Prediction of monthly flows for Três Marias reservoir (São Francisco river basin) using the CFS climate forecast model

Previsão de vazões mensais ao reservatório de Três Marias usando a previsão climática do modelo CFS

Luana Ferreira Gomes De Paiva[1] , Suzana Maria Montenegro[1] & Marcio Cataldi[2]

[1]Universidade Federal de Pernambuco, Recife, PE, Brasil
[2]Universidade Federal Fluminense, Niterói, RJ, Brasil
E-mails: luanag@ons.org.br (LFGP), suzanam@ufpe.br (SMM), mcataldi@id.uff.br (MC)

ABSTRACT

Despite the water crisis in 2016, 76% of the energy in Brazil was generated by hydroelectric plants, which shows that the Brazilian system is still strongly dependent on the hydrological conditions of basins. Therefore, the flow forecasts for these plants subsidize the decision making within the scope of the Electric Sector, since they allow the evaluation of the operational conditions of the hydroelectric and thermoelectric plants through the use of energy optimization models, providing gains in the operations of SIN (Sistema Interligado Nacional – the Brazilian National Interconnected System). The precipitation forecast is of fundamental importance for the elaboration of these hydroelectric flow forecasts. For energy evaluations, the DECOMP and NEWAVE models are used, with the GEVAZP model being applied to generate scenarios through an AR (p) (autoregressive) model. Accordingly, this study shows the impact of precipitation forecast on flow predictions in the climate horizon. For this, a statistical correction was made in the rain predicted by the CFS (Climate Forecast System) model, which tends to overestimate the predicted rain, with rainfall-flow models being calibrated. Tests were performed with this new modeling system and the results, in the form of scenarios, were compared with the scenarios generated by the GEVAZP model, showing the possibility of reducing the generated range by the latter, consequently causing the DECOMP model to not consider ranges with little or no probability of occurrence, which can improve the optimization of the SIN operation planning. This work also shows that the SMAP model exhibited better performance when compared to the Neural Networks model, in terms of the average flow range predicted in relation to the observed flow. There was a clear improvement in the flow predictions with the incorporation of the rain observed one month ahead in the simulations, mainly in the forecast of high flows. Finally, the climate indices had a good relationship with the flow and rain variables.

Keywords: Flow forecast; São Francisco river; Neural Networks (NN); SMAP (Soil Moisture Accounting Procedure); Precipitation forecast; Hydrological simulation.

RESUMO

Mesmo com a crise hídrica, em 2016, 76% da energia no Brasil foi gerada por usinas hidrelétricas, mostrando que ainda se tem um sistema que é fortemente dependente da situação hidrológica das bacias hidrográficas. Com isso, as previsões de vazões a essas usinas subsidiam a tomada de decisão do âmbito do Setor Elétrico, já que permitem a avaliação das condições operacionais das usinas hidrelétricas e termelétricas, através da utilização dos modelos de otimização energética, proporcionando ganhos na operação do SIN – Sistema Interligado Nacional. A previsão de precipitação-é de fundamental importância para a elaboração dessas previsões de vazões das hidrelétricas. Para as avaliações energéticas são utilizados os modelos Decomp e Newave, que utilizam o Modelo Gevazp para a geração dos cenários através de um modelo AR(p). Assim, o presente trabalho objetiva mostrar o impacto da previsão de precipitação na previsão de vazões, no horizonte climático. Para isso fez-se uma correção estatística na chuva prevista pelo Modelo CFS e foram calibrados modelos chuva vazão Após essa calibração, foram realizados testes com esse novo sistema de modelagem e os resultados, em forma de cenários, foram comparados com os cenários gerados pelo Modelo Gevazp, mostrando que há possibilidade de reduzir a faixa gerada por esse último e, consequentemente, fazer com que o modelo Decomp não visite faixas com pouca ou

nenhuma probabilidade de ocorrência, o que pode melhorar tanto a otimização do planejamento da operação do SIN. Este trabalho também mostra também que o modelo SMAP apresentou melhor desempenho quando comparado ao modelo de Redes Neurais, quando comparado com o intervalo de vazão médio previsto em relação ao fluxo observado. Houve uma melhoria clara nas previsões de fluxo com a incorporação da chuva observada um mês à frente nas simulações, principalmente na previsão de vazões altas. Por fim, os índices climáticos apresentaram boa relação com as variáveis vazão e chuva.

Palavras-chave: Previsão de vazões; Rio São Francisco; Redes neurais; SMAP; Previsão de precipitação; Simulação hidrológica.

INTRODUCTION

Despite the water crisis that the Brazilian Northeast Subsystem has been facing since 2013, approximately 76% of the country's total energy generation in 2016 was derived from hydropower. Due to this peculiar characteristic, both the planning and the electroenergetic generation of SIN (Sistema Interligado Nacional – the Brazilian National Interconnected System) exhibit a close correlation with the water stocks in the reservoirs of hydroelectric plants and with the flow of these reservoirs. With this, the information on flow prediction for these plants is extremely important for the planning and programming of the SIN, as they subsidize the decision making of the Brazilian Electric Sector, allowing the evaluation of the operational conditions of the hydroelectric and thermoelectric plants over time, respecting the electrical constraints through the use of different optimization and energy simulation models, providing synergic gains for the SIN operation. These forecasts, when previously performed and carried out with relative accuracy, allow decision making to minimize the effects of scenarios of exceptional flood or shortage of water resources. In this context, it is important to highlight the role of precipitation forecast, taking into account its importance for forecasting hydroelectric plant flows (Cataldi et al., 2007).

As described by Rocha et al. (2015), the Brazilian System flow representation should be divided into three stages: the first refers to the first month of the planning horizon, with weekly discretization and flows being individually treated by each plant. The second stage refers to a period of one or more months ahead (only a one-month period, according to current practices), with monthly discretization and flows individually treated by each plant. The third stage refers to the remaining period until the horizon of five years ahead, with monthly discretization and flows treated in an aggregated way in terms of energy, for each of the electrically-interconnected subsystems, covering all hydroelectric subsystems usage in the form of equivalent reservoirs. Nevertheless, it is important to detach the daily programming process from the Dessem Model (Centro de Pesquisa de Energia Elétrica, 2003).

ONS (Operador Nacional do Sistema Elétrico – the Brazilian National Electric System Operator) is responsible for preparing the forecast of average natural flows on a daily, weekly and monthly basis for all hydroelectric usage sites, in addition to generating monthly scenarios, to be used in the planning and scheduling processes of the SIN operation. The probabilistic runoff scenarios for the horizon of one month up to five years ahead, adopted in the second and third stages of the planning studies, which complement the representation of hydroelectric energy generation by flow, are obtained through the GEVAZP model (Maceira & Mercio, 1997; Jardim et al., 2001); a stochastic model that generates synthetic series, which, are possible realizations of

a certain stochastic process, under the hypothesis of stationarity and ergodicity. The flows of these series are generated through autoregressive modeling, with the inclusion of random noise obtained from the time series of each plant, according to a set of probabilistic laws and without taking into account any past or future climate trends. Nevertheless, it is important to note that this runoff scenario generation does not objectively incorporate any type of meteorological or climate information.

Many studies have been carried out to better understand subseasonal predictions and how to use these in the flow forecast process, for instance, by using neural networks or physical models. Some application examples of these methodologies are discussed ahead.

When analyzing the subseasonal to seasonal (S2S) predictions, White et al. (2017) showed that S2S prediction fills the gap between short-range weather prediction and long-range seasonal outlooks. In this specific study, the emerging operational S2S forecasts are presented to the wider weather and climate applications community by undertaking the first comprehensive review of sectoral applications of S2S predictions, including public health, disaster preparedness, water management, energy and agriculture.

On the other hand, Vitart et al. (2017) developed a database for S2S predictions. The S2S database includes near-real-time ensemble forecasts and reforecasts of up to 60 days from 11 centers. This database will also help to assess the potential of current operational S2S systems to forecast extreme events around the globe.

In turn, Baker et al. (2019) enhanced climate forecast relevance and usability through the development of a system for evaluating and displaying real-time S2S climate forecasts on a watershed scale. The paper described the formulation of S2S climate forecast products based on the Climate Forecast System version 2 (CFSv2) and the North American Multi-Model Ensemble (NMME). Forecast verification reveals to be an appreciable skill in the first two bi-weekly periods (Weeks 1–2 and 2–3) from CFSv2, being useful in the NMME Month 1 forecast, with varying skills at longer lead times, depending on the season. The application of a bias-correction technique (quantile mapping) eliminates forecast bias in the CFSv2 reforecasts, without any significant change of the Pearson correlation skill.

Regarding the flow forecasting process, Tucci et al. (2015) presented an assessment of ensemble inflow forecasts for a hydropower reservoir in Brazil, the Três Marias dam. Inflow forecasts with lead times of 15 days were generated twice a day using a 14-members ensemble obtained from the global numerical weather prediction, run by the Brazilian Weather Forecasting Center (CPTEC), and a large-scale hydrological model. Results are encouraging and it

is believed that ensemble inflow forecasts for reservoirs in Brazil will be used in a near future as inputs to the optimization of the national electric power production system. Therefore, the MGB-IPH hydrological model (Collischonn et al., 2007) was used to conduct the generation of ensemble inflow forecasts.

In turn, Fan et al. (2016) introduced a mass conservative scenario tree reduction in combination with detailed hindcasting and closed-loop control experiments for a multipurpose hydropower reservoir in a tropical region in Brazil - the Três Marias hydropower. In the experiments, precipitation forecasts based on the observed data, as well as deterministic and probabilistic forecasts are used to generate streamflow forecasts in a hydrological model over a period of 2 years. Results for a perfect forecast show the potential benefit of the online optimization and indicate a desired forecast lead time of 30 days. The range of the energy production rate between the different approaches is relatively small, varying between 78% and 80%, which suggests that the use of stochastic optimization combined with ensemble forecasts leads to a significantly higher level of flood protection without compromising energy production.

Regarding neural networks, which are also used in the present work, according to Evsukoff et al. (2012), data-based models such as Artificial Neural Networks (ANNs) and neuro-fuzzy models have been applied on hydrological modeling and have provided good simulation results, although they usually require a great number of parameters for large period simulations.

When analyzing the scenario generation process, Cossich et al. (2015) showed the possibility of improving the method of generating runoff scenarios in the horizon of up to three months, considering, in addition to information on past runoff, GCM climate precipitation forecasts from the atmosphere, using the ECHAM 4.5, CFS, COLA/IRI and CCM3 models. The authors applied a statistical AR model and another ARx (autoregressive) model. The vresults showed that both models represented the seasonal flow behavior well. However, the univariate methodology presented inferior quality in relation to the multivariate methodology, since AR models cannot usually anticipate flow variations, being unable to capture their natural variability. The authors highlighted that the insertion of the climate precipitation forecasts by sets proved to be an important complement to the univariate stochastic modeling. They also pointed out that, even with the errors of climate models in relation to the simulation of the precipitation variable, it can be concluded that their use in scenario generation models is promising, pointing out that these GCMs tend to have better results in the future.

On the other hand, when analyzing the other flow forecasting methodology that will be used in this work, Fernández Bou et al. (2015) developed a model for predicting flows one month ahead, based on precipitation forecasts originating from the CFSv2 model. The model used for rainfall-runoff transformation was SMAP. Precipitation forecasts were corrected (correction coefficients were generated for each month of the year) and compared to rainfall resulting from 12 rainfall monitoring stations from ANA (Agência Nacional de Águas – the Brazilian National Water Agency). The precipitation forecast used was the average of the 25 members. The results achieved better quality flow predictions than the official ONS model, currently used in the PMO (Programa Mensal de Operação - Monthly Operation Program) processes and their revisions.

