SARIMA MODELS WITH MULTIPLE SEASONALITY

Luisa Bisaglia ¹, Francesco Lisi ¹

¹ Department of Statistical Sciences, University of Padua, Italy, (e-mail: luisa.bisaglia@unipd.it, francesco.lisi@unipd.it)

Abstract: SARIMA models and exponential smoothing methods are classical approaches to account seasonal dynamics. However, they typically allow to model just one periodic component, while many empirical time series data show multiple seasonality, possibly interlacing together. To face this case, different decomposition models have been proposed in literature, while SARIMA models have been quite neglected. To fill the gap, in this work we suggest a suitable specification of the SARIMA model, called mSARIMA, able to account multiple seasonality. To study its performance, we compare it with two popular seasonal-trend decomposition approaches, namely the TBATS and MSTL models. A simulation exercise shows that mSARIMA models are more effective in describing the different seasonal components.

Keywords: Time series, Multiple seasonality, mSARIMA, seasonal-trend decomposition models.

1 Introduction

Typically, a SARIMA model allows to account just one periodic component. When multiple cycles arise, REG-SARIMA or SARIMAX models are often considered. In this case, only one seasonal component is treated as stochastic while the other ones are deterministically described using dummy variables, trigonometric functions or spline functions. Alternatively, a large body of literature focuses on time series decomposition techniques such as the Seasonal-Trend decomposition by regression (STR, Documentov & Hyndman, 2022), the Trigonometric Exponential Smoothing State Space model (TBATS, A.M. et al., 2011) and the Multiple Seasonal Trend decomposition using Loess (MSTL, Bandara et al., n.d.). The present work aims at showing that a suitable specification of SARIMA models allows to consider multiple seasonal components and effectively estimate them. We denote this class of models as Multiple Seasonality ARIMA models, briefly mSARIMA. We note, however, that these models are nothing but suitably constrained specifications of the general SARIMA models from which they inherit all properties.
2 The Multiple Seasonality ARIMA model

Let $\varepsilon_t$ be a zero-mean white noise process with variance $\sigma^2$. We denote by mSARMA$(p, q) \times (P_1, Q_1)_{S_1} \times \ldots \times (P_m, Q_m)_{S_m}$ the stationary process $Y_t$

$$\phi(B) \prod_{i=1}^{m} \Phi(B^{S_i}) Y_t = \theta(B) \prod_{i=1}^{m} \Theta(B^{S_i}) \varepsilon_t, \quad (1)$$

where $\phi(B)$ and $\theta(B)$ are the usual AR and MA polynomials in $B$ of degrees, respectively, $p$ and $q$. $\Phi(B^{S_i}) = (1 - \Phi_{i,1}B^{S_i} - \ldots - \Phi_{i,P}B^{P_iS_i})$ is the i-th seasonal AR polynomial of degree $P_i$ in $B^{S_i}$ $(i=1,...,m)$ while $\Theta(B^{S_i}) = (1 - \Theta_{i,1}B^{S_i} - \ldots - \Theta_{i,Q}B^{Q_iS_i})$ is the corresponding MA seasonal polynomial of degree $Q_i$. The polynomials $\phi(B)$ and $\theta(B)$ describe the non-periodic serial dependence of the time series, while the polynomials $\Phi(B^{S_i})$ and $\Theta(B^{S_i})$ model the periodic correlation for the $m$ seasonal components of period $S_i$, $(i = 1, ..., m)$. Just to give an example, the $2 - SARIMA(1, 0) \times (1, 0)_4 \times (1, 0)_7$ is given by:

$$Y_t = \phi_1 Y_{t-1} + \phi_{1,1} Y_{t-4} - \phi_1 \Phi_{1,1} Y_{t-5} + \Phi_{2,1} Y_{t-7} - \phi_1 \Phi_{2,1} Y_{t-8}$$

$$- \Phi_{1,1} \Phi_{2,1} Y_{t-11} + \Phi_{1,1} \Phi_{2,1} Y_{t-12} + \varepsilon_t.$$

It is clear that, although 7 different lags are involved, there are only 3 parameters to be estimated and that it can be also thought as a particular constrained (S)ARMA model. This implies that the stationary conditions for model $(1)$ are those of a standard (S)ARMA, once the constraints are considered. The same holds also for the invertibility conditions. Model $(1)$ can be straightforwardly generalized to the non-stationary case by including suitable unit root polynomials. We can define the non-stationary mSARIMA$(p, d, q) \times (P_1, D_1, Q_1)_{S_1} \times \ldots \times (P_m, D_m, Q_m)_{S_m}$ process, $Y_t$, by:

$$\phi(B)(1-B)^d \prod_{i=1}^{m} \Phi(B^{S_i})(1-B^{S_i})^d Y_t = \theta(B) \prod_{i=1}^{m} \Theta(B^{S_i}) \varepsilon_t \quad (2)$$

