Thailand Statistician
October 2026; 24(4): 841-854
http://statassoc.or.th
Contributed paper

Bayesian Prediction of Double Seasonal ARMA Processes

Ayman A. Amin*
Department of Mathematics, College of Arts and Sciences, Najran University, Najran, Saudi Arabia
Department of Statistics, Mathematics, and Insurance, Faculty of Commerce, Menoufia University, Menoufia, Egypt
*Corresponding author; e-mail: aaabdellah@nu.edu.sa

Received: 2 March 2023
Revised: 14 June 2026
Accepted: 15 June 2026

Abstract

Recent developments in timeseries analysis extended autoregressive moving average (ARMA) models to accommodate datasets exhibiting two seasonal cycles, leading to the formulation of double seasonal ARMA (DSARMA) models. While Bayesian estimation methods have been explored for these models, Bayesian prediction has not yet been formally developed. In the current work, we advance the existing literature by establishing a Bayesian predictive framework for DSARMA models. Our approach assumes normally distributed error terms and employs both normalgamma and Jeffreys priors on the model parameters. Under these assumptions, we derive closedform predictive distributions for future, unobserved values, showing that they follow a univariate-t distribution. This result enables efficient multistepahead forecasting. To assess the forecasting accuracy of the current proposed Bayesian prediction methodology, we conduct a comprehensive simulation study and apply the framework to forecasting hourly electricity load across several European countries.

Keywords: Posterior and predictive analyses, DSARMA models, multiple seasonality, hourly electricity load.

1. Introduction

Autoregressive moving-average (ARMA) models have many applications in different fields (Hyndman and Athanasopoulos, 2018; Abonazel and Abd-Elftah, 2019; Awwad et al., 2021). However; in many practical settings, highfrequency timeseries data are routinely collectedfor example, halfhourly website traffic or hourly electricity demand. Such series often display several seasonal patterns simultaneously, including intraday, intraweek, and intrayear cycles. Because classical seasonal autoregressive moving-average (SARMA) models are designed to capture only a single seasonal component, they are insufficient for representing these richer structures. Consequently, several authors have proposed extensions capable of modeling multiple seasonalities; see, for instance, the nonBayesian contributions of De Livera et al. (2011), Ryu et al. (2017), Taylor and McSharry (2017), Lago et al. (2018), Sulandari et al. (2021), and the Bayesian work of Amin (2018).

It is well known that SARMA models involve multiplicative interactions between nonseasonal and seasonal parameters, rendering the model nonlinear in its coefficients. This nonlinearity leads to an analytically intractable likelihood function, which in turn complicates both posterior inference and predictive analysis (Amin, 2009). To address these challenges, researchers have proposed various strategies for approximating SARMA posterior and predictive distributions. These strategies generally fall into two categories: Markov-chain Monte-Carlo (MCMC)based methods and analytical

approximations. Analytical approaches modify the SARMA posterior and predictive distributions to obtain standard or at least closedform expressions; see Shaarawy and Ali (2003), Amin (2018), Amin (2019c). In contrast, MCMCbased methods approximate the intractable joint-posterior and predictive distributions by simulating from the corresponding conditional distributions; see Vermaak et al. (1998), Ismail and Amin (2014).

Bayesian methods for SARMA models with a single seasonal component are well developed and supported by a substantial body of work. Using MCMCbased approximations, Barnett et al. (1996, 1997) conducted Bayesian analyses of singleseasonal AR (SAR) and ARMA models, while Vermaak et al. (1998) introduced a MetropoliswithinGibbs algorithm to facilitate Bayesian inference for SAR models. On the analytical side, Shaarawy and Ali (2003) proposed an identification procedure for SAR models by deriving an approximate posterior mass function for the model orders. Employing Gibbs sampling as an MCMC technique, Ismail and Amin (2014) developed Bayesian estimation methods for SARMA models, and more recently, Amin (2019a) extended this line of work by presenting Bayesian estimation and forecasting for these models.

Bayesian methodology for double SARMA (i.e. DSARMA) modelsdesigned for timeseries exhibiting two seasonal cyclesremains relatively underdeveloped. Amin and Ismail (2015) initiated this line of research by employing a Gibbs sampler to derive Bayesian estimators for double SAR (i.e. DSAR) models. This work was subsequently expanded by Amin (2017b,a), who developed Bayesian estimation procedures for double seasonal moving-average (DSMA) and DSARMA models, respectively. Using analytical approximation techniques, Amin (2018, 2019c) first proposed Bayesian estimation methods for DSARMA models and later introduced a Bayesian identification approach for DSAR structures. Continuing along this trajectory, Amin (2020, 2024) applied Gibbs sampling to deliver Bayesian estimation, prediction, and a comprehensive analysis of DSAR models.

In practical settings, it is not sufficient to merely estimate timeseries models; reliable prediction is equally essential. To contribute to the growing literature on Bayesian methods for timeseries with two seasonal cycles, this paper develops a Bayesian predictive framework for DSARMA models, allowing practitioners to generate accurate multistepahead forecasts.

Our predictive approach assumes normally distributed DSARMA error terms and adopts normal-gamma and Jeffreys priors for the model parameters. We begin by using nonlinear least squares (NLS) to estimate the unobserved lagged errors arising from the movingaverage component. Based on these estimates, we approximate the marginal-posterior distributions of the DSARMA coefficients and the precision parameter by multivariate-t and gamma distributions, respectively. Under the assumption of normal errors and the specified priors, we derive closedform predictive distributions for future observationsboth onestepahead and multistepaheadwhich take the form of univariate-t distributions. These standard predictive forms facilitate accurate longhorizon forecasting. Given the mathematical complexity of DSARMA models and the typically large size of datasets exhibiting double seasonality, implementing the proposed Bayesian framework poses significant programming and computational challenges. Addressing these challenges is a central objective of the present work.

The remainder of the current manuscript is structured as follows. Section 2 introduces the DSARMA model, and Section 3 develops its posterior analysis. Section 4 presents the proposed Bayesian predictive framework for DSARMA models. In Section 5, we report the results of a simulation study and then illustrate the methodology using hourly electricityload data from several European countries. Section 6 concludes the paper.

2. Double Seasonal ARMA (DSARMA) Models

Let $\{z_t\}$ be an observed time-series that has zero-mean and exhibits two seasonal patterns with s_1 and s_2 as seasonal periods. This time series can be modeled by the DSARMA model with order p, q, P_1, Q_1, P_2 and Q_2, which is shortened as DSARMA$(p,q)(P_1,Q_1)_{s_1}(P_2,Q_2)_{s_2}$, and written as:

$$\boldsymbol{\phi}_p(B)\boldsymbol{\varphi}_{P_1}(B^{s_1})\boldsymbol{\Phi}_{P_2}(B^{s_2})z_t = \boldsymbol{\theta}_q(B)\boldsymbol{\vartheta}_{Q_1}(B^{s_1})\boldsymbol{\Theta}_{Q_2}(B^{s_2})u_t \tag{1}$$

where $\{u_t\}$'s are unobserved DSARMA model errors that we assume to be independent and normally distributed with zero-mean and precision τ, and B is a backshift operator given as $B^k z_t = z_{t-k}$. $\boldsymbol{\phi}_p(B) = \left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)$ and $\boldsymbol{\theta}_q(B) = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right)$ are the nonseasonal autoregressive and moving-average operators with order p and q respectively. As expected, two seasonal autoregressive operators are included in this model:

$\boldsymbol{\varphi}_{P_1}(B^{s_1}) = \left(1 - \varphi_1 B^{s_1} - \varphi_2 B^{2s_1} - \cdots - \varphi_{P_1} B^{P_1 s_1}\right)$ with order P_1 and

$\boldsymbol{\Phi}_{P_2}(B^{s_2}) = \left(1 - \Phi_1 B^{s_2} - \Phi_2 B^{2s_2} - \cdots - \Phi_{P_2} B^{P_2 s_2}\right)$ with order P_2.

