SEASONAL HYDROLOGICAL AND METEOROLOGICAL TIME SERIES

Edilberto Cepeda-Cuervo.[1]
Statistics department
Science Faculty
Universidad Nacional de Colombia, Bogotá, Colombia
Jorge Alberto Achcar [2]
Social medicine department
FMRP, Universidade de São Paulo
Ribeirão Preto, S.P., Brazil
Marinho G. Andrade[3]
Applied mathematics and statistics department
ICMC, Universidade de São Paulo
São Carlos, S.P., Brazil

Summary

Time series models are often used in hydrology and meteorology to model streamflows series in order to make forecasting and generate synthetic series which are inputs for the analysis of complex water resources systems. In this paper we introduce a new modeling approach for hydrologic and meteorological time series assuming a continuous distribution for the data, where both the conditional mean and conditional variance parameters are modeled. Bayesian methods using standard MCMC (Markov Chain Monte Carlo Methods) are used to simulate samples for the joint posterior distribution of interest. Two applications to real data set illustrate the proposed methodology, assuming that the observations come from a normal, a gamma or a beta distribution. A first example is given by a time series of monthly averages of natural streamflows,

[1]email:ecepedac@unal.edu.co
[2]achcar@fmrp.usp.br
[3]email:marinho@icmc.usp.br

measured in the year period ranging from 1931 to 2010 in Furnas hydroelectric dam, Brazil. A second example is given with a time series of 313 air humidity data measured in a weather station of Rio Claro, a Brazilian city located in southeastern of Brazil. These applications motivate us to introduce new classes of models to analyze hydrological and meteorological time series

Keywords: *Hydrology time series data, Meteorological time series, Conditional regression models, Bayesian analysis, MCMC methods.*

1 Introduction

Time series models are often used in hydrology to model streamflow series in order to make predictions and to generate synthetic series which are inputs for the analysis of complex water resources systems (see, for example, Salas et al., 1980, 1982; Hosking, 1984; Hipel & McLeod, 1994; Montanari et al., 1997; Hasebe et al., 2000). In many studies, hydrologists also use time series data for displaying the amount of rainfall that has fallen in a region for the past day, year or a period of 10 years (see for example, Guimaraes & Santos, 2011, and Lee & Lee, 2000).

Modeling hydrological variability is very important in the planning and management of water resources. Many aspects of the hydrologic cycle could be described by time series data. Researches, usually use time series data to evaluate the resources of a water basin. Important variables related to streamflow and watershed describe streamflow properties such as monthly flows or streamflow parameters. A time series model estimates the streamflow parameters. Different time series models as ARMA and higher orders of MA models have been used by some authors when considering hydrologic regionalization of watersheds (see for example, Chiang et al., 2002 a, b). Spectral analysis and forecasting of hydrological time series also is considered by some authors (see, for example, Marques et al., 2006).

Considering hydrological time series, the monthly streamflow series typically have a periodic behavior in the mean and variance and in general, periodic autoregressive models are adopted in de analysis of the data (see, for example, Modal & Wasimi, 2006). In this situation, usually it is assumed that the series flow has a normal or log-normal distribution(see for example, Tesfaye et al., 2006; Wang et al., 2009).

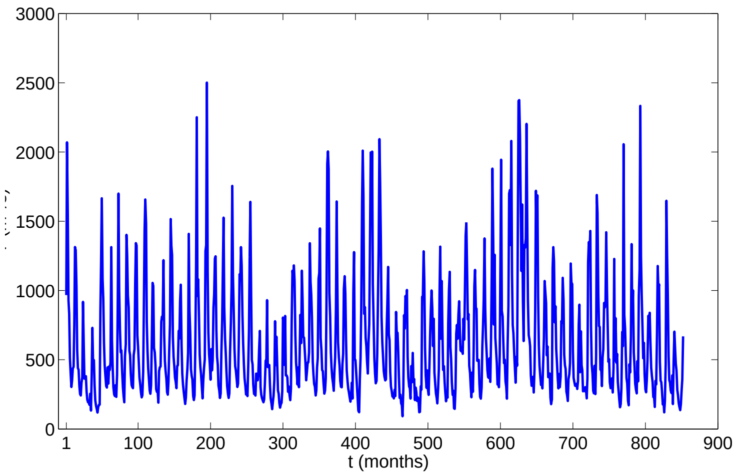

Figure 1: Time series of monthly averages of natural Streamflows, measured in the period 1931 to 2010, in Furnas hydroelectric dam, in southeastern Brazil.

