INFLUENCE OF CLIMATE CHANGE ON HYDROPOWER PLANTS ELECTRICITY PRODUCTION

Răzvan ROMAN[1], Matteo DE FELICE[2], Bogdan POPA[3]

The energy sector can be affected by changing climate conditions through many ways, either for the better or for the worse. Renewable energy endowments refer to a flux of energy, which is closely related to climate conditions. For this reason, it can be expected that climate change may affect renewable sources more intensively than fossil one.

The amount of electricity that can be generated from hydropower plants depends not only on the installed generation capacity, but also on the variation in water inflows to the power plants reservoirs.

Keywords: climate change, hydropower.

1. Introduction

Climate change is caused by factors that include oceanic processes (such as oceanic circulation), biotic processes, variations in solar radiations received by Earth, plate tectonics and volcanic eruptions, and human-induced alterations of the natural world; these latter effects are currently causing global warming, and *climate change* is often used to describe human-specific impacts.

Scientists actively work to understand [1] past and future climate by observations and theoretical models. Borehole temperature profiles, ice cores, floral and faunal records, glacial and periglacial processes, stable isotope and other sediment analysis, and sea level records serve to provide a climate record that spans the geologic past. More recent data are provided by the instrumental record. Physically based general circulation models are often used in theoretical approaches to match past climate data, make future projections, and link causes and effects in climate change.

Past precipitation can be estimated in the modern era with the global network of precipitation gauges. Surface coverage over oceans and remote areas is relatively sparse, but, reducing reliance on interpolation, satellite data has been available since the 1970s. Quantification of climatologic variation of precipitation in prior centuries and epochs is less complete but approximated using proxies such as marine sediments, ice cores, cave stalagmites, and tree rings.

Climate change[2] associated with global warming induced by the increase in carbon dioxide and other trace gases in the atmosphere has been focus of a multitude of scientific investigations for over the past two decades. One of the

[1] Eng. graduated from Power Engineering Faculty, University POLITEHNICA of Bucharest, Romania
[2] Energy and Environment Modeling Researcher at ENEA Rome, Italy
[3] Prof., Power Engineering Faculty, University POLITEHNICA of Bucharest, Romania

most important and immediate effects of global warming would be the changes in local and regional water availability, since the climate system is interactive with the hydrologic cycle. Such effects may include the magnitude and timing of runoff, the frequency and intensity of floods and droughts, rainfall patterns, extreme weather events, and the quality and quantity of water availability; these changes, in turn, influence the water supply, power generation, sediment transport and deposition, and ecosystem conservation. Some of these effects may not necessarily be negative, but they need to be evaluated as early as possible because of the great socio-economic importance of water and other natural resources.

Modeling climate[3] and projected changes in climate is a resource-intensive research activity, usually involve supercomputers and a multidisciplinary approach. These disciplines range from socio-economic (scenarios), to computing, physics, chemistry (climate models) and Earth and life sciences (impact models), as well as statistics and probability (analysis). In the European Union there are a finite number of institutes that conduct research into climate and climate change. The ENSEMBLES project represents the first occasion on which this spectrum of researchers was brought together to work with a single purpose.

Since the 1990s, ensemble forecast have been used operationally (as routine forecasts) to account for the stochastic nature of weather processes – that is, to resolve their inherent uncertainty. Starting in 1992 with ensemble forecasts prepared by the European Centre for Medium Range Weather Forecasts (ECMWF) and the National Centers for Environmental Prediction, model ensemble forecasts have been used to help define the forecasts uncertainty and to extend the window in which numerical weather forecasting is viable farther into the future than otherwise possible. The ECMWF model, the Ensemble Prediction System, uses singular vectors to simulate the initial probability density.

Against this background, the European Commission initiated the ENSEMBLES project to help inform researchers, decision makers, businesses and the public by providing them with climate information obtained through the use of the latest climate modeling and analysis tools. The value, and core, of the ENSEMBLES project is in running multiple climate models (*ensembles*); a method known to improve the accuracy and reliability of forecasts. The project output is a range of future predictions assessed to decide which of the outcomes are more probable than the others. This probabilistic information will assist policy makers, at all levels, in determining future strategies to address climate change.