Also, two seasonal moving-average operators are included:

$\boldsymbol{\vartheta}_{Q_1}(B^{s_1}) = \left(1 - \vartheta_1 B^{s_1} - \vartheta_2 B^{2s_1} - \cdots - \vartheta_{Q_1} B^{Q_1 s_1}\right)$ with order Q_1 and

$\boldsymbol{\Theta}_{Q_2}(B^{s_2}) = \left(1 - \Theta_1 B^{s_2} - \Theta_2 B^{2s_2} - \cdots - \Theta_{Q_2} B^{Q_2 s_2}\right)$ with order Q_2.

The nonseasonal and seasonal autoregressive coefficients vectors are $\boldsymbol{\phi} = (\phi_1, \phi_2, \cdots, \phi_p)^T$, $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \cdots, \varphi_{P_1})^T$, $\boldsymbol{\Phi} = (\Phi_1, \Phi_2, \cdots, \Phi_{P_2})^T$ respectively, and the nonseasonal and seasonal moving-average coefficients vectors are $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_q)^T$, $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \cdots, \vartheta_{Q_1})^T$, $\boldsymbol{\Theta} = (\Theta_1, \Theta_2, \cdots, \Theta_{Q_2})^T$ respectively.

It is possible to show mathematically the stationarity and invertibility conditions for the DSARMA model (1), i.e. DSARMA model is stationary if the roots of $\boldsymbol{\phi}_p(B) = 0$, $\boldsymbol{\varphi}_{P_1}(B^{s_1}) = 0$ and $\boldsymbol{\Phi}_{P_2}(B^{s_2}) = 0$ are located outside the unit circle, and when all the roots of $\boldsymbol{\theta}_q(B) = 0$, $\boldsymbol{\vartheta}_{Q_1}(B^{s_1}) = 0$ and $\boldsymbol{\Theta}_{Q_1}(B^{s_1}) = 0$ fall outside the unit circle the DSARMA process is invertible. Stationarity and invertibility for time-series models are introduced and discussed in details in time-series textbooks, see, for example, Box et al. (2015).

We can simplify the form of the DSARMA model by expanding its compact structure given in eqn (1) to be in the form of using the summation notation as:

$$
\begin{aligned}
z_t \;=\; & \sum_{i=1}^{p} \phi_i z_{t-i} + \sum_{j=1}^{P_1} \varphi_j z_{t-js_1} + \sum_{m=1}^{P_2} \Phi_m z_{t-ms_2} - \sum_{i=1}^{p}\sum_{j=1}^{P_1} \phi_i \varphi_j z_{t-i-js_1} - \\
& \sum_{i=1}^{p}\sum_{m=1}^{P_2} \phi_i \Phi_m z_{t-i-ms_2} - \sum_{j=1}^{P_1}\sum_{m=1}^{P_2} \varphi_j \Phi_m z_{t-js_1-ms_2} + \\
& \sum_{i=1}^{p}\sum_{j=1}^{P_1}\sum_{m=1}^{P_2} \phi_i \varphi_j \Phi_m z_{t-i-js_1-ms_2} - \sum_{i=1}^{q} \theta_i u_{t-i} - \sum_{j=1}^{Q_1} \vartheta_j u_{t-js_1} - \\
& \sum_{m=1}^{Q_2} \Theta_m u_{t-ms_2} + \sum_{i=1}^{q}\sum_{j=1}^{Q_1} \theta_i \vartheta_j u_{t-i-js_1} + \sum_{i=1}^{q}\sum_{m=1}^{Q_2} \theta_i \Theta_m u_{t-i-ms_2} + \\
& \sum_{j=1}^{Q_1}\sum_{m=1}^{Q_2} \vartheta_j \Theta_m u_{t-js_1-ms_2} - \sum_{i=1}^{q}\sum_{j=1}^{Q_1}\sum_{m=1}^{Q_2} \theta_i \vartheta_j \Theta_m u_{t-i-js_1-ms_2} + u_t. \qquad (2)
\end{aligned}
$$

Based on the structure of DSARMA model (2), the lagged errors are unknown and therefore the Bayesian analysis of these models is very complicated and needs numerical analysis to be conducted. However, as it is done with the usual ARMA and SARMA models, the Bayesian analysis can be simplified by estimating the unknown errors. Different methods are proposed in literature to estimate the unknown errors, for instance these unknown errors can be estimated recursively as:

$$
\begin{aligned}
\hat{u}_t \;=\; & z_t - \sum_{i=1}^{p} \hat{\phi}_i z_{t-i} - \sum_{j=1}^{P_1} \hat{\varphi}_j z_{t-js_1} - \sum_{m=1}^{P_2} \hat{\Phi}_m z_{t-ms_2} + \sum_{i=1}^{p}\sum_{j=1}^{P_1} \hat{\phi}_i \hat{\varphi}_j z_{t-i-js_1} \\
& + \sum_{i=1}^{p}\sum_{m=1}^{P_2} \hat{\phi}_i \hat{\Phi}_m z_{t-i-ms_2} + \sum_{j=1}^{P_1}\sum_{m=1}^{P_2} \hat{\varphi}_j \hat{\Phi}_m z_{t-js_1-ms_2}
\end{aligned}
$$

$$
\begin{aligned}
&- \sum_{i=1}^{p}\sum_{j=1}^{P_1}\sum_{m=1}^{P_2} \hat{\phi}_i\hat{\varphi}_j\hat{\Phi}_m z_{t-i-js_1-ms_2} + \sum_{i=1}^{q}\theta_i\hat{u}_{t-i} + \sum_{j=1}^{Q_1}\hat{\vartheta}_j\hat{u}_{t-js_1} \\
&+ \sum_{m=1}^{Q_2}\hat{\Theta}_m\hat{u}_{t-ms_2} - \sum_{i=1}^{q}\sum_{j=1}^{Q_1}\theta_i\hat{\vartheta}_j\hat{u}_{t-i-js_1} - \sum_{i=1}^{q}\sum_{m=1}^{Q_2}\theta_i\hat{\Theta}_m\hat{u}_{t-i-ms_2} \\
&- \sum_{j=1}^{Q_1}\sum_{m=1}^{Q_2}\hat{\vartheta}_j\hat{\Theta}_m\hat{u}_{t-js_1-ms_2} + \sum_{i=1}^{q}\sum_{j=1}^{Q_1}\sum_{m=1}^{Q_2}\theta_i\hat{\vartheta}_j\hat{\Theta}_m\hat{u}_{t-i-js_1-ms_2}, \qquad (3)
\end{aligned}
$$

where, $\hat{\phi}_i \in R^p, \hat{\varphi}_j \in R^{P_1}, \hat{\Phi}_j \in R^{P_2}, \hat{\theta}_i \in R^q, \hat{\vartheta}_j \in R^{Q_1}$ and $\hat{\Theta}_j \in R^{Q_2}$ are the nonlinear least squares (NLS) estimates that we get by minimizing the sum of square errors $\sum u_t^2 = SS(\boldsymbol{\phi}, \boldsymbol{\varphi}, \boldsymbol{\Phi}, \boldsymbol{\theta}, \boldsymbol{\vartheta}, \boldsymbol{\Theta})$ with respect to $\boldsymbol{\phi}, \boldsymbol{\varphi}, \boldsymbol{\Phi}, \boldsymbol{\theta}, \boldsymbol{\vartheta}$, and $\boldsymbol{\Theta}$.