Da Silva et al. (2018) used the SMAP model to make flow forecasts for the Água Vermelha hydroelectric plant, located in the Grande River Basin. In addition, the RegCM Regional model was used for precipitation forecasts, with this model being also used with boundary conditions from the MIROC model. Both models had a bias correction, essential to improve the rain simulation. As a result, it was observed that the model was able to simulate the main patterns observed throughout South America. The simulation showed that the rainfall has an added value when the regional climate model is used, as compared with the global climate model.

Regarding the evaluation of atmospheric climate variables as inputs in neural network models, the distribution of SST (Sea Surface Temperature) anomalies and their influence on the oceans, as well as other climate variables, are useful for society, considering the influence of SST on the climate of all the Earth. Therefore, experts have been continuously taking this information as a basis for their studies, including its relationship with precipitation and river flows.

For instance, Pinto et al. (2006) presented the results of climate indicators used in the probabilistic forecast of semiannual precipitations (October-March) and quarterly flows (OND and JFM) at the Alto São Francisco basin. The climate indicators used were SST anomalies in different regions of the oceans and the SOI (Southern Oscillation Index). The model developed estimates the probabilities of precipitation and seasonal flows to occur in categories defined as normal, below normal and above normal. Consensual forecasts were made with two or three probabilistic models, being of higher quality in relation to forecasts solely based on climatology for both precipitations and flows.

With this in mind, the objectives of this study are to evaluate the generation of runoff scenarios from rainfall-runoff models, using climate prediction, subsequently comparing these scenarios with those generated by GEVAZP, which is a generator of synthetic series without any type of climate conditioning. Accordingly, two rainfall-runoff modeling methodologies were tested: one conceptual, based on the SMAP model, and another based on Artificial Neural Networks (ANN), from a Multi-Layer Perceptron (MLP) model.

These techniques were applied in order to present the gain that can be obtained with the use of precipitation information in the generation of runoff scenarios, through rainfall-runoff modeling, in the horizon of up to one month, and its consequences for the planning and the operation of the SIN. The usage evaluation was carried out for the hydroelectric located in the São Francisco river basin: the Três Marias dam, which together with Sobradinho and the Luiz Gonzaga hydroelectric, represent 96% of the EAR (Energia Armazenada - Stored Energy) in the Brazilian Northeast Subsystem.

MATERIALS AND METHODS

Study area

According to Ana (Agência Nacional de Água, 2004), the São Francisco basin has an area of 639,000 km² and its main course has an extension of 2,700 km between the headwaters, in Serra da Canastra, located in the municipality of São Roque de Minas (state

of Minas Gerais), and the mouth, in the Atlantic Ocean, between the states of Sergipe and Alagoas. The basin area corresponds to approximately 8% of the national territory, covering parts of six Brazilian states and the Federal District. It is also worth noting that the basin comprises a significant part of the Drought Polygon, which is a territory recognized by the Brazilian legislation as being subject to critical periods of prolonged droughts and is mainly located in the Northeastern region, extending to the north of the State of Minas Gerais.

The present study was developed on the drainage area of the Três Marias reservoir, located in the municipality of Alto São Francisco. Rainfall data from ANA and CEMIG (Companhia Energética de Minas Gerais – the Energy Company of the State of Minas Gerais) were used.

Study assumptions

In order to obtain the runoff scenarios, climate forecasts from the CFS models (Saha et al., 2010), operated in National Centers for Environmental Prediction (NCEP), are used. These predictions served as input for the rainfall-runoff models to elaborate the flood scenarios for up to one month ahead, using the SMAP model and a MLP Neural Network model in this rainfall-runoff transformation. Figure 1 shows the flow diagram of the steps adopted in the current study.

A brief description of the CFS atmospheric model, the method for correcting this precipitation forecast, as well as of the rainfall-runoff modeling obtained from the SMAP and MLP Neural Networks models are also shown ahead. Moreover, a brief report on the climate information that was used in this study is also presented.

a) Climate forecast model and forecast correction

The CFS is a climate model developed by the National Centers for Environmental Prediction (NCEP) to simulate the condition of the coupled ocean-atmosphere-land system and sea-surface ice, with high resolution for the period from 1979 to 2010. The global atmospheric model has a horizontal spatial resolution of approximately 38 km with 64 vertical levels. The oceanic model has a latitudinal spacing of 0.25° near the Equator and of up to 0.5° in the tropics, with 40 levels, to a depth of 4,737 meters. The first version of CFS, later called CFSv1, was put into operation at the NCRMSE in August 2004 and was the first global model used at the NCRMSE for seasonal prediction (Saha et al., 2006) of fully coupled atmosphere-ocean-land.

As described by Saha et al. (2014), the second version of CFS (CFSv2) was operationalized at the NCEP in March 2011. This version has brought improvements to almost all aspects of the data assimilation components and system model forecast. A coupled reanalysis was conducted over a 32-year period
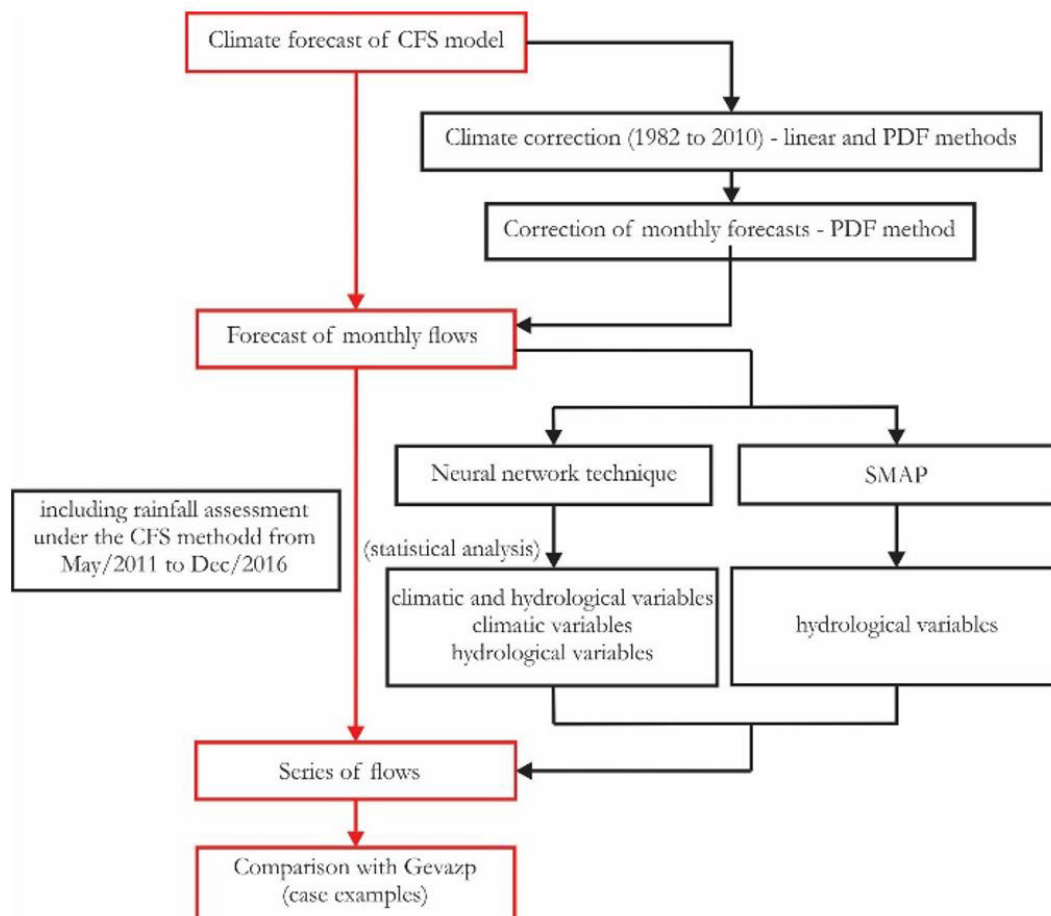

Figure 1. Flowchart diagram of the study steps.

(1979-2010), which provided the initial conditions for performing retrospective forecasts for the period from 1982 to 2010.

For the correction of the precipitation forecast, the methodology described by Wood et al. (2002) will be used, the PDF (Probability Density Function) methodology, in which, for each month of the year and for each forecast grid point of the climate model, two probability distribution curves are developed, for observed and predicted data, respectively . The correction is made based on the equality of the frequency curves, as shown in Figure 2.

The PDF methodology consists of determining the frequency of each variable and the correction through the equality of the frequencies between the observed and the predicted curves, with subsequent calculations of the Pearson correlation coefficient between the two variables being obtained. The linear correction consists of the determination of a coefficient of direct relation between the climate and the average rain observed in the basin.

In order to evaluate the quality of the precipitation forecasts of the CFS model by graphic analysis, the results predicted by this model in the period from 2011 to 2016 were compared with the average rainfall data of the Três Marias reservoir.

Figure 3 shows the location of the reservoirs of Três Marias, in Brazil..

b) Climate variables analysis

This study evaluated the climate variables and possible influences of SST anomalies in the Equatorial Pacific, Tropical Atlantic and other climate variables, as well as their relationship with the flow of the Três Marias reservoir.

As showed in Venables & Smith, 2017, Principal Component Analysis (PCA) aims to simplify the structure of a set of variables and explain their total variance as much as possible by determining values as linear combinations of the original variables. The Principal Components (PCs) should be used to reduce the size of the data, being formed by transforming a set of "p" variables into a set of uncorrelated variables. These new variables are linear combinations of the original variables, being obtained in descending order of importance such that, for example, the first major component accounts for as much of the total variability as possible in the original data.
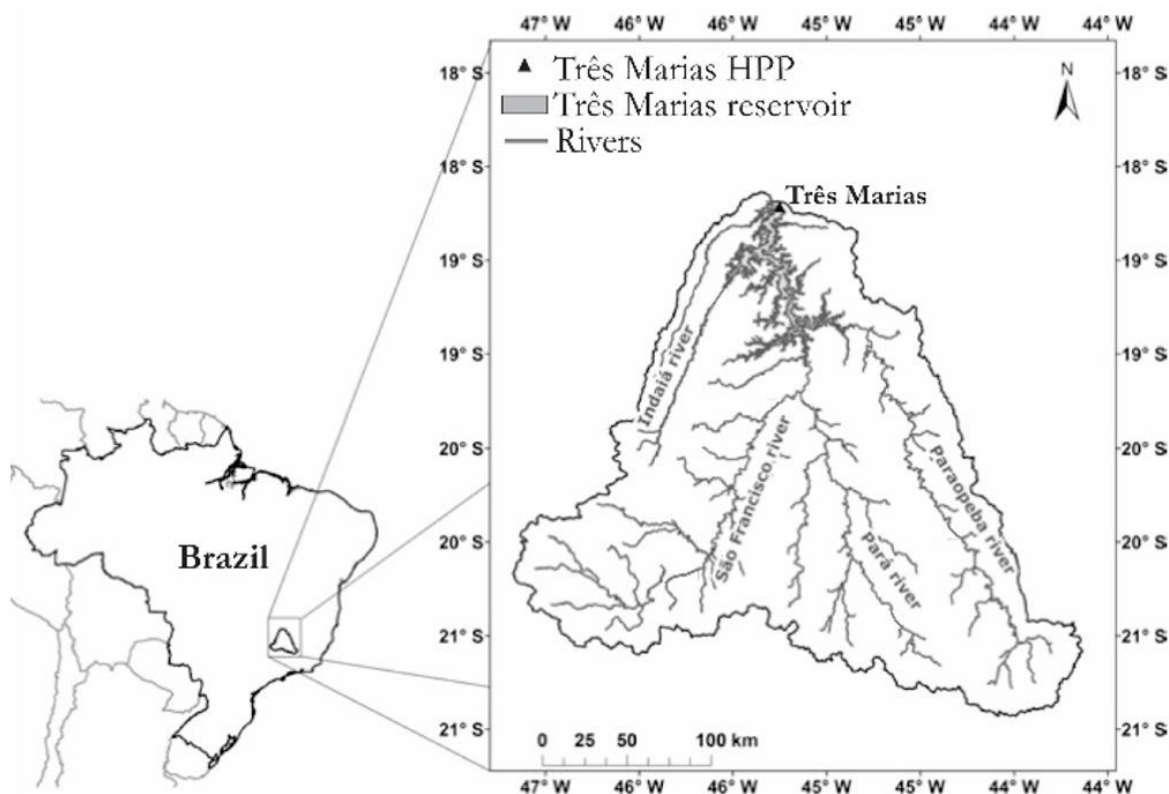

Figure 2. Methodology for precipitation correction.