2. Hydrological models

Hydrologic models are simplified, conceptual representations of a part of the hydrologic cycle. They are primarily used for hydrologic prediction and for understanding hydrologic processes. Two major types of hydrologic models can be distinguished: stochastic models and process-based models.

Stochastic models are black box systems, based on data and using mathematical and statistical concepts to link a certain input (for instance rainfall) to the model output (for instance runoff). Process-based models try to represent the physical processes observed in the real world. Typically, such models contain representations of surface runoff, subsurface flow, evapotranspiration, channel flow, but they can be far more complicated. These models are known as deterministic hydrology models.

Research in hydrologic modeling tries to have more global approach to the understanding of the behavior of hydrologic systems to make better predictions and to face the major challenges in water resources management.

ERA-40 is an analysis of the global atmosphere and surface conditions for 45 years, over the period from September 1957 through August 2002. The three-dimensional ERA-40 data set was created consistently with one version of the Integrated Forecasting System of the ECMWF. The model results are based on not only the routine measurements, but also included all available and appropriate observational data. This assimilated observational data are from diverse sources, such as weather stations, radio soundings, ship measurements and since 1970s also satellite measurements. The data set contains all the important atmospheric parameters. It was build for a time period of 45 years with the same model to be consistent.

The analysis was done in an effort to improve the accuracy of historical weather maps and aid in a more detailed analysis of various weather systems through a period that was severely lacking in computerized data. Each analysis represents a state of the model after iteratively adjusting the background towards observations in a way that is optimal, given estimates of the accuracy of the background and observations.

The conventional observations for ERA-40 originate from many sources, reflecting the evolution of the global observing system. Comparing monthly precipitation totals from ERA-40 and the Global Precipitation Climatology Project (GPCP) shows that ERA-40 precipitation is substantially too large only in the tropics, especially over the oceans. ERA-40 precipitation is in much better agreement with GPCP in the extra tropics, not only with respect to the climatologically means but also with the inter-annual variability of monthly totals.

The ENSEMBLES forced with ERA-40 are: C41, CHMI, CNRM, DMI, ETHZ, GKSS, HC, ICTP, KNMI, OURANOS, SHMI and the Ensemble of the models. These models were realized by numerous research facilities in Europe and the aim was to be used in climate analyze.

Numerical models, General Circulation Models (GCMs), representing physical processes in the atmosphere, ocean, cryosphere and land surface, are the most advanced tools currently available for simulating the response of the global climate system to increasing greenhouse gas concentrations. In 2000 was published *The special Report on Emissions Scenarios* (SRES) by the Intergovernmental Panel on Climate Change, where they present a description of the greenhouse gas emissions scenarios which was used to make projections of possible future climate change.

There are 40 different scenarios, each making different assumptions for future greenhouse gas pollution, land-use and other driving forces. Assumptions about future technological development as well as the future economic development are thus made for each scenario. Most of them include an increase in the consumption of fossil fuels; some variations of B1 have lower levels of consumption by 2100 than in 1990. These emissions scenarios are organized into families, which contain scenarios that are similar to each other in some respects. The six families of scenarios discussed are A1FI, A1B, A1T, A2, B1 and B2. The A1 scenarios are of a more integrated world. The A1 family of scenarios is characterized by: rapid economic growth, a global population that reaches 9 billion in 2050 and the gradually declines, the quick spread of new and efficient technologies, a convergent world income and way of life converge between regions (extensive social and cultural interactions worldwide).

The A1B subset is from the A1 family and is main characteristic is that is based on a balanced emphasis on all energy sources. The data that we used in this paper for the predicted flow was from A1B scenario.

The GCM used models are DMI, DMI-ECHAM5, ETHZ, HadRM3Q0, HadRM3Q3, HadRM3Q16, ICTP, KNMI-ECHAM5-r1, KNMI-ECHAM-r3, KNMI-MIROC, MPI, SMHI-BCM, SMHI-ECHAM5, SMHI-HadCM3Q3 and the Ensemble (a model obtained as average of these models).

3. Methodology

To assess the impact of a new rainfall regime on the electricity generation from hydroelectric power plants (HPPs), it is necessary first to project how it would affect the inflows in the reservoir.

The first analysis performed was the annual analysis. This approach was chosen, even though it may be considered simplistic, to see if a simple analysis of data sets can lead to satisfactory results. Data was extracted from ERA-40 models precipitation and temperature.