Once we estimate the model errors, in order to make the Bayesian analysis much easier for these models, it is preferred to write the DSARMA model (2) in the matrix-form as:

$$
\mathbf{z} \;=\; W\boldsymbol{\Gamma} + \mathbf{u}, \qquad (4)
$$

where, $\mathbf{z}$ is the vector of observed time-series, i.e. $\mathbf{z} = (z_1, z_2, \cdots, z_n)^T$; and $\mathbf{u}$ is the vector of unobserved model errors, i.e. $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$, and W is a design matrix of size $n \times p^\star$, with $p^\star = (1+p)(1+P_1)(1+P_2)(1+P_3) + (1+q)(1+Q_1)(1+Q_2) - 2$. The t^{th} row of the design matrix W is given as:

$$
\begin{aligned}
W_t \;=\; &(z_{t-1}, \ldots, z_{t-p}, z_{t-s_1}, z_{t-s_1-1}, \ldots, z_{t-s_1-p}, \ldots\ldots, z_{t-P_1s_1}, z_{t-P_1s_1-1}, \ldots, z_{t-P_1s_1-p}, \\
&z_{t-s_2}, z_{t-s_2-1}, \ldots, z_{t-s_2-p}, z_{t-s_1-s_2}, z_{t-s_1-s_2-1}, \ldots, z_{t-s_1-s_2-p}, \ldots\ldots, z_{t-P_1s_1-s_2}, \\
&z_{t-P_1s_1-s_2-1}, \ldots, z_{t-P_1s_1-s_2-p}, \ldots\ldots\ldots, z_{t-P_2s_2}, z_{t-P_2s_2-1}, \ldots, z_{t-P_2s_2-p}, z_{t-s_1-P_2s_2}, \\
&z_{t-s_1-P_2s_2-1}, \ldots, z_{t-s_1-P_2s_2-p}, \ldots\ldots, z_{t-P_1s_1-P_2s_2}, z_{t-P_1s_1-P_2s_2-1}, \ldots, z_{t-P_1s_1-P_2s_2-p}, \\
&\hat{u}_{t-1}, \ldots, \hat{u}_{t-q}, \hat{u}_{t-s_1}, \hat{u}_{t-s_1-1}, \ldots, \hat{u}_{t-s_1-q}, \ldots\ldots, \hat{u}_{t-Q_1s_1}, \hat{u}_{t-Q_1s_1-1}, \ldots, \hat{u}_{t-Q_1s_1-q}, \\
&\hat{u}_{t-s_2}, \hat{u}_{t-s_2-1}, \ldots, \hat{u}_{t-s_2-q}, \hat{u}_{t-s_1-s_2}, \hat{u}_{t-s_1-s_2-1}, \ldots, \hat{u}_{t-s_1-s_2-q}, \ldots\ldots, \hat{u}_{t-Q_1s_1-s_2}, \\
&\hat{u}_{t-Q_1s_1-s_2-1}, dots, \hat{u}_{t-Q_1s_1-s_2-q}, \ldots\ldots\ldots, \hat{u}_{t-Q_2s_2}, \hat{u}_{t-Q_2s_2-1}, \ldots, \hat{u}_{t-Q_2s_2-q}, \\
&\hat{u}_{t-s_1-Q_2s_2}, \hat{u}_{t-s_1-Q_2s_2-1}, \ldots, \hat{u}_{t-s_1-Q_2s_2-q}, \ldots\ldots, \hat{u}_{t-Q_1s_1-Q_2s_2}, \hat{u}_{t-Q_1s_1-Q_2s_2-1}, \\
&\ldots, \hat{u}_{t-Q_1s_1-Q_2s_2-q}\big), \qquad (5)
\end{aligned}
$$

$\boldsymbol{\Gamma}$ is the coefficients vector written as:

$$
\begin{aligned}
\boldsymbol{\Gamma} \;=\; &(\phi_1, \ldots, \phi_p, \varphi_1, -\phi_1\varphi_1, \ldots, -\phi_p\varphi_1, \ldots\ldots, \varphi_{P_1}, -\phi_1\varphi_{P_1}, \ldots, -\phi_p\varphi_{P_1}, \Phi_1, -\phi_1\Phi_1, \\
&\ldots, -\phi_p\Phi_1, -\varphi_1\Phi_1, \phi_1\varphi_1\Phi_1, \ldots, \phi_p\varphi_1\Phi_1, \ldots\ldots, -\varphi_{P_1}\Phi_1, \phi_1\varphi_{P_1}\Phi_1, \ldots, \phi_p\varphi_{P_1}\Phi_1, \\
&\ldots\ldots\ldots, \Phi_{P_2}, -\phi_1\Phi_{P_2}, \ldots, -\phi_p\Phi_{P_2}, -\varphi_1\Phi_{P_2}, \phi_1\varphi_1\Phi_{P_2}, \ldots, \phi_p\varphi_1\Phi_{P_2}, \ldots\ldots, \\
&-\varphi_{P_1}\Phi_{P_2}, \phi_1\varphi_{P_1}\Phi_{P_2}, \ldots, \phi_p\varphi_{P_1}\Phi_{P_2} - \theta_1, \ldots, -\theta_q, -\vartheta_1, \theta_1\vartheta_1, \ldots, \theta_q\vartheta_1, \ldots\ldots, \\
&-\vartheta_{Q_1}, \theta_1\vartheta_{Q_1}, \ldots, \theta_q\vartheta_{Q_1}, -\Theta_1, \theta_1\Theta_1, \ldots, \theta_q\Theta_1, \vartheta_1\Theta_1, -\theta_1\vartheta_1\Theta_1, \ldots, -\theta_q\vartheta_1\Theta_1, \\
&\ldots\ldots, \vartheta_{Q_1}\Theta_1, -\theta_1\vartheta_{Q_1}\Theta_1, \ldots, -\theta_q\vartheta_{Q_1}\Theta_1, \ldots\ldots\ldots, -\Theta_{Q_2}, \theta_1\Theta_{Q_2}, \ldots, \theta_q\Theta_{Q_2}, \\
&\vartheta_1\Theta_{Q_2}, -\theta_1\vartheta_1\Theta_{Q_2}, \ldots, -\theta_q\vartheta_1\Theta_{Q_2}, \ldots\ldots, \vartheta_{Q_1}\Theta_{Q_2}, -\theta_1\vartheta_{Q_1}\Theta_{Q_2}, \ldots, -\theta_q\vartheta_{Q_1}\Theta_{Q_2})^T \qquad (6)
\end{aligned}
$$

3. Posterior Analysis of DSARMA Models

By applying a transformation from $\mathbf{z}$ to $\mathbf{u}$, we can formulate and write the conditional-likelihood function of the DSARMA model (4) as:

$$\begin{aligned} L(\mathbf{\Gamma}, \tau \mid \mathbf{z}) &\propto \tau^{\frac{n-q}{2}} \exp\left\{-\frac{\tau}{2}\mathbf{u}^T\mathbf{u}\right\}, \\ &\propto \tau^{\frac{n-q}{2}} \exp\left\{-\frac{\tau}{2}\left(\mathbf{z}-W\mathbf{\Gamma}\right)^T\left(\mathbf{z}-W\mathbf{\Gamma}\right)\right\}. \end{aligned} \tag{7}$$

This likelihood function (7) is conditioned on the first q initial values, i.e. $(z_0, z_{-1}, \cdots, z_{1-q})$, where $q = P_2 s_2 + P_1 s_1 + p$.