Figure 3. Drainage area of the Três Marias reservoirs. Adapted from Tucci et al. (2015).

c) Flow forecast and scenario generation models

There are several methodologies to estimate rainfall-runoff transformation. Physical and stochastic models, as well as those based on the Neural Networks technique, are available. Among the physical or conceptual models, this text highlights the SMAP model, which is widely used in Brazil, either because of its feasibility and few parameters, or due to the quality obtained in its calibration.

As previously shown, Fernández Bou et al. (2015) developed a model to forecast flows one month ahead, based on precipitation forecasts. The precipitation forecasts originated from the CFSv2 model, with SMAP being the model used for rainfall-runoff transformation. Precipitation forecasts were corrected (correction coefficients were generated for each month of the year) and compared to a rainfall resulting from 12 ANA rainfall stations. The precipitation forecast used was the average of 25 members. The results achieved were flow forecasts of better quality than the official ONS model, currently used in PMO processes and their revisions.

The models based on the Neural Networks technique have also been widely used in several sectoral applications and also in the area of water resources, in the rainfall-runoff transformation. This study also highlights the feasibility and the results obtained using this methodology.

In this study, rainfall-runoff models based on NN (Neural Networks) and the SMAP model are used. The approach presented selected a Multi-Layer Perceptrons (MLP) Neural Network with a Levemberg-Marquardt supervised learning algorithm, similar to that used by Gomes (2006). This choice was based on its good performance presented in the previously mentioned study, in which it was used for a similar problem to the one approached in this study.

c.1) Neural Networks models

One of the most important properties of a Neural Network is its learning ability, for improving performance. This occurs through an iterative adjustment process, applied to its weights, known as training. This learning ability occurs from the downward gradient that is performed by the backpropagation algorithm. The goal of this downward gradient is to seek a global minimum through adjustments in the synaptic weights, by a value that is proportional to the opposite direction of the derivative of the error provided by the neuron in relation to the weighting value. The term backpropagation is related to the recursive propagation of errors. Therefore, the training can be understood as follows: initially, the signals are propagated in a progressive way (from the input layer to the output layer). Then, the recursively propagated errors (from the output layer to the input layer) are calculated through the network by determining the derivatives of the error function. Finally, these derivatives are used to perform weight adjustments (Valença, 2005).

Moreover, as previously mentioned, the weight and bias adjustments are optimized according to the Levemberg-Marquardt algorithm, suitable for networks with a moderate amount of data (up to several hundreds of synaptic weights). This method has a higher convergence speed than the networks with downward training gradient, in which network weights and deviations are updated as the performance function has a sharper decrease (Gomes, 2006).

In the training process, the cross-validation method was applied, which uses an independent dataset to determine the optimum stop point during training, in order to especially minimize the risks of super-adjustment or under-adjustment. Thus, the data set was divided into three sub-sets:

– training: patterns used to modify weights;

– validation: patterns used to mainly check the problem of super-adjustment (overfitting);

– test: patterns to test the final model performance.

Figure 4 shows an example of one of the simulations performed, showing that the simulation was interrupted when the error began to increase for the validation dataset, avoiding over adjustment.

c.2) SMAP model

The SMAP (Soil Moisture Accounting Procedure) model is a conceptual model, focused on hydrological simulation of rainfall-runoff transformation, having as main advantages its simplicity, feasibility in obtaining the input data, application for the great majority of SIN basins, ease of understanding the methodology as well as the functioning of the model and its parameters, which allows adjustments/improvements, as well as the use of a small number of parameters.

The SMAP was originally developed by Lopes et al. (1982) for a daily time interval and later adapted to the hourly and monthly versions, with some changes in its structure. In this study, the monthly version was used.

In its monthly version, SMAP consists of two mathematical reservoirs (Rsolo and Rsub), whose state variables are updated monthly.

The model also contains a routine of previous updates of the humidity level, in which, for each time interval, a percentage of the average rain of the month is added in order to use the average moisture content of the month. Such a routine is an increment that considerably improves the results, out of observations,
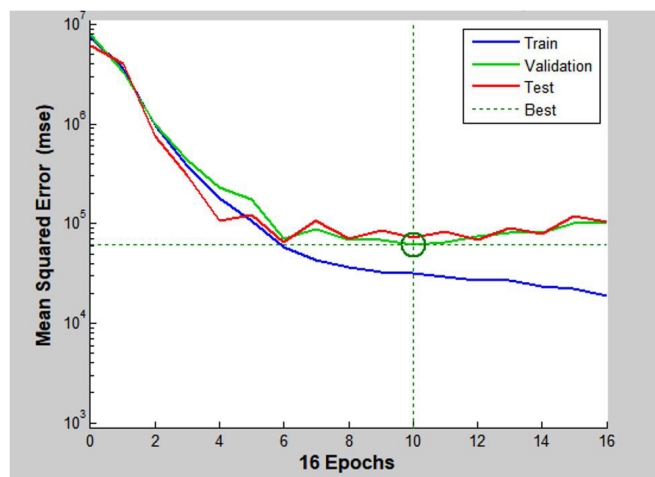

Figure 4. Process for stopping the training of a Neural Network of the study (output example from Matlab).

mainly in regions of great variability in the pluviometric regime (Lopes et al., 1982).

The main parameters obtained in the calibration of the SMAP are (the detailed formulation can be observed in Fernández Bou et al. (2015)):

Str - soil saturation capacity (mm);

K2t - surface flow parameter (dimensionless);

Capc - recharge coefficient (dimensionless), related to the movement of water in the unsaturated zone of the soil and, therefore, a function of soil type;

K - Recession constant (month -1).

The results obtained considered the MAPE (mean absolute percentage error), the RMSE (Root-mean-square error), the MAE (mean absolute error), the Nash Coefficient and Nash-Log. The detailed formulation can be observed in Fernández Bou et al. (2015) and Gomes (2006).

c.3) Generation scenario

The results of the flow forecasts by the SMAP and NN models were compared with the flows generated by GEVAZP through two example applications for the months of September 2012 and January 2013, and for the period from January 2015 until December 2016. GEVAZP (Maceira & Mercio, 1997; Jardim et al., 2001) is a stochastic model of synthetic series generation, in which realizations of a stochastic process are possible, under the hypotheses of stationarity and ergodicity. The flows of these series are generated from autoregressive modeling, with the inclusion of random noise, obtained from the time series of each plant and according to a set of probabilistic laws.

RESULTS AND DISCUSSION

Climate and correction forecast model

Correction of the precipitation forecast was made for the drainage area of the Três Marias reservoir. This correction is detailed below.

Figure 5 shows a comparison of the average rainfall predicted by the CFS with 1, 2, and 3 months in advance, compared to the average rainfall observed. This average forecast corresponds to the ensemble forecasts.

In Figure 5, it can be observed that the CFS model has a good rainfall forecast for the study area, in relation to its seasonal variability. However, it is observed that the model tends to overestimate the forecasts in the periods of little or no rain.

Due to this overestimation, it was necessary to make a correction in the precipitation forecast of the CFS model. Initially, a climate correction was made using the linear model and the model proposed by Wood et al. (2004), here referred as PDF. Subsequently, the different months of each year were corrected by the same methodology used in the climate forecast.

As shown by Fernández Bou et al. (2015) and Silva et al. (2018), in order to improve climate forecasts, it is necessary to make a climate correction on the rain forecast results of the models. Using the PDF and Linear methods, this correction was performed based on a comparison between the predicted and observed average rain in a given period. In this study, the period for climate correction was from 1982 to 2010 - the available period of precipitation forecast data.

Figure 6 shows the comparison between the climate forecast obtained for the CFS and the average rainfall observed at the Três Marias drainage area. It can be observed that the CFS model tends to overestimate the forecast for the study area, with more or less intensity, in all months of the year, maintaining the seasonality, which is important.

Figure 7 shows the application of the PDF methodology for the Três Marias drainage area. Table 1 shows the coefficients obtained. Therefore, it can be observed that the values obtained tend to be close, with the average of the period from January to December being of 0.70 for both cases, with no tendency to be above one or the other coefficient. However, in the months of May, July, August and September, these coefficients change to values with larger differences.

The monthly correction of the rain predicted by the CFS model was also carried out through the PDF methodology.
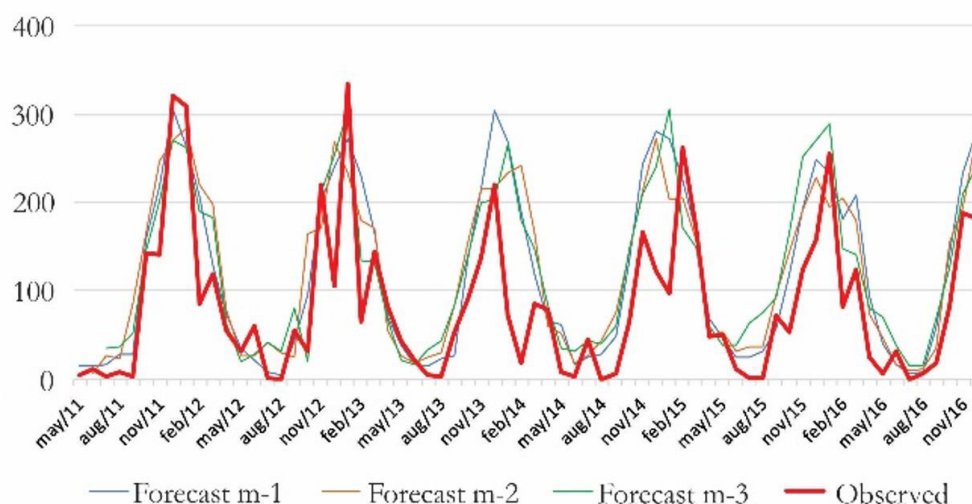

Figure 5. Comparison of predicted rainfall by the CFS model with 1, 2, and 3 months in advance, compared to the average observed rainfall in the drainage area of the Três Marias reservoir.