where $(1-B^{S_i})^d = (Y_t - Y_{t-d})^d$ is the $d-$seasonal difference of $Y_t$. As for standard ARIMA models, the process $X_t = (1-B)^d \prod_{i=1}^{m} (1-B^{S_i})^d Y_t$ is a stationary mSARMA. For building an mSARIMA model, the classical Box-Jenkins approach (Box & Jenkins, 1976) can be applied, with some simple and intuitive modifications needed to account the presence of more than one seasonal component. For the estimation step maximum likelihood methods can still be used taking care that the mSARMA model is a constrained one. This implies that the user has to write the specific likelihood to be maximized.
In this section we compare the mSARMA with two popular seasonal-trend decomposition models, namely the TBATS and MSTL models. We analyse, through a simple Monte Carlo exercise, their ability in whitening the residuals’ autocorrelation function. To this end we simulated 500 independent realizations of length $n = 200$ and 500 from the three following mSARMA specifications. All models include two seasonal components: of periods 4 and 7, the first two, and of periods 4 and 12, the third one. In the last case, cycles overlap.

1. Model 1: $2 - \text{SARIMA}(1,0,0) \times (1,0,0)_4 \times (1,0,0)_7$, with $\phi_1 = 0.4$, $\Phi_{1,1} = 0.3$, $\Phi_{2,1} = 0.35$ and $\sigma^2 = 1$;
2. Model 2: $2 - \text{SARIMA}(1,0,0) \times (1,0,0)_4 \times (2,0,0)_7$, with $\phi_1 = 0.4$, $\Phi_{1,1} = 0.3$, $\Phi_{2,1} = 0.25$, $\Phi_{2,2} = 0.35$ and $\sigma^2 = 1$;
3. Model 3: $2 - \text{SARIMA}(1,0,0) \times (1,0,0)_4 \times (1,0,0)_12$, with $\phi_1 = 0.4$, $\Phi_{1,1} = 0.3$, $\Phi_{2,1} = 0.4$ and $\sigma^2 = 1$.

For each series we estimated an mSARMA model, a MSTL model and a TBATS model and we analyzed the residuals time series to check if the multi-seasonal serial dependence has been completely accounted. To assess the residuals’ appropriateness we propose a modification of the Pierce test (Pierce, 1978) able to account for multiple seasonality. When 2 periodic components are present, the hypothesis system to be verified is $H_0: \rho_{S_1} = \cdots = \rho_{k,S_1} = \rho_{S_2} = \cdots = \rho_{k,S_2} = 0$ against $H_1: \bar{H}_0$. When applied to the residuals of a model, it tests the model’s adequacy in describing both seasonal components. The test statistics is:

$$mQ_{S_1,S_2}(k) = n \cdot (n+2) \left( \sum_{j=1}^{k} \frac{1}{n-j \cdot S_1} \rho_{j,S_1}^2 + \sum_{j=1}^{k} \frac{1}{n-j \cdot S_2} \rho_{j,S_2}^2 \right)$$ (3)

where $\rho_j$ is the correlation coefficient at lag $j$ of the considered series. Under the null hypothesis of no seasonal autocorrelation, it follows a $\chi^2_{2k-\text{par}}$ distribution, where $\text{par}$ is the number of estimated parameters.

In our exercise we computed $mQ_{4,7}(5)$ (for the first two models) and $mQ_{4,12}(5)$ (for the third model) on the residual series of our three models, i.e. mSARIMA, MSTL and TBATS, and we counted the percentage of times the null hypothesis is not rejected at a significance level of 5%. Results are given in Table ??: it is clear that the mSARIMA model produce (seasonally) uncorrelated residuals most of times, while the other two models do not, particularly when the sample
size increases. One could argue that it is not fair considering time series generated only by mSARMA models: we agree with this point. These results are very preliminary and other generating processes must be taken into account, in particular processes with deterministic components. However, we showed that, when multiple seasonality is generated by an ARMA process, MSTL and TBATS model are not appropriate to describe this dynamics.

Table 1. Percentage of times, the Pierce test for multiple seasonality does not reject the hypothesis of seasonal uncorrelation in the residuals of the mSARMA, MSTL and TBATS models. The level of the test is $\alpha = 5\%$.

<table>
<thead>
<tr>
<th>Model</th>
<th>$n = 200$</th>
<th>$n = 500$</th>
<th>$n = 200$</th>
<th>$n = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>mSARIMA</td>
<td>93.2</td>
<td>92.7</td>
<td>92.5</td>
<td>91.6</td>
</tr>
<tr>
<td>MSTL</td>
<td>2.3</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>TBATS</td>
<td>42.3</td>
<td>17.3</td>
<td>6.3</td>
<td>0.0</td>
</tr>
</tbody>
</table>

References


BANDARA, K., HYNDMAN, R.J., & C., BERGMEIR. MSTL: A Seasonal-Trend Decomposition Algorithm for Time Series with Multiple Seasonal Patterns. *International J Operational Research*.