Similar behavior have meteorological time series. In this case, given that the relative air humidity is a random variable taking values in the open interval $(0, 1)$, usually it is assumed a beta distribution to analyze the data. Another possibility to analyze the data set is to consider a transformation of the data and to assume a normal distribution to the transformed data. As especial case, we could assume a logistic transformation.

In this paper, a more general assumption is considered in the analysis of the hydrological or meteorological time series conditional to the historical available information: it is assumed that the data is generated from a normal, lognormal, gamma or beta distribution, with conditional mean and variance, given respectively, by $E(Y_t|Y_{t-1})$ and $V(Y_t|Y_{t-1})$. Thus a general model is proposed to analyze hydrological or meteorological time series, assuming that the observations come from to the continuous biparametric exponential family of distribution.

To illustrate and motivate the use of the proposed models, we first consider a data set consisting of the time series of monthly averages of natural streamflows,

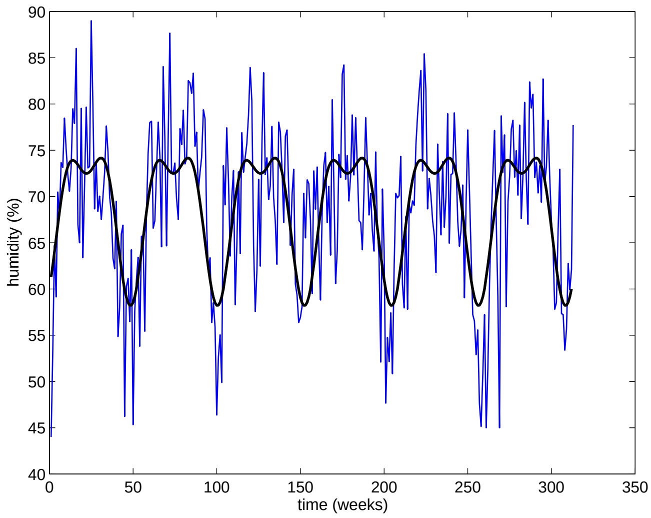

Figure 2: The air humidity time series data and the fitted periodical mean.

measured in the year period ranging from 1931 to 2010, in Furnas hydroelectric dam, located in southeastern Brazil. This time series is shown in Figure 1. From Figure 1, we observe that the streamflow series have a periodic behavior in the mean and variance and in general periodic autoregressive models are assumed in the analysis of this kind of time series data (Modal & Wasimi, 2006).

To take into account this heteroscedasticity in the time series of streamflows, the model proposed in this paper assumes seasonal and autoregressive terms in the modeling of the mean and variance parameters. In this way, we propose a periodic and heteroscedastic model, in which the variance also presents a autoregressive structure.

A second example, with behavior similar to the assumed hydrological time series, we consider a meteorological time series given by weekly averages of air humidity, measured in the city of Rio Claro, São Paulo state, Brazil. This time series is displayed in Figure 2. In this case, given that the relative air humidity is a random variable taking values in the open interval $(0, 1)$, it is assumed a beta distribution in the analysis of the data. Thus, a joint mean and variance beta regression model, including seasonal and autoregressive terms in both, mean and variance models, is proposed in this paper to analyze this type of data sets.

This paper is structured as follows: in section 2, the seasonality analysis of the time series is introduced. In section 3, seasonal autoregressive models are proposed. Section 4 present the results of the analysis of the hydrological time series obtained using the proposed models, assuming normal and gamma distributions. In section 5, the results of the analysis of air humidity time series are presented. Finally, in section 6 some conclusions and future research topics are included.