The months of the year were grouped into two sets of data: wet (with data from November to April) and dry (from May to October). This grouping was carried out because only a few years exhibited information from the CFS, with the month-to-month separation probably leading to lower results than those grouped in the wet and dry periods. The results of this analysis are shown
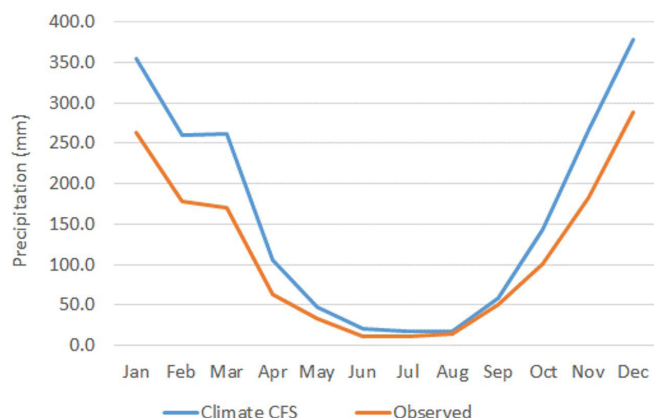

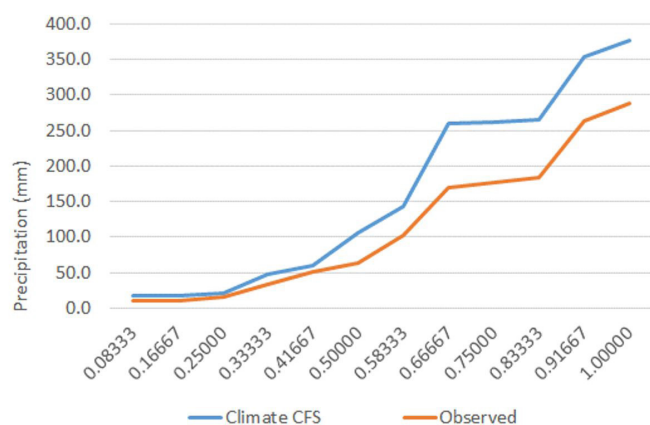**Figure 6.** CFS climate x average rainfall at the Três Marias drainage area, from 1982 to 2010.

Figure 7. PDF analysis, comparison of CFS climate x average rainfall at the Três Marias drainage area, from 1982 to 2010.

Table 1. Relationship between the climate and the average rainfall at the Três Marias area, Linear Coefficient and PDF.

Três Marias	CFS Climate	Verif	Clinear	Cpdf
Jan	354.0	263.8	0.75	0.62
Feb	260.1	177.6	0.68	0.63
Mar	261.6	169.6	0.65	0.70
Apr	106.4	63.0	0.59	0.72
May	47.2	33.9	0.72	0.85
Jun	21.3	11.2	0.52	0.59
Jul	17.6	10.5	0.59	0.71
Aug	16.9	14.9	0.89	0.65
Sep	59.3	50.5	0.85	0.68
Oct	142.9	101.7	0.71	0.69
Nov	265.7	183.5	0.69	0.75
Dec	377.9	288.2	0.76	0.76
Average	**160.9**	**114.0**	**0.70**	**0.70**

below, with the data originating from the climate correction by the PDF methodology and the Linear Coefficient.

As established in the methodology, the frequency curves were then elaborated and the precipitation was corrected by replacing these values with the observed value, both at the same frequency. As an example, Figure 8 shows the results obtained from the climate corrected by the PDF methodology.

Table 2 shows a summary of the CFS model correction results. The best average forecast obtained was the one that had the PDF climate correction.

Corrections for the wet and dry periods were also made for all 25 members of the CFS model. The results obtained show certain changes in the forecasts carried out for each day, by each of the 25 members of the model, though without any great changes from one member to another.

Climate variables analysis

As previously discussed, a statistical analysis was performed considering the climate variables, average rainfall and basin runoff. A monthly average data from 1982 to 2017 was used, which corresponds to the available meteorological data. It should be emphasized that this evaluation aims to define which variables are best related to the natural flow of the Três Marias reservoir, which is the variable to be predicted in this study.

In order to better characterize the order of magnitude, Table 3 presents the average, minimum, maximum and standard deviations of the variables under analysis.

Figure 9 shows the relationship between the average rainfall of Três Marias and all climate variables: without a temporal lag (t), with a time lag of 1 month (t-1), 2 months (t-2) and 3 months (t-3). For most variables, it was observed that, in absolute terms, an increase in the time lag causes an upsurge of the Pearson correlation of the variables. None of the lags showed a good Pearson correlation between SOI and the Três Marias average rainfall, as well as the PDO (Pacific Decadal Oscillation). At time t, there is a positive Pearson correlation with SATL and Nino 1+2. At the same time t, the negative Pearson correlation of the Três Marias average rainfall for variables NATL, Nino 4 and Nino 3.4 was highlighted. In relation to the Pearson correlation with the Três Marias average rainfall in other lags, a high Pearson correlation is still observed at times t-1 (0.58) and t-2 (0.34). Regarding the Pearson correlation with the natural flow of Três Marias, a high Pearson correlation (0.74) was observed at time t, reducing t-1, t-2, and t-3 times, as expected.

Figure 10 shows an analysis of the relationship between the natural flow of Três Marias and the various climate variables in analysis: without a temporal lag (t), with a time lag of 1 month (t-1), 2 months (t-2) and 3 months (t-3). The variables SOI, PDO,

Table 2. Results obtained with the correction of precipitation, Três Marias area.

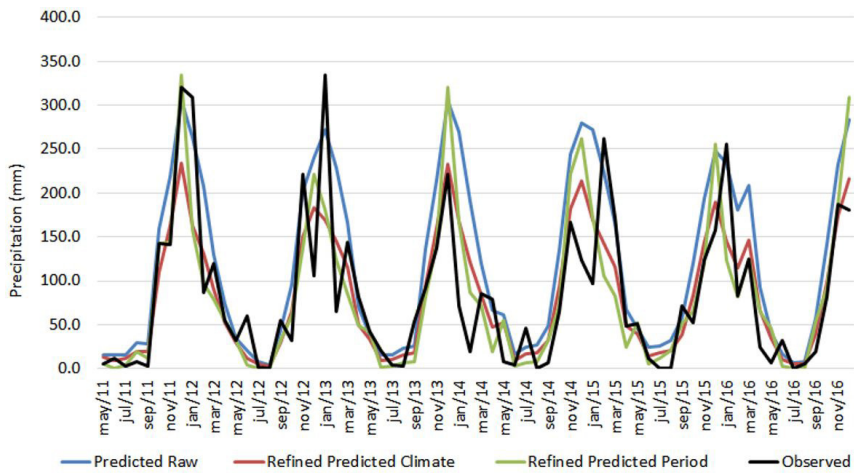| Deviation (mm)	Predicted Raw	Refined Predicted Climate	Refined Predicted Period
PDF	47.2	**34.5**	37.5
Linear Coef.		35.3	36.8

Figure 8. Comparison of predicted, corrected predicted and observed precipitation.
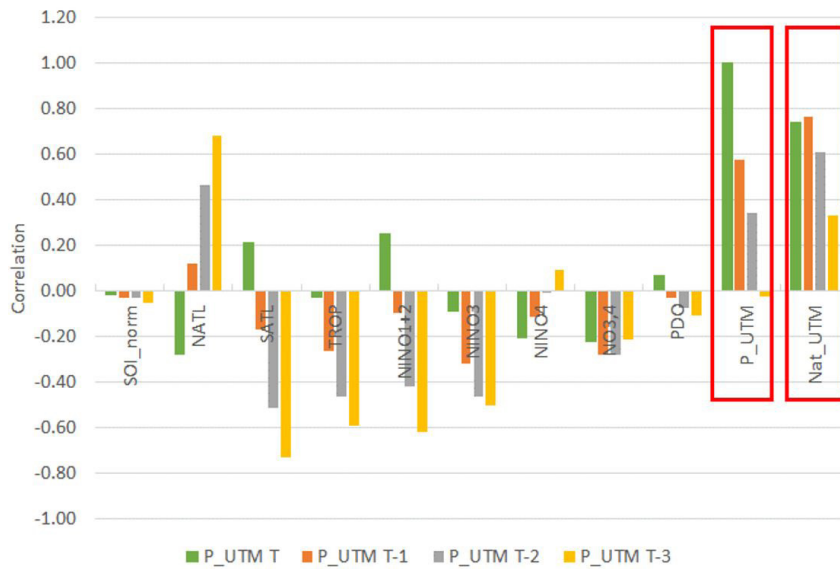

Figure 9. Pearson correlation of the Três Marias average rainfall with the climate variables for t, t-1, t-2 and t-3 lags.
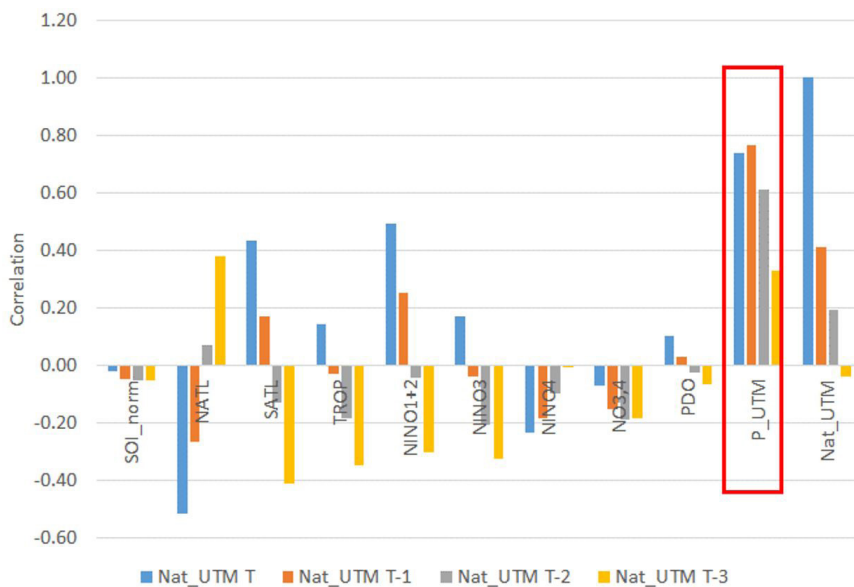

Figure 10. Pearson correlation of the natural flow of Três Marias with the climate variables for t, t-1, t-2 and t-3 lags.

and Nino 3.4 do not exhibit a good Pearson correlation with the natural flow of Três Marias in none of the lags.

At time t, this study highlights the positive Pearson correlation of this flow with the variables SATL and Nino 1+2. At the same time t, the negative Pearson correlation with the NATL and Nino 4 variables was highlighted. Regarding the Pearson correlation with the natural flow of Três Marias in other lags, high Pearson correlations at t-1 (0.67) and t-2 (0.37) were observed. With respect to the Pearson correlation with the average rainfall at Três Marias, the values are high for all the lags analyzed, mainly t, t-1 and t-2, as expected.

It should be noted that, based on the results obtained in the Pearson correlation analysis of the Três Marias average rainfall, the flows at Três Marias, the variables SOI_norm, Nino 4 and PDO presented a low Pearson correlation and were not used as inputs in the flow prediction process.

A Principal Component Analysis (PCA) was performed with the series under analysis. Five components were required to explain 94.9% of the series variance. The PC1 component explains 45.18% and the PC2 explains 32.10% of the total series variance.

Table 3. General information of the variables under analysis.