In the first step a correlation analyzes was made between the analyzed rainfall values and the inflow. For the distribution of flow values to be almost normal in correlation procedure was used logarithmic values of the flow [5].

To correlate these values it was used the 'correcoef' function from MATLAB. This function computes the Pearson correlation coefficient. The Pearson product-moment correlation coefficient is a measure of the correlation (linear dependence) between two variables X and Y, giving a value between "1" and "-1" inclusive. After this result, the next step of this analyze was to implement a bootstrapping and a cross-validation procedure. Bootstrapping is a very common method use in calculating estimators for small datasets. This technique allows estimation of the sampling distribution of almost any statistic using very simple methods. Using bootstrap procedure was estimated the correlation between the inflow and precipitation for the ERA-40 models.

The next step was to apply a cross-validation procedure to obtain a linear model. For this analyze it was also considered the temperature as a parameter, because was tried to see what is its importance. Cross-validation is a technique for assesing how the results of a statistical analysis will generalize to an independent data set.

All this analyze lead to the conclusion that in the future step it will be used data from the nearest point of the hydropower endowment.

For the second step of this analyze was used as source the article [4], where it is presented an analyze made for the Brazilian energy sector focusing on hydropower energy production and liquid biofuels production.

It was analyzed a linear relation between the historical flow and precipitation data obtained from ERA-40 models. The main hypothesis which was taken in consideration is that the variations of the rainfall regime would only impact the mean and variance of the flow. It was assumed also that the flow time series, f have a log-normal distribution. In this case ln$f$ has a normal distribution, N (μ, σ^2).

If it's considered that $y_{it} = (\ln f - \mu)/\sigma$, where $y_{it} \approx N(0,1)$ it was easy to estimate a model that captures the underlying structure of the stochastic process.

Therefore, maintaining the historical mean and variance, the projected flow for the 1971-2000 period would be equal to:

$$f^{1971-2000} = e^{[(y_{it}*\sigma)+\mu]} . \tag{1}$$

Based on the mean and standard deviation for each month, a cross-sectional time series linear regression model was estimated. In order to improve the analysis a dummy variable, DT, was introduced in the regression. This was introduced to obtain the seasonal pattern, especially for the dry period which we considered to be from July to December. In this period, it was considered that the soil is still saturating, therefore an increase in rain has a smaller effect on the runoff than in the January-June period, when the soil is already saturated and more water from rain runs straight to the river flow. To obtain the equations for the monthly mean flow and the monthly standard deviation was used the following relations:

$$\mu_m((\ln f)_m) = \alpha^{av} + \beta_1^{av} * avg(\ln(precip)_m) + \beta_2^{av} * DT * avg((\ln(precip)_m) , \tag{2}$$

$$\sigma_m((\ln f)_m) = \alpha^{sd} + \beta_1^{sd} * sd(\ln(precip)_m) + \beta_2^{sd} * DT * sd(\ln(precip)_m) , \tag{3}$$

in where was considered μ_m((ln f).m) and σ_m((ln f).m) to be the average and the standard deviation for each month of the historical flow, avg(ln(precip)$_m$) and sd(ln(precip)$_m$) represent the average and the standard deviation of the logarithm precipitation for the month m, α^{av},β_1^{av}, β_2^{av} are the estimated regression coefficients for the mean, in particularly β_1^{av} or $\beta_1^{av}+\beta_2^{av}$ being the elasticity of monthly average of natural logarithm of flow, respectively α^{sd},β_1^{sd}, β_2^{sd} are the estimated regression coefficients for the standard deviation, in particularly β_1^{sd} or $\beta_1^{sd}+\beta_2^{sd}$ being the elasticity of monthly standard deviation of natural logarithm of flow.

The relevant regression coefficients for the analysis are the β coefficients. These represent the elasticity of average of logarithm flow respect to average logarithm precipitation and elasticity of the standard deviation. When a month is in the DT period, its elasticity for the mean is equal to ($\beta_1^{av}+\beta_2^{av}$), in the other case just β_1^{av}. The same goes for the elasticity for standard deviation.

This analyze was made for the ERA-40 models, the Ensemble, the three weather stations and the average data from the weather stations and also for the GCM models. The goal was to have as result the regression coefficients, the R^2 value, the T-statistic values and the significance of each regression coefficient.