2 A Period model

In this section, we introduce a new modeling approach that includes seasonality terms which better describes time series of monthly averages of natural stream flows, denoted by Y_t. As illustration, we consider a time series of monthly averages of natural Streamflows in Furnas hydroelectric dan, introduced in section 1 (see Figure 1). A spectral analysis is developed for this time series to determine the time periods to be considered in the mean and variance model formulation. Thus, if in the spectral analysis, the number of observations is $T = 2q + 1$, where q is a positive integer number, the Fourier time series model given by,

$$y_t = \alpha_0 + \sum_{i=1}^{q} (\alpha_{1i} \cos(2\pi f_i t) + \alpha_{2i} \sin(2\pi f_i t)) + e_i \quad (1)$$

is fitted, where $f_i = i/T$ is the ith harmonic of the fundamental frequency $1/T$ and, α_{1i} and α_{2i}, $i = 1, \ldots, q$, are the related coefficients and e_i is an error term assumed to have a particular parametric distribution.

Observe that the highest frequency is 0.5 cycle per month (time interval) since the smallest period is 2 months. In the time series formulations to be introduced in Section 3, only frequencies whose periods have a higher intensity are considered.

Usually, in the case of monthly time series, these periods are given by periods of 6 and 12 months. These periods will be presented in Section 4, where the specification of the parameters related to the periods in done jointly with the specification of the parameters for the autoregressive models assumed for the mean and variance. If some of these parameters show to be not significative after a preliminary statistical analysis, the terms related to these parameters are excluded from the model to develop a further analysis.

3 The Proposed Seasonal Autoregressive Model

In hydrological or meteorological time series, we assume that the observations of interest variable are generated from a conditional continuous probability distribution function. As special cases, we could assume that the observations were generated from standard conditional probability distributions as normal,log-normal, gamma, beta or exponential conditional density functions, denoted by $f(y_t|H_{t-1})$, $t = 1, 2, \ldots, n$, where H_{t-1} is the available information up time $t-1$ and thus Y_t has conditional means and variances given respectively by $\mu_t = \mathrm{E}(Y_t|H_{t-1})$ and $h_t = \mathrm{Var}(Y_t|H_{t-1})$, following the models:

$$\mu_t = \alpha_0 + \sum_{i=1}^{q}(\alpha_{1i}\cos(2\pi f_i t) + \alpha_{2i}\sin(2\pi f_i t)) + \sum_{i=0}^{p}\phi_i y_{t-i} \quad (2)$$

$$\log(h_t) = \lambda_0 + \sum_{i=1}^{s}(\lambda_{1i}\cos(2\pi f_i t) + \lambda_{2i}\sin(2\pi f_i t)) + \sum_{i=0}^{r}\theta_i y_{t-i} + \epsilon_i \quad (3)$$

where $\boldsymbol{\beta} = \{\phi_0, \phi_1, \ldots, \phi_p,\ \alpha_{11}, \ldots, \alpha_{1q},\ \alpha_{21}, \ldots, \alpha_{1q}\}$, is the vector of parameters for the mean model and $\boldsymbol{\gamma} = \{\theta_0, \theta_1, \ldots, \theta_r,\ \lambda_{11}, \ldots, \lambda_{1s},\ \lambda_{21}, \ldots, \lambda_{1s}\}$ is the vector of the parameters of the variance model, $f_i = i/T$ the ith harmonic of the fundamental frequency $1/T$, and $\epsilon_i \sim N(0, \tau^2)$, where the variance τ^2 is an unknown parameter, $N(a, b^2)$ denotes a normal distribution with mean equals to a and variance equals to b^2. These parameters are estimated using a Bayesian approach.

In order to illustrate the proposed methodology, we also include the equations to relate the mean and variance parameters in the gamma and beta distributions.

1. If Y_t, $t = 1, 2, \ldots, n$, follows a gamma conditional distribution $G(p_t, q_t)$, where $G(p, q)$ denotes a gamma distribution with mean pq and variance

pq^2, the conditional mean and variance are related to the original parameters by the equations $\mu_t = p_t q_t$ and $h_t = \mu_t q_t$.