Variable	Average	Minimum	Maximum	Standard deviation
SOI_norm	0.0	-3.6	2.9	1.0
NATL	26.9	24.8	28.9	1.0
SATL	25.0	22.3	27.9	1.5
TROP	27.8	26.8	29.3	0.5
NINO1+2	23.3	18.6	29.2	2.4
NINO3	26.0	23.0	29.1	1.4
NINO4	28.6	26.4	30.3	0.8
NO3,4	27.1	24.3	29.6	1.1
PDO	-0.2	-3.1	2.4	1.1
P_UTM	108.8	0.0	508.1	107.8
Nat_UTM	655.5	40.0	3744.0	631.7

Where: SOI_norm, NATL, SATL, TROP, NINO1+2, NINO3, NINO4, NINO3,4, PDO represents the meteorological variables; P_UTM represents the rainfall in the area and Nat_UTM, the flow.

Figure 11 shows the distribution of these 5 factors that explain 94.9% of the series variance. It can be observed that the rain variables are mostly explained by components 1, 2, and 3. In terms of climate variables, principal component 1 is mainly present in the variables NATL, SATL, TROP, Nino 1+2, Nino 4 (all with a large proportion) and Nino 3.4 (small proportion).

Principal component 2 is also present in all climate variables, with the variables TROP, Nino 1+2, Nino 4 and Nino 3.4 exhibiting the largest proportions of this variable. Component 3 is mainly present in the NATL and Nino 3.4 variables.

Regarding flows, it can be observed in Figure 11 that they are mostly explained by principal components 1 and 4, although all contain explanatory factors. Principal component 4 is mainly present in the SATL and Nino 3.4 variables.

Flow forecast models and scenario generation

a) Model calibration

a.1) SMAP

The evaluation of the SMAP methodology for the natural flow of Três Marias is shown below. The data from January 1987 to April 2011 were used in the calibration, with the period from May 2011 to December 2016 used for validation (with the rain forecast by the CFS model). Figure 12 shows the daily observed and calculated hydrograms and the hydrograms of the basic flow from the underground reservoir of the model, as well as the series of observed precipitation accumulated in 24 hours, obtained from the average rain weight in the sub-basin considering the weights adjusted by the optimization routine.

The average calculated flow was of 645 m³/s, and the observed flow was of 652 m³/s, showing an average volume deviation of 1.1%. Table 4 shows the main parameters obtained in the calibration of the SMAP.

Table 5 presents the results regarding the metrics used for the performance evaluation of the calibrated model.
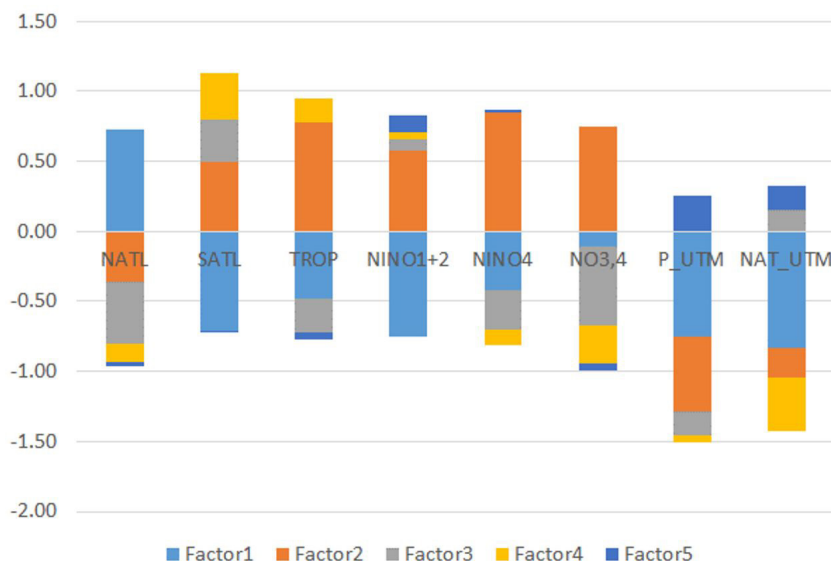

Figure 11. Principal component analysis factors by variables.

a.2) Neural Network

The application of the rainfall-runoff models based on the Neural Networks technique for the prediction of the natural flow of Três Marias are shown ahead. The calibration was carried out with the observed rainfall and an evaluation of the performance of the model was performed using the rain predicted by the CFS model for the period from May 2011 to December 2016, due to several correction methodologies of this precipitation forecast.

Table 6 shows the simulated input variables in the wet period for the flow prediction for Três Marias, in which t corresponds to the variable on the forecast day, t-1 with a one-month lag, t-2 with a two-month lag, and so on.

The list of these variables with the respective acronym adopted in this study are presented as follows:

- Climate variables: North Atlantic Temperature (NATL), South Atlantic Temperature (SATL), region temperature as shown by Pinto et al. (2006): temperature in the tropical region (TROP) and temperature in Pacific regions (Nino 1+2, Nino 3, Nino 4, Nino 3.4);
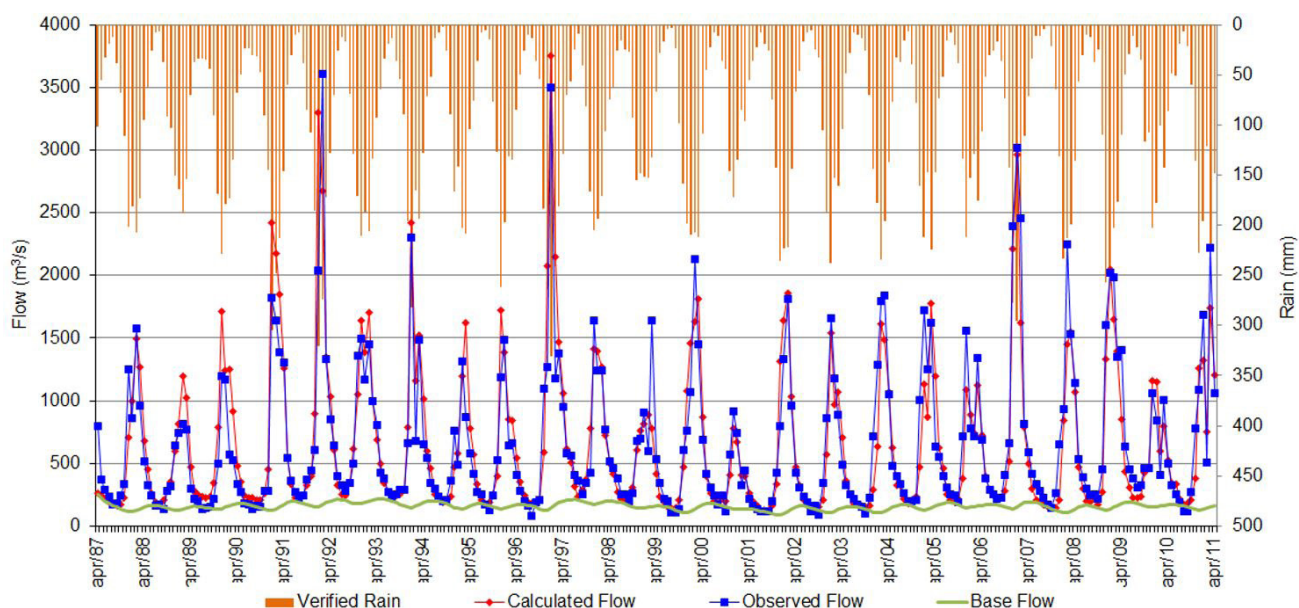
Table 4. Main parameters obtained in SMAP calibration.

Str	K2t	Capc	K
524	5.21	7.80	6.0

Table 5. Performance index found in the calibration.

MAPE	RMSE	MAE	Nash	Nash-Log	R2
24.7%	237.5	144.5	0.84	0.89	0.85

Table 6. Input variables considered in the MLP Neural Network for the wet period.

Case	Number Variables	Variables								
Simul1	8	NATL t-1	SATL t-1	TROP t-1	NINO1+2 t-1	NINO3 t-1	NINO4 t-1	NO3,4 t-1	P_UTM t-1	
Simul2	18	NATL t-2	NATL t-1	SATL t-2	SATL t-1	TROP t-2	TROP t-1	NINO1+2 t-2	NINO1+2 t-1	NINO3 t-2
		NINO3 t-1	NINO4 t-2	NINO4 t-1	NO3,4 t-2	NO3,4 t-1	P_UTM t-2	P_UTM t-1	Nat_UTM t-2	Nat_UTM t-1
Simul3	4	P_UTM t-2	P_UTM t-1	Nat_UTM t-2	Nat_UTM t-1					
Simul4	27	NATL t-3	NATL t-2	NATL t-1	SATL t-3	SATL t-2	SATL t-1	TROP t-3	TROP t-2	TROP t-1
		NINO1+1 t-3	NINO1+2 t-2	NINO1+2 t-1	NINO2 t-3	NINO3 t-2	NINO3 t-1	NINO3 t-3	NINO4 t-2	NINO4 t-1
		NO3,4 t-3	NO3,4 t-2	NO3,4 t-1	P_UTM t-3	P_UTM t-2	P_UTM t-1	Nat_UTM t-3	Nat_UTM t-2	Nat_UTM t-1
Simul5	12	SATL t-3	SATL t-2	SATL t-1	NINO1+1 t-3	NINO1+2 t-2	NINO1+2 t-1	P_UTM t-3	P_UTM t-2	P_UTM t-1
		Nat_UTM t-3	Nat_UTM t-2	Nat_UTM t-1						

Figure 12. SMAP model calibration for the Três Marias area.

- Average rainfall in the drainage area of the Três Marias reservoir (P_UTM);

- Natural flow at the Três Marias reservoir (Nat_UTM).

Table 7 shows the deviations of the flow forecasts one month ahead for the Três Marias reservoir. It can be observed that simulation 5 provided the best results, which consists of a selection of the best correlated climate variables, of the average rainfall at the reservoir and of the natural flow at Três Marias, all with up to 3 months of lag,. Comparing the simulations, the effect of the incorporation of the natural flow for Três Marias is highlighted. It is observed that simulation 1, which does not have the natural flow at Três Marias as an input variable, had very low-quality results. Comparing simulations 2 and 4, it was noticed that the addition of time t-3 in the simulation variables is important. Comparing the simulations from 1 to 5, it was observed that simulation 5 best represented the predicted flows, having the best performance indices.

Figure 13 shows the predicted vs. observed graph of simulation 5 (data used in the Neural Network), in which no error is observed in the forecast lag. In general, it is also observed that the simulation represented well the low flows.

In order to evaluate the contribution of the rainfall forecast, the climate predictions of the CFS model were added in the predictions of flows one month ahead for the Três Marias reservoir. These forecasts were carried out for the climate correction (average of the period from 1982 to 2010) and then by the monthly correction. Considering that no systematic error was observed in the rainfall forecasts, the observed rain of the given month, representing the perfect rainfall, was initially used in the calibrations. With the addition of this variable in simulation 5, the results shown in Table 8 (simulation 6) and represented in Figure 14 were obtained.

Comparing the results of simulations 5 and 6, an improvement in the forecasts can be clearly observed, as a result of the incorporation of the observed rainfall in m+1 in the simulations. For instance, MAPE decreased from 39.2% to 21.8%, and the Nash coefficient increased from 0.73 to 0.89. Comparing

Figures 13 and 14, a substantial improvement was observed mainly in the forecast of high flows, such as in the wet period of 1997.