With the relation (1) it is possible to create a flow time series for which was build histograms, comparing the histogram of the historical flow data with the projected flow. Some models could recreate the original histogram pattern. In this way it was assessed this approach and could establish for the next step that the data from the weather stations is very reliable related to the other data that we have and we can use it in the further step.

The next step of the analysis was to compute the mean and the standard deviation for the predicted flow for the period between 2071 and 2099. As data the GCM models were used and the average data from the weather stations. The used relations are the following:

$$\mu_m^{2071-2099} = \mu_m * \left\{1 + \left[E^{av} * \left(\frac{avg(precip^{2071-2099})}{avg(precip_m)} - 1\right)\right]\right\}, \tag{4}$$

$$\sigma_m^{2071-2099} = \sigma_m * \left\{1 + \left[E^{sd} * \left(\frac{sd(prec^{2071-2099})}{sd(prec_m)} - 1\right)\right]\right\},$$

where $\mu_m^{2071\text{-}2099}$ and $\sigma_m^{2071\text{-}2099}$ is the mean and the standard deviation for the flow for each month, μ_m and σ_m is the mean and the standard deviation from the historical flow data, E^{av} and E^{sd} represent the elasticity as described early and the relative difference between historical and projected precipitation average and standard deviation.

4. Vidraru hydropower development – study case

The Arges River is located in the South of the Carpathians and forms one of the most important hydrographical basins of Romania having an area of 12,500 sqr. Km. Commissioned on December 9, 1966 the Vidraru hydropower plant has since used the hydropower potential of Arges on a 28 km long sector between Cumpana and Oiesti, developing a 324 m total head.

The development scheme represents the implementation of some modern concepts through the catchment of a 745 km^2 hydrographical basin. Therefore there was an average intake flow of 19.7 m^3/s, flow provided both from Arges River – 7.5 m^3/s and from the neighbor basins Topolog, Vaslan, Doamnei, through 10 secondary catchments with underground galleries of a total length of 29 km – 12.2 m^3/s. The storage lake has a total volume of 465 million m^3, of which 320 million m^3 represents the efficient volume. The normal top-water level is 830 mASL, the lake having an area of 870 ha and a length of 14 km at this elevation.

As data was used, beside the ERA-40 and the GCM models a row of mean monthly values of flow into Vidraru reservoir during 1971-2000 and also from Bradet, Bahna Rusului and Voina weather stations was used monthly precipitation data for 2008, 2009, 2010 and 2011. All these weather stations are found in vicinity of Vidraru.

5. Results

In the monthly analyze for the correlation coefficients the results are shown in table 1 and for the ERA-40 model which has the highest correlation coefficient it is presented a correlation map, figure 1 and the chart comparing the original yearly flow average values with the ones obtained with the 'Polyfit' function, figure 2.

Using bootstrap procedure it was estimated the correlation between the inflow and precipitation for the models. The results were not more different than the originals correlation values. The point having the highest correlation coefficient modified from the initial position for only three models, but this position is very close to the first one.

Table 1

Results obtained in the monthly analyze for the correlation coefficients

Model name	Best point			Coefficient		'Polyfit' coefficients	
	Correlation	Latitude	Longitude	RMSE [-]	MAPE [%]	a	b
MPI-M-REMO	0.73	44.3919	27.2901	2.20	9.86	5.80E+05	10.032
SMHIRCA	0.62	45.1744	25.8659	2.59	11.63	4.49E+05	7.7757
METNOHIRHAM	0.73	44.3917	27.29	2.20	9.86	5.80E+05	10.032
DMI-HIRHAM5	0.60	45.5155	29.3895	2.71	11.40	4.79E+05	11.9486
C4IRCA3_CTL	0.61	44.126	20.2145	2.58	11.01	3.60E+05	8.9538
C4IRCA3	0.62	45.1744	25.866	2.59	11.63	4.49E+05	7.7757