2. If Y_t, $t = 1, 2, \ldots, n$, follows a beta distribution function $B(p_t, q_t)$, the reparametrization of the beta distribution density as function of the mean and precision, $\phi_t = p_t + q_t$, result to be appropriate in order to define the joint mean and precision beta regression models as appear in Cepeda(2001). This reparametrization, where $\phi = p + q$, $p = \mu\phi$ and $q = \phi(1 - \mu)$, has been intensively used in the literature following the joint modeling approach for the mean and precision beta parameters introduced by Cepeda-Cuervo (2001) and Cepeda and Gamerman (2005), under a Bayesian approach. It is important to point out that Ferrari and Cribari-Neto (2004) also introduced modeling for the mean but considering constant precision parameters, under a classical approach. In all of these cases, ϕ can be interpreted as a precision parameter in the sense that, for fixed values of μ, larger values of ϕ correspond to smaller values for the variance of Y. This is not an easy interpretation. Thus, in this paper, we use the mean and variance reparametrization of the beta distribution function in the definition of joint mean and variance beta regression models, taking into account $\mu(1-\mu) > \sigma^2$, samples of the posterior distribution of the parameters should be simulated in the subspace of parameters that satisfy this property. Although this reparametrization result in a complex expression for the beta distribution, it leads to a best and more easily interpretation for the statistical analysis results in the applications. In this reparametrization,

$$p_t = \frac{(1-\mu_t)\mu_t^2 - \mu_t\sigma_t^2}{\sigma_t^2} \quad (4)$$

$$q_t = \frac{(1-\mu_t)[\mu_t - \mu_t^2 - \sigma_t^2]}{\sigma_t^2} \quad (5)$$

where the autoregressive seasonal beta regression model have the conditional mean and variance model given by the equations (2) and (3).

Special cases can be proposed easily from this model. A first one, is a seasonal mean model, with mean given by (2) and autoregressive variance without seasonal terms. A second one, a seasonal mean model, with mean given by (2) and seasonal variance without autoregressive terms. A third one, the autoregressive mean and variance model, without seasonal terms in the mean and in the variance. A fourth one, an autoregressive model, with constant variance.

4 Hydrological Time Series

In this section we also consider as a motivation for the introduction of new modeling for hydrological time series, the Furnas dam hydroelectric hydrological time series introduced in section 1, assuming autoregressive conditional heteroscedastic models.

As in Cepeda-Cuervo et al. (2012), the first step in the proposed analysis is to determine the period for the harmonics of higher intensity in the spectral analysis of the Stream flows data. In this way, we note that the harmonics of higher intensity corresponds to the cycle $(1/f_i)$ of 6 and 12 months. Thus, the seasonal term to be included in the mean equation model of the streamflow series are given by: $\cos\left(\frac{2\pi}{6}t\right)$, $\sin\left(\frac{2\pi}{6}t\right)$, $\cos\left(\frac{2\pi}{12}t\right)$ and $\sin\left(\frac{2\pi}{12}t\right)$.

Many autoregressive models could be assumed to analyze this data set. As special cases, in section 4.1, we assume a normal conditional distribution and in section 4.2, we assume a conditional gamma distribution. In order to apply the Bayesian methodology independent normal prior distributions $N(0, 10^k)$, with $k = 2$, are assumed for the parameters associated with the seasonal terms. For the other parameter in the model, independent normal prior distributions with $k = 5$ are assumed. For the parameter σ_e^2, the variance of the error term introduced in equation (1), we assume a gamma prior distribution, $G(0.001, 0.001)$. Observe that we are assuming very non-informative prior distributions for all parameters.

4.1 Normal seasonal time series

In this section, we present the results of the analysis of the time series of monthly averages of natural streamflows, measured in the period 1931 to 2010, in hydroelectric dam introduced in Section 1, assuming joint mean and variance autoregressive normal models. From the model given by equations (2) and (3), many autoregressive models were considered and the obtained model with smallest DIC (Deviance Criterion Information, introduced by Spiegelhalter et al, 2002) value was the heteroscedastic normal regression model with conditional mean and variance models given respectively by

$$\begin{aligned} \mu_t &= \beta_0 + \beta_1 cos(2\pi t/6) + \beta_2 sin(2\pi t/6) \\ &\quad + \beta_3 cos(2\pi t/12) + \beta_4 y_{t-1} + \beta_5 y_{t-2} \qquad (6) \\ h_t^2 &= \exp(\gamma_0 + \gamma_1 y_{t-1} + \gamma_2 y_{t-3} + \gamma_3 cos(2\pi t/12) + e_t) \qquad (7) \end{aligned}$$

Samples of the joint posterior distribution of interest were simulated using standard MCMC (Markov Chain Monte Carlo) methods and the free available WinBugs software (Spiegelhalter et al, 2003). In each of the cases, many samples were generated starting from different initial values. All of them showed the same behavior, after a small burn-in period consisting of 3000 or 5000 generated samples.Convergence of the simulation algorithm was observed from trace plots of the generated Gibbs samples.