In order to further improve flow forecasts, following the methodology for a daily flow forecast shown by Gomes (2006), two Neural Networks were adjusted: one for the wet period (November to April) and another for the dry period (May to October). Table 9 and Figure 15 show the results obtained through simulation 6. Comparing simulations 6 and 7, it should be noted that the latter, with separate calibrations for the wet and dry periods, showed an improvement of 5.8% in the MAPE, 31 m³/s in the RMSE, 29 m³/s in the MAE, 0.03 in the Nash coefficient and 0.02 in the Nash-Log coefficient.

b) Evaluation of the Models

The results were subsequently evaluated, replacing the observed rainfall data at time m+1 (subsequent month) by the forecast rainfall data. These simulations were performed in the

Table 7. Results obtained for the one-month lag forecasts for the Três Marias reservoir.

Simulation	MAPE	RMSE	MAE	Nash	Nash-Log
Simul1	42.6	384	253	0.64	0.75
Simul2	34.3	360	219	0.69	0.80
Simul3	38.0	378	240	0.65	0.79
Simul4	35.7	331	211	0.73	0.80
Simul5	39.2	333	218	0.73	0.79

Table 8. Results of simulation 6 that incorporate the observed rainfall.

Simulation	MAPE	RMSE	MAE	Nash	Nash-Log
Simul6	21.8	211	133	0.89	0.92

Table 9. Results of simulation 7, with one network for the wet period and another for the dry period.

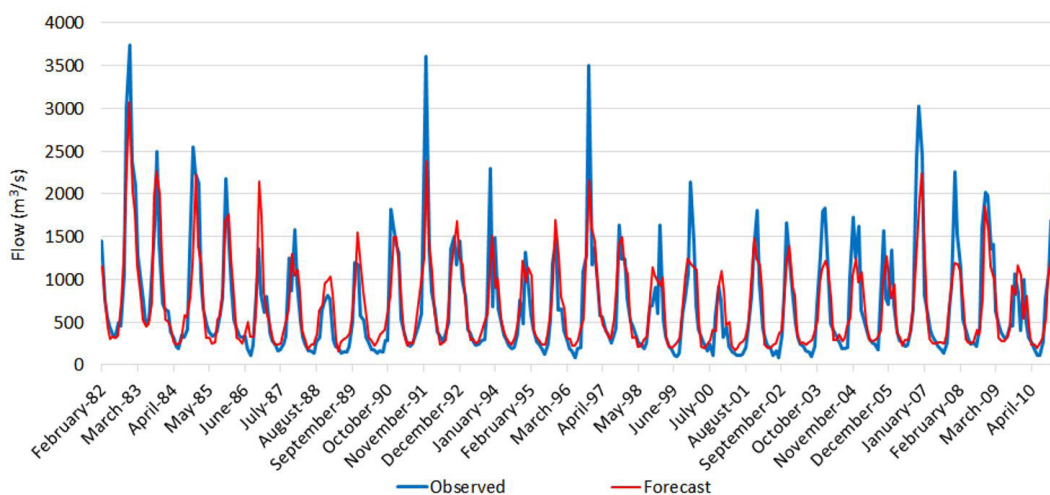| Simulation	MAPE	RMSE	MAE	Nash	Nash-Log
Simul7	16.0	180	103.7	0.92	0.94

Figure 13. Observed flows x predicted flows by simulation 5 using NN (Neural Networks).
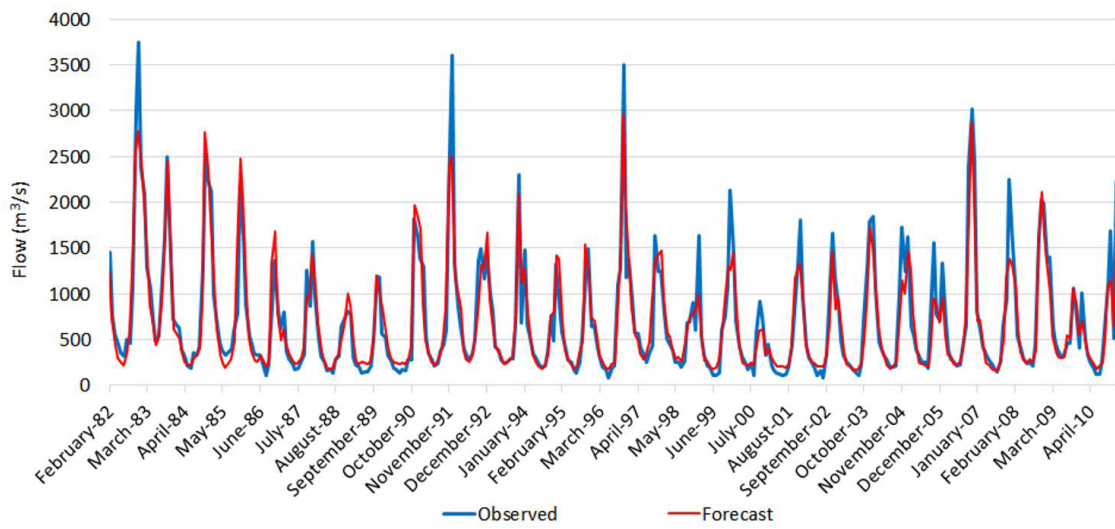
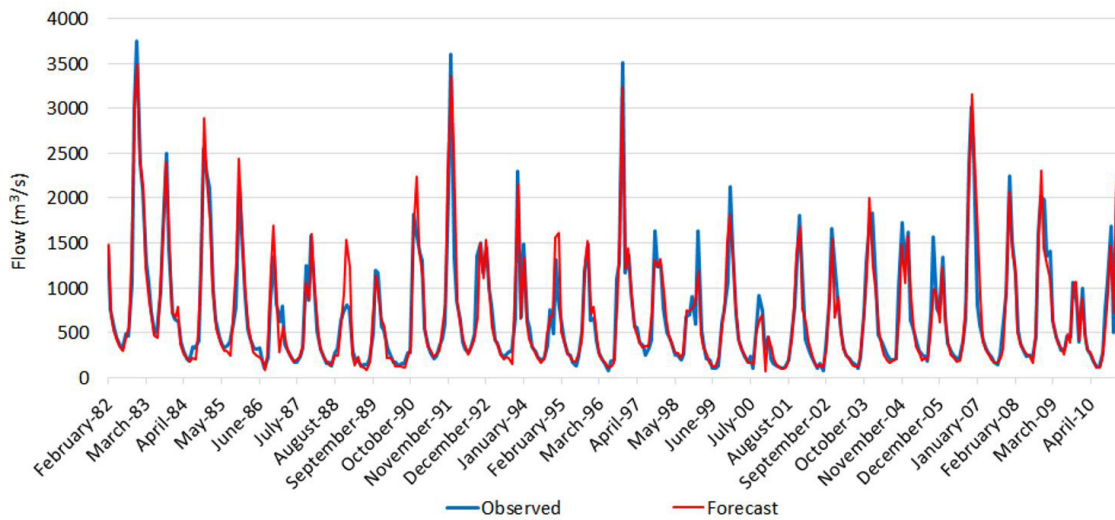**Figure 14.** Observed flows x predicted flows by simulation 6 that incorporate the observed rainfall.

Figure 15. Observed flows x predicted flows by simulation 7 that incorporate observed rainfall.
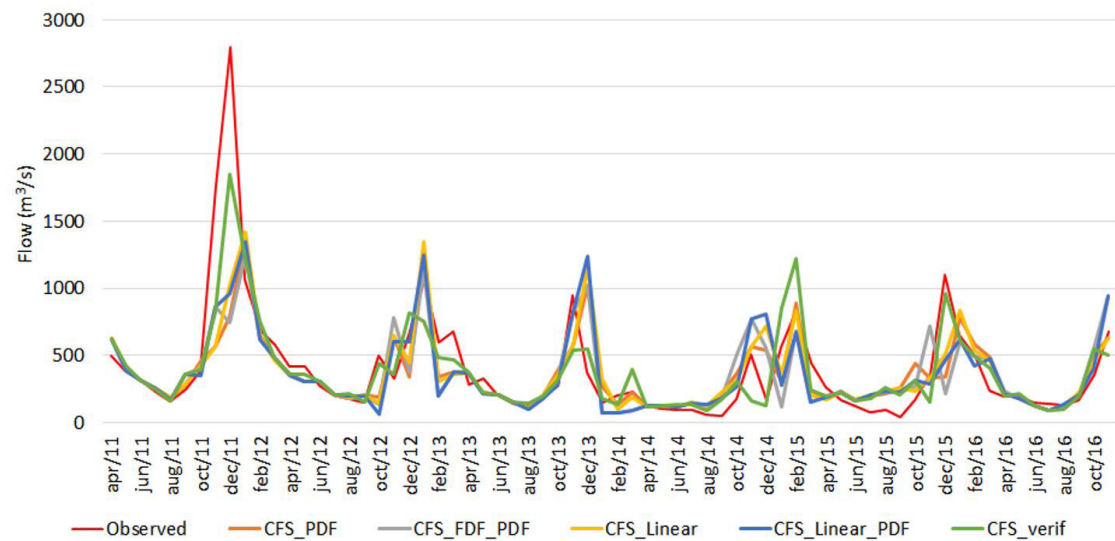

Figure 16. Results of simulation 7 incorporating rain forecasts of the CFS model (ensemble average).

period from April 2011 to December 2016, the available period of the CFS forecasts (test period).

Figure 16 shows the results of these simulations, being observed that the simulations with a climate correction by the Linear method (Simul7C), except for the MAPE index, had slightly better results than the others. By analyzing this Figure, it can be observed that none of the simulations were able to make a good prediction for the low flows in the data of the wet period, nor the simulation with the observed rain, showing the difficulty of the calibration for this type of forecast. However, in the dry season network, low flow rates were easily predicted by all models. This same graph shows that there is no phase error in the forecasts, that is, no time lag is observed.

A simulation was also separately carried out for all 25 members of the CFS model, with all being corrected with the same methodology used for calculating the average ensemble. As an example, Figure 17 shows the forecasts obtained based on the rain predicted by the CFS corrected by the PDF/PDF methodology (climate correction with the PDF methodology and monthly correction with the PDF methodology).

In Figure 17, it can be observed that the results of flow forecast with the Neural Networks vary considerably in some cases, depending on the methodology adopted for the correction of the precipitation forecast. For instance, in the forecast performed for January 2013, the observed flow was of 1059 m³/s and the forecast flow varied from 839 to 1913 m³/s with rain correction by the PDF/PDF methodology; from 594 to 1069 m³/s using only climate correction with the PDF methodology; and from 872 to 1957 m³/s using the rain correction with the Linear/PDF methodology.

As previously shown, in parallel to the application of Neural Networks, the SMAP methodology was evaluated for the natural flows of Três Marias.

With this calibration, this study made an assessment of the model's performance using the rainfall predicted by the CFS model for the period from May 2011 to December 2016, using several methodologies to correct this forecast precipitation. Figure 18 shows the results found.

