirtschaftswissenschaftler*. Springer Gabler, Wiesbaden.

Zhao, J., Cai, J., and Zheng, W. (07022018). Research on railway freight volume prediction based on arima model. In Wang, X., Zhang, Y., Yang, D., and You, Z., editors, *CICTP 2018*, pages 428–437, Reston, VA. American Society of Civil Engineers.