For the model given by equations (6) and (7), the value of the logarithm of the likelihood function evaluated at the obtained estimates for the parameters of the model was given by $-2logL = 10479.200$ and the obtained DIC criterion value used in Bayesian discrimination of models was given by $DIC = 10876.500$. Monte Carlo estimates of the posterior means for each parameter based on the generated Gibbs samples and their respective standard deviations are given, respectively, in Tables 1 and 2 for this model including both autoregressive and seasonal terms in the conditional mean and variance terms.

Parameter	β_0	β_1	β_2	β_3	β_4	β_5
Mean	77.44 (6.409)	62.1 (5.557)	40.45 (5.597)	162.7 (9.71)	0.674 (0.03023)	0.124 (0.02258)

Table 1: Double Normal seasonal model: Mean parameters estimates.

Parameter	γ_0	γ_1	γ_2	γ_3	σ_e^2
Mean	7.926 (0.1563)	0.002357 $(2.097E-4)$	$5.559E-4$ $(2.452E-4)$	1.743 (0.1351)	1.005 (0.1768)

Table 2: Normal seasonal regression model:Variance parameters estimates.

4.2 Gamma seasonal time series

In this section we present the results of the analysis of the time series for monthly averages of natural streamflows, measured in the period 1931 to 2010, in Furnas hydroelectric dam, assuming joint mean and variance autoregressive gamma models, that is, we assume that the observations of the interest are generated from a conditional gamma density function given by $f(y_t|H_{t-1})$, where H_{t-1} is the information up to time $t-1$ and Y_t has conditional mean and conditional variance given respectively by $\mu_t = \mathrm{E}(Y_t|H_{t-1})$ and $\sigma_t^2 = \mathrm{Var}(Y_t|H_{t-1})$, respectively, and defined by (6) and (7). For this model, the logarithm of the likelihood function evaluated at the estimates for the parameters of interest is given by $-2logL = 10649$ and the DIC value is given by 10862.300. Using the DIC criterion to discriminate the two models (normal seasonal time series and gamma seasonal time series), we observe better fit of the data for the gamma seasonal time series model, since we have smaller DIC value for this model. The posterior parameter estimates of the parameter together with the corresponding standard deviation are given in Tables 3 and 4.

Parameter	β_0	β_1	β_2	β_3	β_4	β_5
Mean	78.74 (7.126)	60.2 (5.645)	41.59 (5.567)	176.6 (8.615)	0.7239 (0.03278)	0.1181 (0.0238)

Table 3: Gamma seasonal model: mean parameters estimates.

Although the mean estimates given in Tables 1 and 3, and variance estimates given in tables 2 and 4, show some agreement between the conditional normal

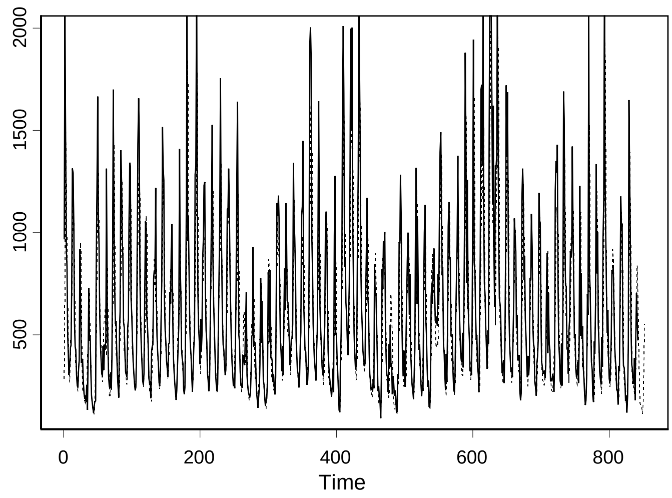

Figure 3: Time series of monthly averages of natural Streamflow and normal fit mean estimates.)