The purpose of this assessment was to evaluate the range of flows generated by the SMAP model, which was compared
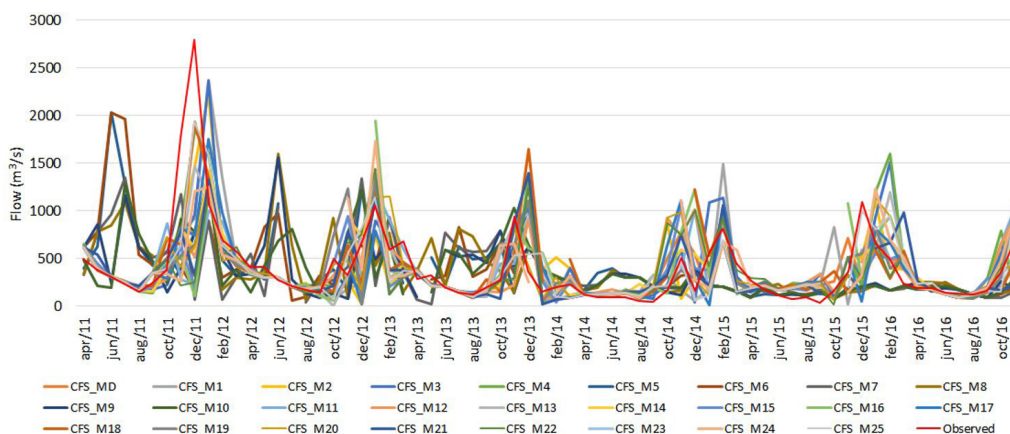

Figure 17. Outflow prediction with NN (Neural Networks) using the rain forecasts of the CFS model, PDF/PDF correction.
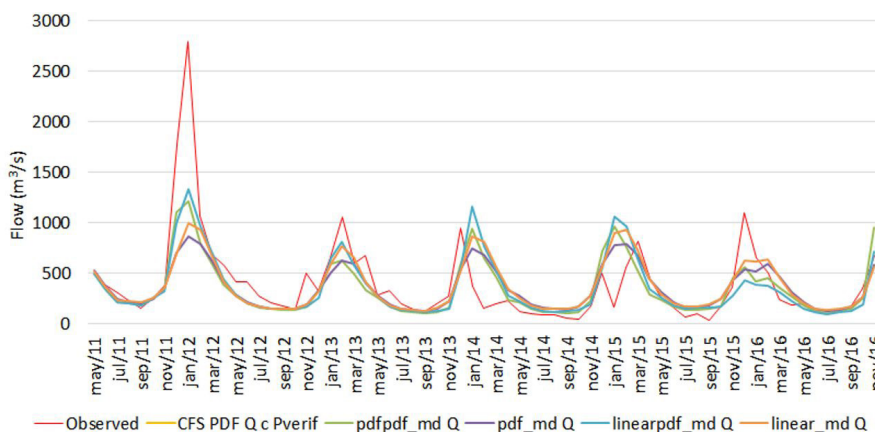

Figure 18. Application of predicted rain by the CFS for the evaluation of the SMAP model for the Três Marias area. Where: Qverif represents the observed flow; Q c Pverif represents the predicted simulated flow with the observed rainfall (only for the comparison with the previous simulations in the same data period); pdfpdf_md Q represents the predicted flow with the average rainfall of the ensemble and with both climate correction and monthly correction by the PDF method; pdf_md Q represents the predicted flow with the average rain of the ensemble and with climate correction by the PDF method; linearpdf_md Q represents the predicted flow with the average rain of the ensemble and with climate correction by the Linear method and monthly correction by the PDF method; linear_md Q represents the predicted flow with the average rain of the ensemble and with climate correction by the Linear method.

to the flows obtained in the previous item, using the Neural Networks methodology.

Figure 19 shows the range of flows generated by the SMAP model and by the Neural Network model, compared to the observed flow. By analyzing the data that generated this Figure, it is observed that in 23 cases, which represent 34%, the observed flows are outside the range of flows predicted by the Neural Networks model. In 52 cases, which represent 76%, the SMAP forecasts failed to reproduce the range. In 26 cases, which represent 38%, SMAP minimum flows represent smaller values than those obtained with Neural Networks models. In 5 cases, which represent 7%, SMAP maximum flows represent higher values than those obtained with the Neural Networks models. It is worth pointing out that in only 5 cases (7%), the predicted flow by the Neural Networks model did not present values in the observed range, although the SMAP model would have improved the accuracy of this range – in almost all cases, the flow would have reduced its range. In other words, in only 7% of cases, SMAP would be able to improve the flow forecast performed by the Neural Networks model.

Table 10 and Figure 20 show the comparison of the average range generated by SMAP and Neural Networks predictions. Although the main objective of this study is not to correct a prediction but to locate the probable range of its occurrence, it should be noted that the SMAP model performed better than the Neural Networks model, when comparing the average flow range predicted in relation to the observed flow.

In Figure 21, it is observed that in the range generated by the GEVAZP model, the flows produced by the Neural Networks model are already present, while in the range generated by the latter, the SMAP flows are present.

Table 10. Deviations from the average flow range predicted by SMAP and Neural Networks models for the period from May 2011 to December 2016.

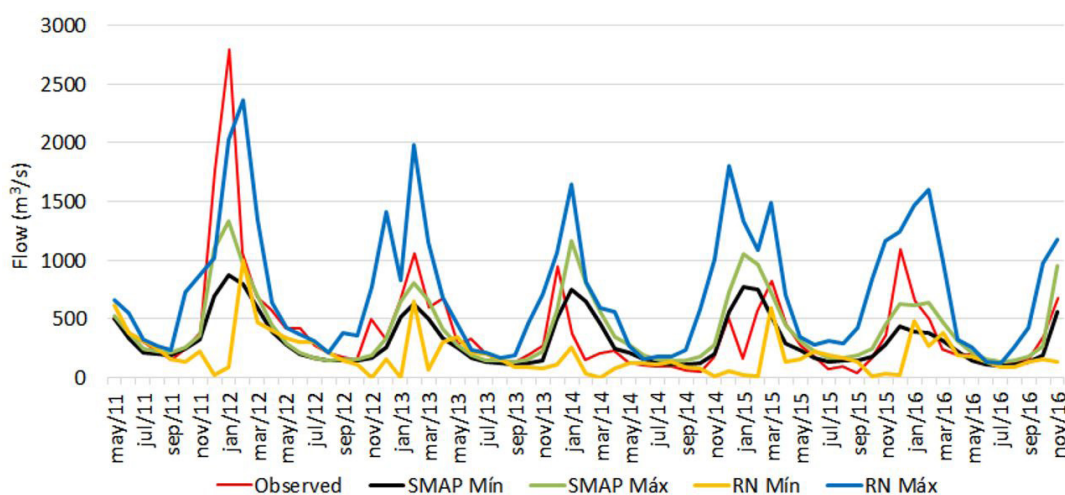| Simulation	MAPE	RMSE	MAE	Nash	Nash-Log
Smap	51.8	303	157	0.49	0.65
Neural network	73.5	339	198	0.36	0.52

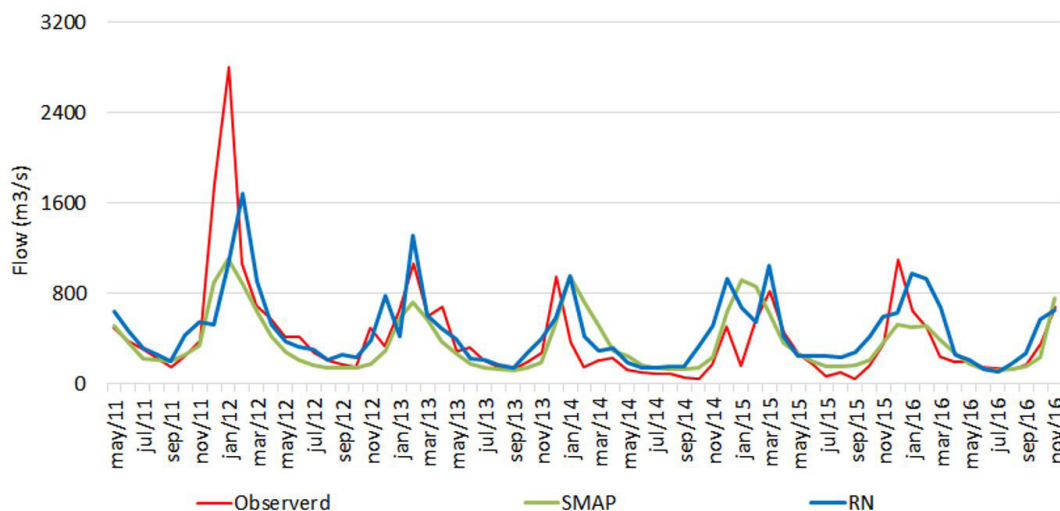**Figure 19.** Evaluation of the flow range of the flow predictions for the Três Marias reservoir.

Figure 20. Comparison of the average ranges predicted by SMAP and Neural Networks models.

In Figure 22, it can be observed that in the range generated by the GEVAZP model, the flows produced by the Neural Networks model are already present. However, in the range generated by the latter, the SMAP flows (the flow range predicted by SMAP was from 628 to 810 m³/s and the flow range predicted by the Neural Networks model was from 649 to 1987 m³/s) are no longer present.

In Figures 21 and 22, it is interesting to note that the runoff scenarios predicted by GEVAZP contain flows well above the observed ones and those predicted by the different rainfall scenarios of the CFS, especially in the predicted flows for the wet period, as is the case of Figure 22 with predicted flows for February 2013.

Figures 21 and 22 show that GEVAZP generates several series out of the likely range. The NEWAVE rounds, in determining future cost, visit all of these series generated by GEVAZP. If this GEVAZP range is reduced, the NEWAVE future cost function will not visit these series out of schedule, enabling it to deal with uncertainty (Cataldi et al., 2012). This lower opening of the scenario decreases the chance that the future cost function will go somewhere in the domain that has no chance of being accomplished.

In addition, the flows generated by the Neural Networks model, SMAP, were compared and verified with the flows generated by Gevazp, for January 2015 until December 2016, as observed in Figure 23. Thus, it can be seen that Gevazp shows scenarios that cannot be verified.
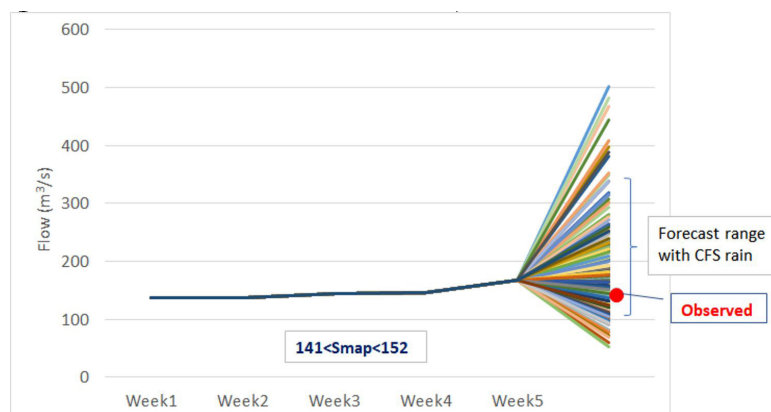

Figure 21. Scenarios predicted by GEVAZP, along with the range of the CFS model and observed flow, PMO September 2012 (predicted flows for October 2012) with SMAP range.
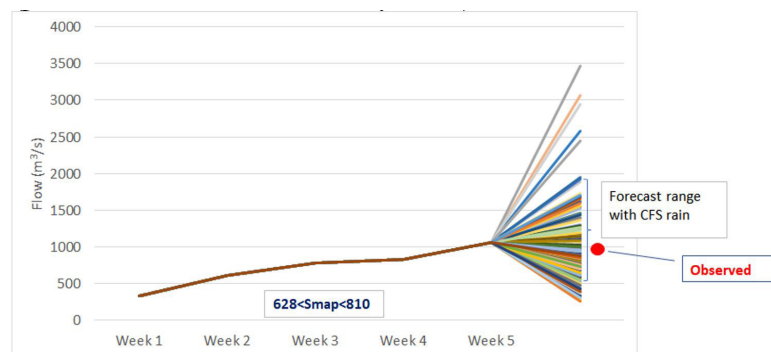

Figure 22. Scenarios predicted by GEVAZP, along with the CFS model range and observed flow, PMO January 2013 (predicted flows for February 2013) with SMAP range.
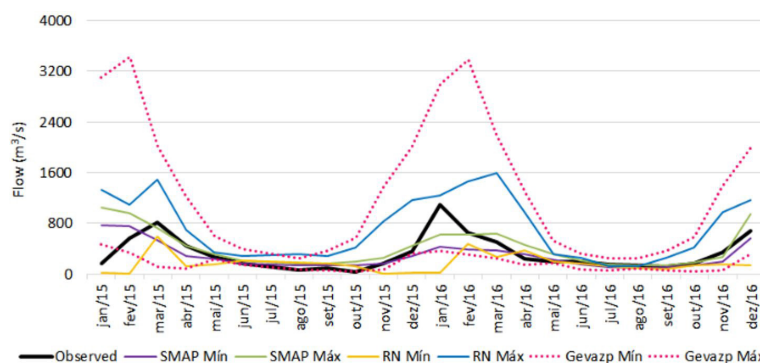

Figure 23. Scenarios predicted by GEVAZP, with the Neural Networks model range, SMAP and the observed flow, for the period from Jan./2015 to Dec./2016.