To ease the posterior analysis of DSARMA model, we treat the products of the nonseasonal and seasonal coefficients as independent, free parameters. Also, we employ two priors for the DSARMA parameters Γ and τ: normal-gamma and Jeffreys' priors. Regrading the normal-gamma prior on $\mathbf{\Gamma}$ and τ; let τ follows Gamma distribution with parameters $\frac{\nu}{2}$ and $\frac{\lambda}{2}$ and $\mathbf{\Gamma}$ follows normal distribution, i.e. $N_{p^\star}(\mu_\Gamma, \tau^{-1}\Sigma_\Gamma)$. Thus, we can formulate the normal-gamma prior distribution of $\mathbf{\Gamma}$ and τ as:

$$\zeta_n\left(\mathbf{\Gamma}, \tau\right) \propto \tau^{\left(\frac{\nu+p^\star}{2}-1\right)} \exp\left\{-\frac{\tau}{2}\left[\lambda + \left(\mathbf{\Gamma}-\mu_\Gamma\right)^T\Sigma_\Gamma^{-1}\left(\mathbf{\Gamma}-\mu_\Gamma\right)\right]\right\}, \tag{8}$$

where $\mu_\Gamma, \Sigma_\Gamma, \nu$ and λ are hyper-parameters that need to be determined or estimated.

When no or little information is available about $\mathbf{\Gamma}$ and τ, we can assume Jeffreys' prior on $\mathbf{\Gamma}$ and τ that can be given as:

$$\zeta_j\left(\mathbf{\Gamma}, \tau\right) \propto \tau^{-1}, \ \ \tau > 0. \tag{9}$$

To derive the marginal-posteriors of the DSARMA parameters $\mathbf{\Gamma}$ and τ, we derive first the joint-posterior of both $\mathbf{\Gamma}$ and τ, and then obtain the marginal-posterior of one of them by analytically integrating out the other. By multiplying the likelihood function (7) with each of the two priors in (8) and (9), we get directly the joint-posterior of parameters $\mathbf{\Gamma}$ and τ.

First, for assuming the normal-gamma prior, we get the joint-posterior of parameters $\mathbf{\Gamma}$ and τ as:

$$\begin{aligned} \zeta_n\left(\mathbf{\Gamma}, \tau \mid \mathbf{z}\right) \propto\ & \tau^{\left(\frac{n-q+\nu+p^\star}{2}-1\right)} \exp\left\{-\frac{\tau}{2}\left[\lambda + \left(\mathbf{\Gamma}-\mu_\Gamma\right)^T\Sigma_\Gamma^{-1}\left(\mathbf{\Gamma}-\mu_\Gamma\right) + \right.\right. \\ & \left.\left.\left(\mathbf{z}-W\mathbf{\Gamma}\right)^T\left(\mathbf{z}-W\mathbf{\Gamma}\right)\right]\right\}. \end{aligned} \tag{10}$$

Second, for assuming Jeffreys' prior, we get the joint-posterior of $\mathbf{\Gamma}$ and τ as:

$$\zeta_j\left(\mathbf{\Gamma}, \tau \mid \mathbf{z}\right) \propto \tau^{\left(\frac{n-q}{2}-1\right)} \exp\left\{-\frac{\tau}{2}\left(\mathbf{z}-W\mathbf{\Gamma}\right)^T\left(\mathbf{z}-W\mathbf{\Gamma}\right)\right\}. \tag{11}$$

From the joint-posterior (10), Amin (2018) showed that the marginal-posterior of the coefficients vector $\mathbf{\Gamma}$ follows a multivariate-t distribution with: mean vector $\mu_n = A_n^{-1}B_n$, covariance matrix $V_n = \frac{1}{v_n-2}(C_n - B_n^T A_n^{-1} B_n)A_n^{-1}$ and degrees of freedom $v_n = (n+\nu-q)$; and also showed that the marginal-posterior of the DSARMA precision τ follows a gamma distribution, i.e. $Gamma(\frac{v_n}{2}, \frac{(C_n - B_n^T A_n^{-1} B_n)}{2})$, where

$$A_n = (W^T W + \Sigma_\Gamma^{-1}), \ \ B_n = (W^T\mathbf{z} + \Sigma_\Gamma^{-1}\mu_\Gamma), \ \text{ and } C_n = (\mathbf{z}^T\mathbf{z} + \lambda + \mu_\Gamma^T\Sigma_\Gamma^{-1}\mu_\Gamma). \tag{12}$$

Similarly, Amin (2018) derived from the joint-posterior (11) the marginal-posterior of the coefficients vector $\mathbf{\Gamma}$ to follow a multivariate-t distribution with mean vector $\mu_j = A_j^{-1}B_j$, covariance matrix $V_j = \frac{1}{v_j-2}(C_j - B_j^T A_j^{-1} B_j)A_j^{-1}$ and degrees of freedom $v_g = (n-q-p^\star)$; and derived the marginal-posterior of the DSARMA precision τ to follow a gamma distribution, i.e. $Gamma(\frac{v_j}{2}, \frac{(C_j - B_j^T A_j^{-1} B_j)}{2})$, where

$$A_j = W^T W, \ \ B_j = W^T\mathbf{z}, \ \text{ and } C_j = \mathbf{z}^T\mathbf{z}. \tag{13}$$

4. Predictive Analysis of DSARMA Models

Given the observed time-series $\mathbf{z}$, the joint predictive density of the k future observations, i.e. $\mathbf{z}_f = (z_{n+1}, z_{n+2}, \ldots z_{n+k})$, can be formulated as a conditional density of future observations $\mathbf{z}_f$ given $\mathbf{z}$ to be used as a basis to predict k future observations. Thus, we can obtain this joint predictive density by averaging the distribution of future values $\mathbf{z}_f$ with respect to the joint-posterior of $\mathbf{\Gamma}$ and τ as:

$$\begin{aligned} \zeta(\mathbf{z}_f \mid \mathbf{z}) &= \int_{\tau}\int_{\Gamma} \zeta\left(\mathbf{z}_f, \mathbf{\Gamma}, \tau \mid \mathbf{z}\right)\, d\Gamma\, d\tau \\ &= \int_{\tau}\int_{\Gamma} \zeta\left(\mathbf{z}_f \mid \mathbf{\Gamma}, \tau\right) \zeta\left(\mathbf{\Gamma}, \tau \mid \mathbf{z}\right)\, d\Gamma\, d\tau. \end{aligned} \tag{14}$$

This joint predictive density specified in Equation (14) is very complicated and cannot be derived in a closed-form. In particular, given $\mathbf{z}$, we can first derive the predictive density of one-step-ahead z_{n+1} that is used to predict z_{n+1}. Then, given $\mathbf{z}$ and z_{n+1} we derive the conditional predictive density of second-step-ahead z_{n+2}. Since z_{n+1} is unknown at time n, we can use its predicted value $\hat{z}_{n+1}$ to approximate the conditional predictive density of z_{n+2}, and so on. Thus, given $\mathbf{z}$ and the predicted values of $z_{n+1}, z_{n+2}, ..., z_{n+k-1}$ we can approximate the conditional predictive density of k^{th}-step-ahead z_{n+k} using the chain rule basis, as proposed by Broemeling and Land (1984) for AR models.

The one-step-ahead future observation z_{n+1} can be written from DSARMA model (4) as

$$z_{n+1} = W_{n+1}\mathbf{\Gamma} + u_{n+1}, \tag{15}$$

and also the k^{th}-step-ahead future observation z_{n+k} for $k > 1$ can be written as

$$z_{n+k} = W_{n+k}\mathbf{\Gamma} + u_{n+k}, \tag{16}$$

where W_{n+1} and W_{n+k} are W_t given in Equation (5) after replacing t by $(n+1)$ and $(n+k)$ respectively, and u_{n+1} and u_{n+k} are the one- and k^{th}-step-ahead future DSARMA model errors that are assumed to be normally distributed with zero-mean and precision τ.