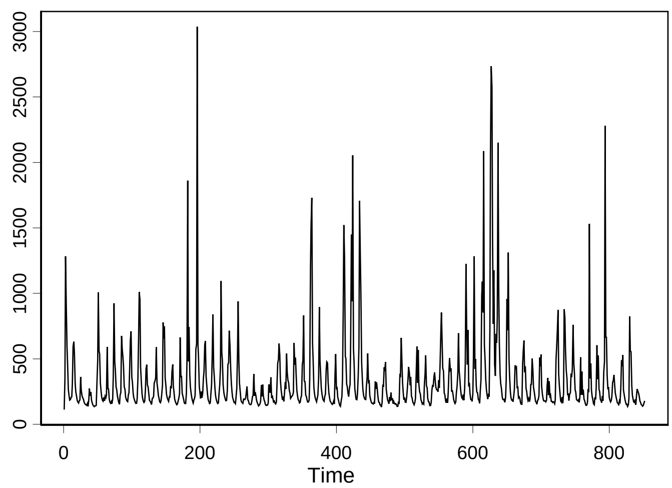

Figure 4: Squared root of the expected volatility in the normal time series.

Parameter	γ_0	γ_1	γ_2	γ_3	σ_e^2
Mean	7.942 (0.1394)	0.002513 (1.829$E-4$)	$5.467E-4$ (2.151$E-4$)	1.595 (0.1156)	0.3342 (0.107)

Table 4: Gamma seasonal model: variance parameters estimates.

and gamma estimates, the DIC value of the conditional gamma models is smaller than the DIC value of the conditional heteroscedastic models, showing that the second model is better fitted by the monthly averages of natural streamflows data.

5 Beta mean and variance seasonal applied to time series analysis

In this section it is assumed that the time series data come from a beta distribution $B(p_i, q_i)$, with mean and precision given respectively by (2) and (3). To illustrate the application of the model we consider the time series of weekly averages air humidity, measured in Rio Claro, located in southeastern Brazil, from 18/10/2002 to 08/10/2008. This time series introduced in section 1, is show in Figure 2.

5.1 Spectral analysis

As in Cepeda et al., (2012), the first step in the proposed analysis is to determine the period of harmonics of higher intensity in the spectral analysis of the Stream flows data. From this paper, we note that the harmonics of higher intensity corresponding to the cycle $(1/f_i)$ of 26 and 52 days. Thus, the seasonal term to be included in the mean equation model of the streamflow series are given by: $\cos\left(\frac{2\pi}{26}t\right)$, $\sin\left(\frac{2\pi}{26}t\right)$, $\cos\left(\frac{2\pi}{52}t\right)$ and $\sin\left(\frac{2\pi}{52}t\right)$.

5.2 Seasonal mean and conditional variance models

In this section, double seasonal beta repression models are proposed to analyze the air humidity time series data. In this way, we assume the following models for the mean and dispersion parameters:

$$\begin{aligned} \text{logit}(\mu_{t_i}) &= \beta_0 + \beta_1 \cos(2\pi t_i/52) + \beta_2 \sin(2\pi t_i/52) \\ &\quad + \beta_3 \cos(2\pi t_i/26) + \beta_4 \sin(2\pi t_i/26) + \beta_5 \text{logit}(Y_{i-1}) \quad (8) \\ \log(h_{t_i}) &= \gamma_0 + \gamma_1 \cos(2\pi t_i/52) + \gamma_2 \sin(\pi/52) + \gamma_3 \sin(\pi/26) + e_i \quad (9) \end{aligned}$$