CONCLUSION

- The CFS climate forecasts tend to overestimate the predicted rain mainly in the months of little or no rain;

- The corrections made by the CFS model for the ensemble average, led to better results when compared with the climate correction and the monthly correction by the PDF methodology. It is worth noting that, in some cases, the results of flow forecast vary considerably depending on the methodology used to correct the precipitation forecast;

- Simulations with Neural Networks showed that, for the Três Marias reservoir, 17% of the observed flows would be outside the range generated by the model. In these cases, the flow generated by SMAP would improve the accuracy of this range in 3% of the cases. It should be noted that the Neural Networks model performed well in this flow range, considering that even when it did not reach the given range, it generated values very close to it. Although the main objective of this study is not the correction of a prediction, but the probable range of its occurrence, it should be noted that the SMAP model performed better than the Neural Networks model, when compared to the average flow range predicted in relation to the observed flow;

- There was a clear improvement in the flow predictions with the incorporation of the rain observed one month ahead in the simulations, mainly in the forecast of high flows, as for the wet period of 1992;

- The study showed that it is possible to safely reduce the range predicted by the GEVAZP model, which prevents the DECOMP and NEWAVE energy models from visiting these scenarios with very little or no probability of occurrence. It should be noted that the flow forecast for 2 months ahead was evaluated and the results were also promising and do not significantly change the expected range presented in this work;

- The variables NATL, SATL, TROP, Nino 1+2, Nino 3, Nino 4 and Nino 3.4 had a good relationship with the flow and rain variables. However, the SOI and PDO variables did not present a good Pearson correlation with the flows and rainfall of Três Marias;

- In the historical arrangement of the natural flow data of Três Marias with the main correlated variables, which are the temperature in the Nino 1+2 area and in the South Atlantic, it was observed that the up and down cycles of these three variables are similar. It is also observed that the temperature amplitude in the Nino 1+2 region is much higher than the SATL. The natural flow at Três Marias clearly shows the water crisis that has been faced by the São Francisco basin since 2012. However, no great historical variation can be observed in the Nino 1+2 and SATL data.

ACKNOWLEDGEMENTS

REFERENCES

Agência Nacional de Água – ANA. (2004). *Plano de recursos hídricos da bacia do São Francisco* (Resumo Executivo). Brasília: ANA.

Baker, S. A., Wood, A. W., & Rajagopalan, B. (2019). Developing subseasonal to seasonal climate forecast products for hydrology and water management. *Journal of the American Water Resources Association*, *55*(4), 1-14. http://dx.doi.org/10.1111/1752-1688.12746.

Cataldi, M., Braga, R. S., Dias, T. L., Sá, R. V., Rocha, V. F., & Souza Filho, F. A. (2012). Assessment of precipitation climate forecasting of models ECHAM 4.5, CFS, COLA/IRI and CCM3 for hydro basins in southwest Brazil. In *Proceedings of the CLIVAR VAMOS Workshop on Modeling and Predicting Climate in the Americas*. Petropolis: CLIVAR.

Cataldi, M., Machado, C. O., Guilhon, L. G. F., Chou, S. C., Gomes, J. L., & Bustamante, J. F. (2007). Análise de Previsões de Precipitação obtidas com a utilização do modelo ETA como insumo para modelos de previsão semanal de vazão natural. *Revista Brasileira de Recursos Hídricos*, *12*(3), 5-12. http://dx.doi.org/10.21168/rbrh.v12n3.p5-12.

Centro de Pesquisa de Energia Elétrica – CEPEL. (2003). *Manual de referência do Modelo Dessem*. Rio de Janeiro: CEPEL.

Collischonn, W., Allasia, D., Silva, B. C., & Tucci, C. M. (2007). The MGB-IPH model for large scale rainfall-runoff modeling. *Hydrological Sciences Journal*, *52*(5), 878-895. http://dx.doi.org/10.1623/hysj.52.5.878.

Cossich, W., Cataldi, M., & Rotunno Filho, O. C. (2015). Avaliação do desempenho da geração de cenários de afluências em reservatórios utilizando previsões de precipitação por conjunto. *Ciência e Natura*, 37, 55-62. http://dx.doi.org/10.5902/2179460X16215.

Evsukoff, A. G., Cataldi, M., & Lima, B. S. L. P. (2012). A multi-model approach for long-term runoff modeling using rainfall forecasts. *Expert Systems with Applications*, *39*(5), 4938-4946. http://dx.doi.org/10.1016/j.eswa.2011.10.023.

Fan, F. M., Schwanenberg, D., Alvarado, R., Assis dos Reis, A., Collischonn, W., & Naumman, S. (2016). Performance of deterministic and probabilistic hydrological forecasts for the short-term optimization of a tropical hydropower reservoir. *Water Resources Management*, *30*(10), 3609-3625. http://dx.doi.org/10.1007/s11269-016-1377-8.

Fernández Bou, A. S., De Sá, R. V., & Cataldi, M. (2015). Flood forecasting in the upper Uruguay River basin. *Natural Hazards*, *79*(2), 1239-1256. http://dx.doi.org/10.1007/s11069-015-1903-7.

Gomes, L. F. C. (2006). *Previsão de vazões naturais diárias ao reservatório de Três Marias usando a técnica de redes neurais*. (Dissertação de mestrado).

Programa de Pós-graduação em Engenharia Civil, Universidade Federal de Pernambuco, Recife.

Jardim, D. L. D. D., Maceira, M. E. P., & Falcão, D. M. (2001). Stochastic streamflow model for hydroeletric systems using clustering techniques. In *Proceedings of the IEEE Porto Power Tech Conference* (Vol. 3). Porto: IEEE.

Lopes, J. E. G., Braga, B. P. F., & Conejo, J. G. L. (1982). SMAP – A simplified hydrological model. In V. P. Singh (Ed.), *Applied modeling in catchment hydrology* (pp. 167-176). Littleton: Water Resources Publications.

Maceira, M. E. P., & Mercio, C. M. V. B. (1997). Stochastic streamflow model for hydroelectric systems. In *Proceedings of the V International Conference PMAPS - Probabilistic Methods Applied to Power Systems*. Vancouver, Canada: IEEE.

Pinto, E. J. A., Naghettini, M., & Abreu, M. L. (2006). Metodologia para previsão sazonal de vazões na bacia do alto rio são francisco, incorporando as estimativas das prováveis trajetórias temporais de precipitação associadas às condições de indicadores climáticos. *Revista Brasileira de Recursos Hídricos*, *11*(4), 121-135. http://dx.doi.org/10.21168/rbrh.v11n4.p121-135.

Rocha, V. F., Braga, R. G. S., França, E. H. F., & Cataldi, M. (2015). Uso de informação meteorológica no planejamento da operação do Sistema Interligado nacional. In *Anais do XXIII Seminário Nacional de Produção e Transmissão de Energia Elétrica*. Foz do Iguaçu: Grupo de estudos de operação de sistemas elétricos – GOP.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., & Goldberg, M. (2010). *The NCEP climate forecast system reanalysis*. Washington: American Meteorical Society. https://doi.org/10.1175/2010BAMS3001.1.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., & Becker, E. (2014). The NCRMSE climate forecast system, Version 2. *Journal of Climate*, *27*(6), 2185-2208. http://dx.doi.org/10.1175/JCLI-D-12-00823.1.

Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., Van Den Dool, H. M., Pan, H.-L., Moorthi, S., Behringer, D., Stokes, D., Pena, M., Lord, S., White, G., Ebisuzaki, W., Peng, P., & Xie, P. (2006). The NCRMSE climate forecast system. *Journal of Climate*, *19*(15), 3483-3517. http://dx.doi.org/10.1175/JCLI3812.1.

Silva, F. N. R., Alves, J. L. D., & Cataldi, M. (2018). Climate downscaling over South America for 1971-2000: application in SMAP rainfall-runoff model for Grande River Basin. *Climate Dynamics*, 52, 681-696.

Tucci, C. E. M., Collischonn, W., & Fan, F. M. (2015). Hydropower forecasting in Brazil. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, J. C. Schaake. *Handbook of hydrometeorological ensemble forecasting* (pp. 1-22). Berlin, Heidelberg: Springer. http://dx.doi.org/10.1007/978-3-642-40457-3_55-1.

Valença, M. J. S. (2005). *Aplicando redes neurais: um guia completo*. Olinda: Editora Livro rápido.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Deque, M., Ferranti, L., Fucile, E., Fuentes, M., Hendon, H., Hodgson, J., Kang, H.-S., Kumar, A., Lin, H., Liu, G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mastrangelo, D., Maclachlan, C., Mclean, P., Minami, A., Mladek, R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A. W., Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D., Woolnough, S., Wu, T., Won, D.-J., Xiao, H., & Zaripov, R., & Zhang, L. (2017). *The Subseasonal to seasonal (S2S) prediction project database*. Washington: American Meteorogical Society.

White, C. J., Carlsen, H., Robertson, A. W., Klein, R. J. T., Lazo, J. K., Kumar, A., Vitart, F., Perez, E. C., Ray, A. J., Murray, V., Bharwani, S., MacLeod, D., James, R., Fleming, L., Morse, A. P., Eggen, B., Graham, R., Kjellström, E., Becker, E., Pegion, K. V., Holbrook, N. J., McEvoy, D., Depledge, M., Kirkpatrick, S. P., Brown, T. J., Street, R., Jones, L., Remenyi, T. A., Johnston, I. H., Buontempo, C., Lamb, R., Meinke, H., Arheimers, B., & Zebiakd, S. E. (2017). Review potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorological Applications*, *24*(3), 315-325. http://dx.doi.org/10.1002/met.1654.

Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic implications of dynamical and statistical approaches to downscale climate model outputs. *Climatic Change*, *62*(1-3), 189-216. http://dx.doi.org/10.1023/B:CLIM.0000013685.99609.9e.

Wood, A. W., Maurer, E. P., Kumar, A., & Lettenmaier, D. P. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research*, 107(D20), 1-15. https://doi.org/10.1029/2001JD000659.

Authors contributions

Luana Ferreira Gomes De Paiva: Data collection, analysis of information, simulations and general writing of the paper.

Suzana Maria Montenegro: General revision of the text in Portuguese and English, general orientation of the paper structure, analysis and simulations, final revision.

Marcio Cataldi: General revision of the text in Portuguese and English, general orientation of the article structure, analysis and simulations, final revision.