In the subsequent theorem and corollary, we consider different prior specifications for the coefficient vector $\mathbf{\Gamma}$ and the precision τ, and demonstrate that the resulting one-step-ahead predictive density for z_{n+1} follows a univariate-t distribution.

Theorem 1 *Given the conditional-likelihood function introduced in Equation (7) for DSARMA model and using the normal-gamma prior on the DSARMA model coefficients vector $\mathbf{\Gamma}$ and precision τ specified in Equation (8), then the predictive density of the one-step-ahead future observation z_{n+1} follows a univariate-t distribution with v_n degrees of freedom, $F_n^{-1}E_n$ mean, and $\frac{1}{v_n-2}F_n^{-1}(L_n - E_n^T F_n^{-1} E_n)$ variance, where*

$$\begin{aligned} F_n &= (1 - W_{n+1}(A_n + W_{n+1}^T W_{n+1})^{-1} W_{n+1}^T) \\ E_n &= W_{n+1}(A_n + W_{n+1}^T W_{n+1})^{-1} B_n \\ L_n &= C_n - B_n^T (A_n + W_{n+1}^T W_{n+1})^{-1} B_n \end{aligned} \tag{17}$$

Proof: Given the joint-posterior of $\mathbf{\Gamma}$ and τ specified in Equation (10), then the predictive density of z_{n+1} can be formulated as

$$\begin{aligned} \zeta_n\left(z_{n+1} \mid \mathbf{z}\right) \propto \int_{\tau}\int_{\Gamma} \tau^{\left(\frac{n-q+\nu+p^{\star}+1}{2}-1\right)} \exp\Big\{-\frac{\tau}{2}\Big[\lambda + \left(\mathbf{\Gamma}-\mu_{\Gamma}\right)^T \Sigma_{\Gamma}^{-1}\left(\mathbf{\Gamma}-\mu_{\Gamma}\right) + \\ \left(\mathbf{z} - W\mathbf{\Gamma}\right)^T\left(\mathbf{z} - W\mathbf{\Gamma}\right) + \left(z_{n+1} - W_{n+1}\mathbf{\Gamma}\right)^2\Big]\Big\}\, d\Gamma\, d\tau. \end{aligned} \tag{18}$$

With some manipulations, this predictive density of z_{n+1} can be rewritten as

$$\zeta_n\left(z_{n+1} \mid \mathbf{z}\right) \propto \int_{\tau} \tau^{\left(\frac{n-q+\nu+1}{2}-1\right)} \exp\left\{-\frac{\tau}{2}\left[C_n - B_n^T A_n^{-1} B_n + (z_{n+1} - W_{n+1}\Gamma)^2\right]\right\} \times \int_{\Gamma} \tau^{\frac{p^\star}{2}} \exp\left\{-\frac{\tau}{2}\left[\left(\mathbf{\Gamma} - A_n^{-1}B_n\right)^T A_n \left(\mathbf{\Gamma} - A_n^{-1}B_n\right)\right]\right\} d\Gamma\, d\tau, \quad (19)$$

where A_n, B_n and C_n are defined in Equation (12). It is easy to figure that the integration with respect to $\mathbf{\Gamma}$ is proportional to the integration of multivariate normal and also the integration with respect to τ is proportional to the integration of gamma distribution. This leads to that the predictive density of the one-step-ahead future observation z_{n+1} is a univariate-t distribution with above given parameters.

Corollary 1 *Given the conditional-likelihood function introduced in Equation (7) for DSARMA model and using the Jeffreys prior on the DSARMA model coefficients vector $\mathbf{\Gamma}$ and precision τ specified in Equation (9), then the predictive density of the one-step-ahead future observation z_{n+1} follows a univariate-t distribution with v_j degrees of freedom, $F_j^{-1}E_j$ mean, and $\frac{1}{v_j-2}F_j^{-1}(L_j - E_j^T F_j^{-1} E_j)$ variance, where*

$$\begin{aligned} F_j &= (1 - W_{n+1}(A_j + W_{n+1}^T W_{n+1})^{-1} W_{n+1}^T) \\ E_j &= W_{n+1}(A_j + W_{n+1}^T W_{n+1})^{-1} B_j \\ L_j &= C_j - B_j^T (A_j + W_{n+1}^T W_{n+1})^{-1} B_j. \end{aligned} \quad (20)$$

Proof: By setting $\lambda = 0$, $\Sigma_\beta^{-1} = \mathbf{0}$, and $\nu = -p^\star$ we directly obtain this corollary result from Theorem (1), and A_j, B_j and C_j are defined in Equation (13).

In the subsequent theorem and corollary for $k > 1$ and using various priors for the model coefficients vector $\mathbf{\Gamma}$ and precision τ, we derive the predictive density of the k^{th}-step-ahead future observation z_{n+k} given the (k-1) future values $(z_{n+1}, z_{n+2}, ..., z_{n+k-1})$.

Theorem 2 *Given the (k-1) future values $(z_{n+1}, z_{n+2}, ..., z_{n+k-1})$, the conditional-likelihood function introduced in Equation (7) for DSARMA model and using the normal-gamma prior on the DSARMA model coefficients vector $\mathbf{\Gamma}$ and precision τ specified in Equation (8), then the conditional predictive density of the k^{th}-step-ahead future observation z_{n+k} follows a univariate-t distribution with (v_n+k-1) degrees of freedom, $F_{k_n}^{-1}E_{k_n}$ mean, and $\frac{1}{v_n+k-3}F_{k_n}^{-1}(L_{k_n} - E_{k_n}^T F_{k_n}^{-1} E_{k_n})$ variance, where*

$$\begin{aligned} F_{k_n} &= (1 - W_{n+k}(A_n + W_{n+k}^T W_{n+k} + W_{k-1}^T W_{k-1})^{-1} W_{n+k}^T) \\ E_{k_n} &= H_{n+k}(A_n + A_{n+k}^T A_{n+k} + W_{k-1}^T W_{k-1})^{-1} B_n \\ L_{k_n} &= C_n + \mathbf{z_{k-1}^T z_{k-1}} - B_{k_n}^T (A_n + W_{n+k}^T W_{n+k} + W_{k-1}^T W_{k-1})^{-1} B_{k_n} \\ B_{k_n} &= (W_{k-1}^T \mathbf{z_{k-1}} + B_n). \end{aligned} \quad (21)$$

W_{k-1} is a matrix of size $(k-1) \times p^\star$, with the i^{th} row is just W_t given in Equation (5) after replacing t by $(n+i)$ and $\mathbf{z_{k-1}} = (z_{n+1}, z_{n+2}, ..., z_{n+k-1})^T$.

Proof: Given the joint-posterior of $\mathbf{\Gamma}$ and τ specified in Equation (10), then the conditional predictive density of z_{n+k} given $(z_{n+1}, z_{n+2}, ..., z_{n+k-1})$ can be formulated as

$$\zeta_n\left(z_{n+k} \mid \mathbf{z}, z_{n+1}, ..., z_{n+k-1}\right) \propto \int_{\tau}\int_{\Gamma} \tau^{\left(\frac{n-q+\nu+p^\star+k-1}{2}-1\right)} \exp\left\{-\frac{\tau}{2}\left[\lambda + (\mathbf{\Gamma} - \mu_\Gamma)^T \Sigma_\Gamma^{-1}(\mathbf{\Gamma} - \mu_\Gamma) + (\mathbf{z} - W\mathbf{\Gamma})^T(\mathbf{z} - W\mathbf{\Gamma}) + (z_{n+k} - W_{n+k}\mathbf{\Gamma})^2 + \sum_{i=1}^{k-1}(z_{n+i} - W_{n+i}\Gamma)^2\right]\right\} d\Gamma\, d\tau. \quad (22)$$

With some manipulations and following the same steps as in Theorem 1, we derive the predictive density of the k^{th}-step-ahead future observation z_{n+k} to follow a univariate-t distribution with above given parameters.