were considered, where $e_i \sim N(0, \sigma_e^2)$. Assuming independent normal prior distributions for the regression parameters, $\beta_i \sim N(0, 100)$, $\gamma_j \sim N(0, 100)$, for $i = 0, 1, \ldots, 5$, and $j = 0, 1, \ldots, 4$, and a gamma prior distribution for the variance of the error, $\sigma_e^2 \sim G(0.001, 0.001)$, 20000 samples of the posterior distribution were also generated using the WinBugs software (Spiegelhalter et al, 2003). Monte Carlo estimates for the posterior means of each parameter were obtained from the final simulated Gibbs sample after an initial burn-in sample period of 2000 samples.This "burn-in" sample was discarded to eliminate the effect of the initial values in the iterative procedure. After this "burn-in" sample period we simulated another 20,000 Gibbs samples choosing every 20th iteration to get approximately non correlated samples, which gives a final sample of size 1,000 used to get the posterior summaries of interest. The posterior summaries of interest are given in Table 5, for the mean parameters, and in Table 6, for the variance parameters. The assumed initial values were $\beta_0 = 1$, $\beta_1 = 0.2$, $\beta_2 = 0.01$, $\beta_3 = 0.2$, $\beta_4 = 0$, $\beta_5 = 0$, $\gamma_0 = -4$, $\gamma_1 = 0$, $\gamma_2 = 0$, $\gamma_3 = 0$, $\sigma_e^2 = 0, 12$.

Parameter	β_0	β_1	β_2	β_3	β_4	β_5
Mean	0.5845 (0.0469)	-0.2091 (0.03039)	0.09499 (0.02583)	-0.06574 (0.02488)	0.08705 (0.02481)	0.2741 (0.05449)

Table 5: Beta regression model: mean parameters estimates.

Parameter	γ_0	γ_1	γ_2	γ_3	σ_e^2
Mean	-5.521 (0.07953)	0.2095 (0.1161)	0.2327 (0.1195)	-0.213 (0.1173)	0.02826 (0.05044)

Table 6: Beta regression model: variance parameter estimates.

The logarithm of the likelihood function evaluated at the obtained estimates for the parameters of the model is given by $logL = -853.082$ and the DIC criterium has value equal to -825.525. In Figure 5, we observe a good agreement between data and the fitted mean, showing the good performance of the proposed model. In Figure 6, we observe a good agreement for the variances, observing that smaller the variances, smaller are the means. This behavior was observed in the original time series. That is, we conclude that this model is very well fitted by the time series data.

6 Conclusions

In this paper, we introduced a new class of time series models assuming continuous random variables within the exponential family.Special cases were considered assuming, normal, gamma and beta distributions.The proposed methodology was illustrated considering hydrological time series and weather time series.

Under a Bayesian approach and using recent MCMC simulation procedures and free available software, as the WinBugs software we observed that this new class of models gives a great flexibility of fit for times series data, as observed in some applications considering Brazilian hydrological and meteorological data. These results could be a great interest in applications to hydrological and meteorological time series.

References

[1] Cepeda-Cuervo, E. (2001). Variability Modeling in Generalized Linear Models, *Unpublished Ph.D. Thesis. Mathematics Institute, Universidade Federal do Rio de Janeiro. http://www.docentes.unal.edu.co/ecepedac/docs/MODELAGEM DA VARIABILIDADE.pdf*

[2] Cepeda C. E. and Gamerman D. (2005). Bayesian methodology for modeling parameters in the two parameter exponential family. *Estadística*, **57**, 93-105.

[3] Cepeda-Cuervo, E., Andrade, M. G., & Achcar, J. A. (2012, October). A seasonal and heteroscedastic gamma model for hydrological time series: A Bayesian approach. In AIP Conference Proceedings (Vol. 1490, p. 97).

[4] Chiang, S. M., Tsay, T. K. & Nix, S. J. (2002a). Hydrologic regionalization of watersheds. i: Methodology development. *Journal of Water Resources Planning and Management*, **128**(1), 3-11.

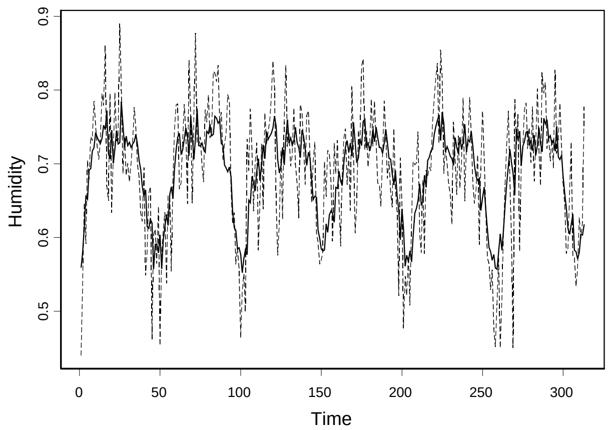

Figure 5: Air humidity time series data (continuous line) and fitted mean (diseased line) for double seasonal model.