Corollary 2 *Given the (k-1) future values ($z_{n+1}, z_{n+2}, ..., z_{n+k-1}$), the conditional-likelihood function introduced in Equation (7) for DSARMA model and using the Jeffreys' prior for the DSARMA model coefficients vector $\mathbf{\Gamma}$ and precision τ specified in Equation (9), then the conditional predictive density of the k^{th}-step-ahead future observation z_{n+k} is a univariate-t distribution with $(v_j + k - 1)$ degrees of freedom, $F_{k_j}^{-1}E_{k_j}$ mean, and $\frac{1}{v_j+k-3}F_{k_j}^{-1}(L_{k_j} - E_{k_j}^T F_{k_j}^{-1}E_{k_j})$ variance, where:*

$$\begin{aligned} F_{k_j} &= (1 - W_{n+k}(A_j + W_{n+k}^T W_{n+k} + W_{k-1}^T W_{k-1})^{-1} W_{n+k}^T) \\ E_{k_j} &= W_{n+k}(A_j + W_{n+k}^T W_{n+k} + W_{k-1}^T W_{k-1})^{-1} B_j \\ L_{k_j} &= C_j + \mathbf{z_{k-1}^T z_{k-1}} - B_{k_j}^T (A_j + W_{n+k}^T W_{n+k} + W_{k-1}^T W_{k-1})^{-1} B_{k_j} \\ B_{k_j} &= (W_{k-1}^T \mathbf{z_{k-1}} + B_j). \end{aligned} \tag{23}$$

Proof: By setting $\lambda = 0$, $\Sigma_\beta^{-1} = \mathbf{0}$, and $\nu = -p^\star$ we directly obtain this corollary result from Theorem 2.

Using Theorems 1 and 2 and Corollaries 1 and 2, we can conduct Bayesian predictions. In particular, we directly use the predictive means to provide point predictions for future values, and the prediction intervals can computed using the univariate-t distribution. For instance, using Theorem 1, for z_{n+1} its point prediction is given as $\hat{z}_{n+1} = F_n^{-1}E_n$ and its $(1-\alpha)\%$ prediction interval is given as:

$$F_n^{-1}E_n - t_{\frac{\alpha}{2},v_n}\sqrt{V_f} \leq z_{n+1} \leq F_n^{-1}E_n + t_{\frac{\alpha}{2},v_n}\sqrt{V_f} \tag{24}$$

where $V_f = \frac{1}{v_n-2}F_n^{-1}(L_n - E_n^T F_n^{-1}E_n)$ and $t_{\frac{\alpha}{2},v_n}$ is the univariate-t distribution $(1-\frac{\alpha}{2})$ percentile with v_n degrees of freedom.

5. Simulation Study and Application

This section has two parts, in the first one we introduce a simulation study to demonstrate how the proposed Bayesian prediction of DSARMA processes can be executed and also to evaluate its accuracy. In addition, we apply the proposed Bayesian prediction to real-world electricity load time series.

5.1. Simulation study

Aiming to simulate time series with different seasonality patterns, we generate time-series of size 1,010 from three DSARMA models, and the simulation design of these three DSARMA models is shown in Table 1.

Table 1 Simulation design

Model	ϕ_1	ϕ_1	φ_1	Φ_1	θ_1	θ_2	ϑ_1	Θ_1	τ
I	0.60	-	0.50	0.40	0.50	-	0.30	0.40	1.00
II	1.30	-0.60	0.40	0.30	-0.40	-	0.50	0.30	1.00
III	0.40	0.30	0.40	0.30	0.60	0.30	0.50	0.30	1.00

In order to apply our proposed Bayesian prediction of DSARMA processes, we employ Jeffreys' prior on the DSARMA model parameters, because of the difficulty to specify the hyperparameters for the normal-gamma prior. Also, we split each generated time-series dataset into both training

and testing sets; where the training set includes all the dataset except the last 10 observations that is considered as a testing set. We use the training set to estimate the DSARMA model and then we use the testing set to evaluate the accuracy of Bayesian prediction of DSARMA models for multiple-step-ahead predictions. After generating each time-series from designed DSARMA models, we apply the Bayesian prediction of DSARMA model in the following steps:

- We first estimate the coefficients and precision of identified DSARMA model using the marginal-posterior of the DSARMA model coefficients and precision, as derived in Section 3, including the posterior mean and posterior standard deviation along with the 95% credible interval.
- Using the predictive density of DSARMA models derived in Section 5, we predict the k future observations, including point and interval predictions for $k = 10$.

To evaluate the prediction accuracy of the DSARMA models, we compute the root mean square errors (RMSE) measure for the prediction results as

$$RMSE = \sqrt{mean(z_{n+i} - \hat{z}_{n+i})^2} \tag{25}$$

where z_{n+i} and $\hat{z}_{n+i}$ are the actual and forecast values respectively of the time series at time $n+i$ for $i = 1, 2, ..., k$. It measures how much the predictions deviate from the actual values of the time series. The smaller the value of the RMSE, the closer the corresponding prediction to the actual value. Thus, the DSARMA model with smaller RMSE values is more accurate.

We present the Bayesian estimates results of these simulated DSARMA models in Table 2, which shows that how the point Bayesian estimates are relatively close to the true values and also 95% credible intervals include corresponding true values. Similar result is obtained for the Bayesian predictions for $k = 10$ as presented in Table 3, and the computed RMSE values are 1.54, 1.24, and 1.38 for Models I, II, and III respectively, which confirm the forecasting accuracy of the fitted DSARMA models.

Table 2 Bayesian estimates of simulated DSARMA models

	ϕ_1	ϕ_1	φ_1	Φ_1	θ_1	θ_2	ϑ_1	Θ_1	τ
					Model-I				
Actual	0.60	-	0.50	0.40	0.50	-	0.30	0.40	1.00
$\hat{\mu}$	0.61	-	0.53	0.39	0.45	-	0.23	0.32	1.00
$\hat{\sigma}$	0.02	-	0.03	0.04	0.04	-	0.04	0.05	0.03
L	0.56	-	0.46	0.32	0.37	-	0.14	0.23	0.94
U	0.66	-	0.59	0.47	0.52	-	0.32	0.41	1.07
					Model-II				
Actual	1.30	-0.60	0.40	0.30	-0.40	-	0.50	0.30	1.00
$\hat{\mu}$	1.22	-0.53	0.46	0.28	-0.31	-	0.37	0.26	1.00
$\hat{\sigma}$	0.09	0.06	0.03	0.05	0.10	-	0.04	0.06	0.03
L	1.03	-0.66	0.40	0.17	-0.50	-	0.29	0.14	0.93
U	1.40	-0.41	0.52	0.38	-0.11	-	0.45	0.38	1.06
					Model-III				
Actual	0.40	0.30	0.40	0.30	0.60	0.30	0.50	0.30	1.00
$\hat{\mu}$	0.18	0.48	0.43	0.23	0.81	0.25	0.40	0.31	1.00
$\hat{\sigma}$	0.19	0.16	0.03	0.06	0.19	0.04	0.04	0.07	0.03
L	-0.20	0.17	0.37	0.11	0.43	0.17	0.31	0.17	0.93
U	0.56	0.80	0.49	0.35	1.19	0.33	0.48	0.45	1.06

5.2. Application on hourly electricity load

In this section of applications we aim to evaluate the accuracy of proposed Bayesian prediction of DSARMA models by conducting the Bayesian prediction to real time-series data with two seasonal

Table 3 Bayesian predictions of simulated DSARMA models

	Model-I				Model-II				Model-III			
k	z_{n+k}	$\hat{z}_{n+k}$	L	U	z_{n+k}	$\hat{z}_{n+k}$	L	U	z_{n+k}	$\hat{z}_{n+k}$	L	U
1	-1.87	-0.42	-2.44	1.60	-0.02	-1.15	-3.19	0.90	-1.71	0.81	-1.34	2.97
2	-0.32	0.34	-1.68	2.35	1.16	0.30	-1.73	2.33	-1.22	-0.38	-2.42	1.65
3	-0.72	1.94	-0.09	3.98	0.71	-0.19	-2.23	1.86	-1.21	-0.18	-2.26	1.89
4	0.31	2.23	0.21	4.25	1.54	3.86	1.81	5.90	0.22	2.27	0.20	4.34
5	0.52	2.76	0.75	4.78	1.16	2.22	0.19	4.25	0.28	2.68	0.63	4.74
6	2.10	1.84	-0.17	3.86	2.73	2.44	0.41	4.47	2.69	3.18	1.14	5.23
7	1.96	0.40	-1.61	2.40	1.89	-0.34	-2.37	1.68	2.36	2.07	0.00	4.14
8	-1.36	-1.93	-3.95	0.08	-1.23	-1.91	-3.93	0.11	0.25	0.34	-1.70	2.38
9	-2.75	-2.26	-4.27	-0.26	-2.26	-2.63	-4.66	-0.61	-0.83	-0.95	-3.00	1.11
10	-2.79	-1.32	-3.32	0.69	-2.54	-1.97	-4.00	0.05	-1.59	-0.87	-2.92	1.19

patterns. We apply our proposed Bayesian prediction of DSARMA models to hourly electricity load in European countries: Austria, Great Britain, Germany, and Spain[1]. These electricity load datasets are collected during seven weeks, each hour from Saturday 25 June 2011 to Friday 12 August 2011, and accordingly the collected time-series exhibit two seasonality layers: intraday (with first seasonal period s_1 = 24) and intraweek (with second seasonal period s_2 = 168), as it is depicted in Figure 1. We present the descriptive statistics of these electricity load datasets in Table 4, including minimum (Min.), mean, standard deviation (St. Dev.) and maximum (Max.). For more information about these electricity load datasets, the reader are referred to Amin (2018), where these datasets are first introduced and analyzed.

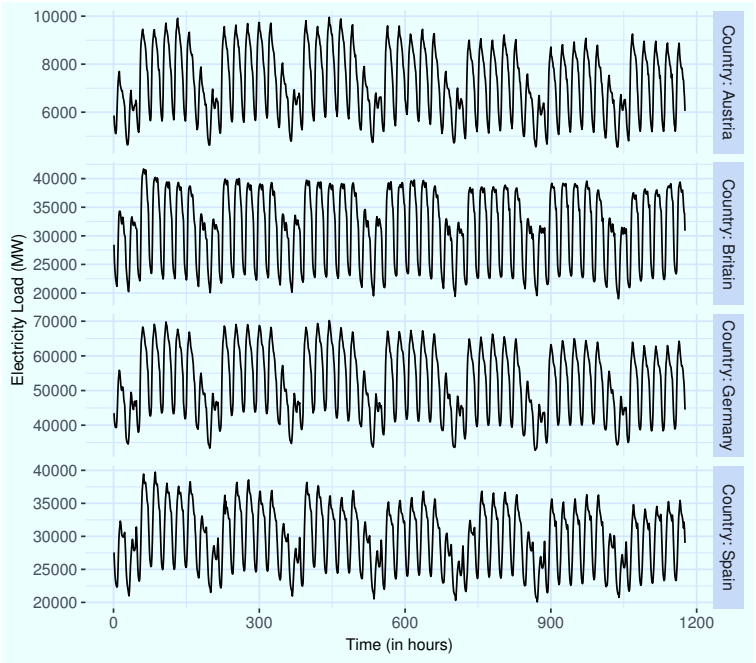


Figure 1 Hourly electricity load in European countries during seven weeks

[1]https://www.entsoe.eu/

Table 4 Summary statistics for hourly electricity load in Gigawatt

Time-series	Country	Min.	Mean	St. Dev.	Max.
1	Austria	4.54	7.21	1.38	9.95
2	Britain	19.00	31.37	6.23	41.70
3	Germany	32.75	52.04	9.72	70.20
4	Spain	20.07	29.61	4.61	39.73

In the work of Amin (2018), the corrected Akaike's information criterion (AIC_c) is applied to identify the best order of DSARMA model order for these electricity load datasets, and DSARMA identification results are presented in Table 5. In addition, For more details on the Bayesian estimation results for these electricity load datasets, the reader are referred to Amin (2018), where the Bayesian estimation is proposed for these DSARMA models and applied to these real datasets.

Table 5 Identified models for hourly electricity load datasets

		DSARMA order					
Dataset	Country	p	P_1	P_2	q	Q_1	Q_2
1	Austria	2	2	1	0	1	2
2	Britain	2	1	1	2	1	1
3	Germany	1	2	1	0	2	2
4	Spain	2	1	2	1	0	1

In order to apply our proposed Bayesian prediction of DSARMA models, we employ Jeffreys' prior on the DSARMA model parameters, because of the difficulty to specify the hyperparameters for the normal-gamma prior. Also, we split each electricity load dataset into both training and testing sets; where the training set includes all the dataset except the last day (24 observations) that is considered as a testing set. We use the testing set to evaluate the accuracy of Bayesian prediction of DSARMA models for one- and multiple-step-ahead predictions. In addition to the RMSE measure introduced in the simulation study, we also compute the mean absolute percentage error (MAPE) measure to present the prediction accuracy in percentage and unit free, since MAPE is computed as

$$MAPE = 100 * mean(|(z_{n+k} - \hat{z}_{n+k})/z_{n+k}|). \tag{26}$$

We present the Bayesian predictions for the future day of testing set (24 observations) in Figure 2, and also we present the computed RMSE and MAPE values for all the Bayesian predictions of these hourly electricity load datasets in Table 6. Although we use the DSARMA models to provide high step-ahead predictions, still most of the computed RMSE and MAPE values are small which reveal the forecasting accuracy of the current proposed Bayesian prediction. This conclusion is more highlighted in Figure 2, since the figure show that the prediction values are relatively very close to actual values. In general, all these results confirm the applicability of the current proposed Bayesian prediction to model and predict real-world time series with double seasonality and show its prediction accuracy even for high step-ahead predictions.