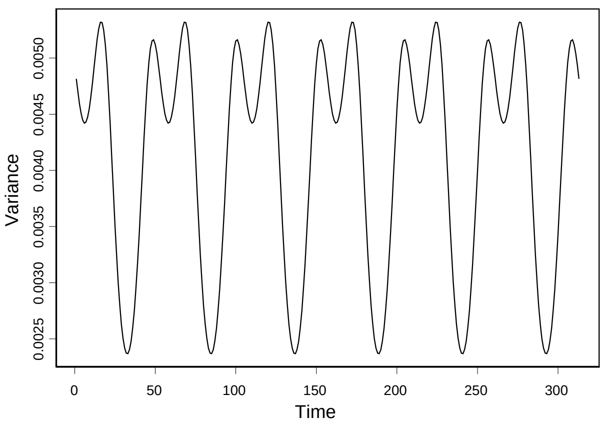

Figure 6: Fitted conditional variance time series.

[5] Chiang, S. M., Tsay, T. K. & Nix, S. J. (2002b). Hydrologic regionalization of watersheds. ii: Applications. *Journal of Water Resources Planning and Management*, **128**(1), 12- 20.

[6] Ferrari, S., Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions, *Journal of Applied Statistics* **31**, 799-815.

[7] Guimaraes, R. & Santos, E. G. (2011). Principles of stochastic generation of hydrologic time series for reservoir planning and design: A case study. *Journal of Hydrologic Engineering*. Accepted Manuscripts.

[8] Hasebe, M., Dandou, T., Kumekawa, T. & Neijou, S. (2000). Time series analysis of monthly rainfall, mean air temperature and carbon dioxide. In W. Z. Y. & S. X. Hu, editors, *Proceedings of the eighth International Symposium on Stochastic Hydraulics*, pages 533-537, Beijing, China. International Symposium on Stochastic, International Association for Hydraulic Research

[9] Hipel, K. W. & McLeod, A. E. (1994). *Time series modeling of water resources and environmental systems.* Elsevier, Amsterdam, The Netherlands.

[10] Hosking, J. R. M. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research*, **20**(12), 1898-1908.

[11] Lee, J. Y. & Lee, K. K. (2000). Use of hydrologic time series data for identification of rechargemechanism in a fractured bedrock aquifer system. *Journal of Hydrology*, **229**, 190-201.

[12] Marques, C. A. F., Ferreira, J. A., Rocha, A., Castanheira, J. M., Melo-Goncalves, P., Vaz, N. & Dias, J. M. (2006). Singular spectrum analysis and forecasting of hydrological time series. *Physics and Chemistry of the Earth*, Parts A/B/C, **31**(18), 1172-1179.

[13] Modal, M. S. & Wasimi, S. A. (2006). Generating and forecasting mointhly flows of the ganges river with par model. *Journal of Hydrology*, **323**(1-4), 41-66.

[14] Montanari, A., Rosso, R. & Taqqu, M. S. (1997). Fractionally differenced arima models applied to hydrologic time series: Identification, estimation and simulation. *Water Resources Research*, **33**(1-4), 1035-1044.

[15] Salas, J. D., Delleur, J. W., Yevjervich, V. & Lane, W. L. (1980). *Applied modeling of hydrologic time series.* Water Resources Publications, Littlton, USA.

[16] Salas, J. D., Boes, D. C. & Smith, R. A. (1982). Estimation of arma models with seasonal parameters. *Water Resources Research*, **18**(4), 1006-1010.

[17] Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B, **64**(4), 583-639.

[18] Spiegelhalter, D. J., Thomas, A., Best N. G., Gilks, W. R. (2003). WinBUGS User Manual (version 1.4). MRC Biostatistics Unit, Cambridge, U.K.

[19] Tesfaye, Y. G., Meerschaert, M. M. & Anderson, P. L. (2006). Identification of periodic autoregressive moving average models and their application to the modeling of river flows. Water Resources Research, 42(W01419), 1-11.

[20] Wang, Q. J., Robertson, D. E. & Chiew, F. H. S. (2009). A bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. Water Resources Research, 45(W05407), 1-18.