Table 6 Prediction measures for hourly electricity load time-series

Time-series	Country	RMSE	MAPE
1	Austria	0.90	10.17
2	Britain	4.07	9.44
3	Germany	3.74	6.49
4	Spain	2.58	7.82

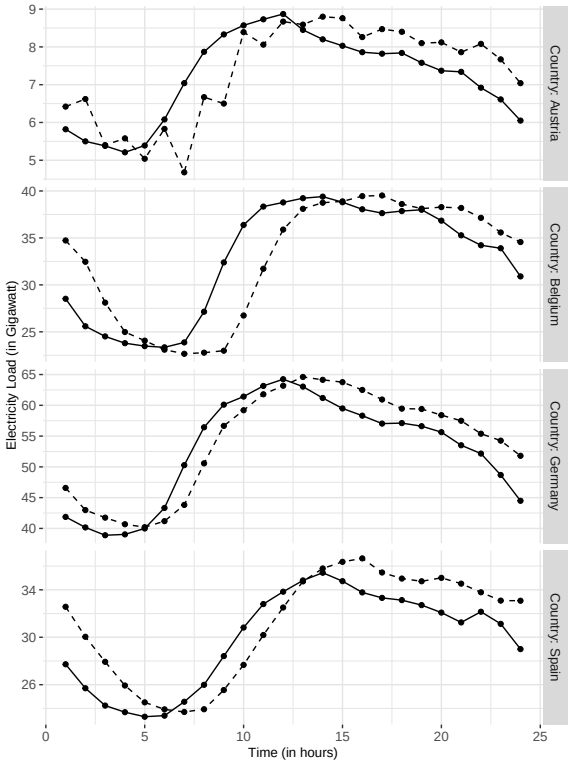


Figure 2 Bayesian predictions of hourly electricity load: actual (solid) and Bayesian prediction (dashed) values for one day in four European countries

6. Conclusion

In this paper, we presented the Bayesian prediction of DSARMA models with objective to model and predict time-series data with two seasonal cycles. To facilitate the derivation of the DSARMA predictive density, we assumed the errors term of DSARMA model is normally distributed and also we specified various priors for the parameters of DSARMA model. We first applied the nonlinear least squares (NLS) method to estimate the unknown lagged errors inherited in the moving-average component, and then we approximated the marginal-posteriors to be multivariate-t and gamma distributions for the DSARMA coefficients and precision respectively. Given these normally distributed DSARMA errors and specified parameters priors, we derived the predictive distributions of unobserved future observations, including one-step-ahead and multiple-step-ahead predictions, to be a univariate-t distribution. Since these predictive distributions are standards, they enable researchers to conduct accurately high step-ahead predictions. We conducted a simulation study and applied the proposed methodology to forecast real hourly electricity–load datasets from four European countries. The results confirm the forecasting accuracy of the current proposed Bayesian prediction approach. We plan to introduce a Bayesian identification for DSARMA models and also an extension to multivariate-time-series models.

Conflict of interest statement

There is no conflict of interest to disclose.

References

Abonazel MR and Abd-Elftah AI. Forecasting Egyptian GDP using ARIMA models. Rep. Econ. Finance. 2019; 5(1): 35-47.

Amin AA. Bayesian inference for seasonal ARMA models: A Gibbs sampling approach. Master's thesis, Statistics Department, Faculty of Economics and Political Science, Cairo University, Egypt, 2009.

Amin AA and Ismail MA. Gibbs sampling for double seasonal autoregressive models. Commun Stat Appl Methods. 2015; 22(6): 557-573.

Amin AA. Gibbs sampling for double seasonal ARMA models. In: Proceedings of the 29th annual international conference on statistics and computer modeling in human and social sciences, 2017.

Amin AA. Bayesian inference for double seasonal moving average models: A gibbs sampling approach. Pak J Stat Oper Res. 2017; 483-499.

Amin AA. Identification of double seasonal autoregressive models: a bayesian approach. In: Proceedings of the 52nd Annual International Conference of Statistics, Computer Science and Operations Research, Egypt, 2017.

Amin AA. Sensitivity to prior specification in Bayesian identification of autoregressive time series models. Pak J Stat Oper Res. 2017; 699-713.

Amin AA. Bayesian inference for double SARMA models. Commun Stat Theory. 2018; 47(21): 5333-5345.

Amin AA. Gibbs Sampling for Bayesian Prediction of SARMA Processes. Pak J Stat Oper Res. 2019; 397-418.

Amin AA. Kullback-Leibler divergence to evaluate posterior sensitivity to different priors for autoregressive time series models. Commun Stat Simulat. 2019; 48(5): 1277-1291.

Amin AA. Bayesian identification of double seasonal autoregressive time series models. Commun Stat Simulat. 2019; 48(8): 2501-2511.

Amin AA. Bayesian analysis of double seasonal autoregressive models. Sankhya B. 2020; 82(2): 328-352.

Amin AA. Bayesian Inference of Triple Seasonal Autoregressive Models. Pak J Stat Oper Res. 2022; 853-865.

Amin AA. Gibbs sampling for Bayesian estimation of triple seasonal autoregressive models. Commun Stat Theory. 2022; doi:10.1080/03610926.2022.2043379.

Amin AA. A Kullback-Leibler divergence based comparison of approximate Bayesian estimations of ARMA models. Commun Stat Appl Methods. 2022; 29(4): 471-486.

Amin AA. Full Bayesian analysis of double seasonal autoregressive models with real applications. J Appl Stat. 2024; 51(8): 1524-1544.

Awwad FA, Mohamoud MA and Abonazel MR. Estimating COVID-19 cases in Makkah region of Saudi Arabia: Space-time ARIMA modeling. PLoS ONE. 2021; 16(4): e0250149.

Barnett G, Kohn R and Sheather S. Bayesian estimation of an autoregressive model using Markov chain Monte Carlo. J Econom. 1996; 74(2): 237-254.

Barnett G, Kohn R and Sheather S. Robust Bayesian Estimation Of Autoregressive–Moving-Average Models. J Time Ser Anal. 1997; 18(1): 11-28.

Box GEP and Tiao GC. Bayesian inference in statistical analysis. Addison-Wesley, 1973.

Box GEP, Jenkins GM, Reinsel GC and Ljung GM. Time series analysis: forecasting and control. John Wiley and Sons, 2015.

Broemeling L and Land M. On forecasting with univariate autoregressive processes: A Bayesian approach. Commun Stat Theory. 1984; 13(11): 1305-1320.

Broemeling LD. Bayesian analysis of linear models. CRC Press, 1985.

Deb C, Zhang F, Yang J, Lee SE and Shah KW. A review on time series forecasting techniques for building energy consumption. Renew Sustain Energy Rev. 2017; 74: 902-924.

De Livera AM, Hyndman RJ and Snyder RD. Forecasting time series with complex seasonal patterns using exponential smoothing. J Am Stat Assoc. 2011; 106(496): 1513-1527.

Hyndman RJ and Athanasopoulos G. Forecasting: principles and practice. OTexts, 2018.

Ismail MA and Amin AA. Gibbs sampling for SARMA models. Pak J Stat. 2014; 30(2): 153-168.

Lago J, De Ridder F and De Schutter B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. Appl Energy. 2018; 221: 386-405.

Ryu S, Noh J and Kim H. Deep neural network based demand side short term load forecasting. Energies. 2017; 10(1): 3.

Shaarawy S and Broemeling L. Bayesian inferences and forecasts with moving averages processes. Commun Stat Theory. 1984; 13(15): 1871-1888.

Shaarawy SM and Ali SS. Bayesian identification of seasonal autoregressive models. Commun Stat Theory. 2003; 32(5): 1067-1084.

Sulandari W, Suhartono S and Rodrigues PC. Exponential Smoothing on Modeling and Forecasting Multiple Seasonal Time Series: An Overview. Fluct Noise Lett. 2021; 2130003.

Taylor JW and McSharry PE. Univariate Methods for Short-Term Load Forecasting. Adv Electr Power Energy Syst. 2017; 17-40.

Vermaak J, Niranjan M and Godsill SJ. Markov chain Monte Carlo estimation for the seasonal autoregressive process with application to pitch modelling. Department of Engineering, University of Cambridge, 1998.