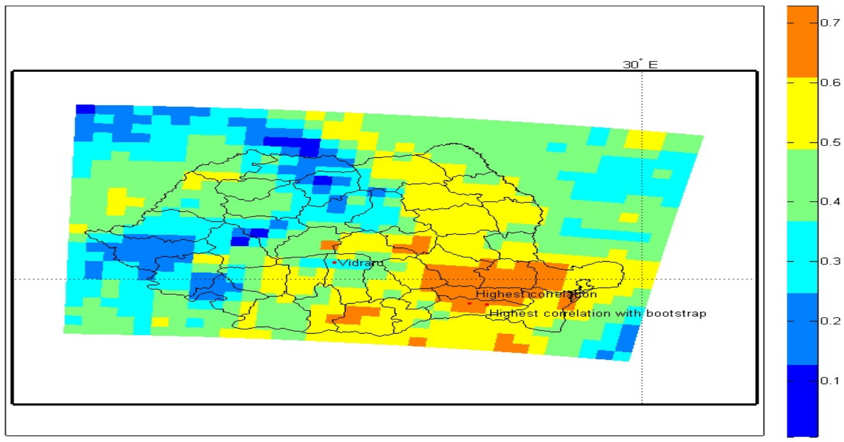

Fig. 1. MPI-M-REMO – correlation map

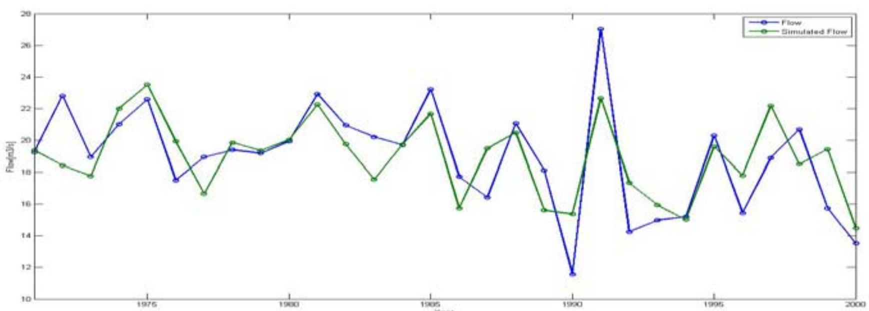

Fig. 2. MPI-M-REMO – flow time series

To obtain a better linear model a cross validation procedure was implemented and the results are shown in table 2, where it is shown also the results in the analyzed cases which were taken in consideration (a: nearest point; b: point with the highest correlation coeff. on precipitation; c: point with highest correlation coeff. on temperature; d: point with the highest correlation coeff. on temperature and the point with the highest correlation coeff. on precipitation; e: nearest five points; f: highest three correlation points; g: highest correlation on precipitation and highest correlation on temperature and nearest point).

Table 2

RMSE values for the analyzed cases

	Method/ Model	RMSE					Avg.
		SMHIRCA	MPI-M-REMO	METNOHIRHAM	DMI-HIRHAM5	C4IRCA3_CTL	
Precip + Temp	**a**	2.7	2.83	2.83	3	2.63	2.798
	b	2.61	**2.24**	**2.23**	**2.61**	**2.37**	2.412
	c	2.67	2.82	2.72	3.11	2.61	2.786

	d	**2.61**	**2.24**	2.24	2.62	**2.37**	2.416
	e	3.78	3.27	3.05	4.34	3.84	3.656
	f	3.02	2.66	2.43	2.89	2.71	2.742
	g	2.57	2.81	2.98	3.25	2.6	2.842

The time series flow obtained with cross-validation for the cases of the nearest point and the case which used data from the point with the highest correlation on temperature and on precipitation are represented here, in figures 3 and 4.

The cases which were chosen to be presented are the ones for which the data was taken into account from nearest point of Vidraru and the point of best correlation of the precipitation and temperature. As with previous analyzes best results were obtained in the case of the MPI-M-REMO. It shows a major improvement in results through the use of cross-validation procedure compared to the previously used linear model, which is why the implementation of the method was used in the first place.

In the next phase of the methodology, in the monthly analyze was used only data from the nearest point to Vidraru from both ERA-40 and GCM models. The main parameter that was given the highest importance was the R^2 coefficient, both for the average and the standard deviation, figure 5.

Some models gave as result some good R^2 values both for the mean and for the standard deviation. We considered the R^2 to be the most important value in this analyze. In the plot below we can observe these values better. Basically we have two areas: a good one, which includes also the weather stations and a bad area where the results are not conclusive. For the standard deviation the R^2 values was approx. 0.5 but in for the mean some models like DMI and ICT have an approx 0.1.

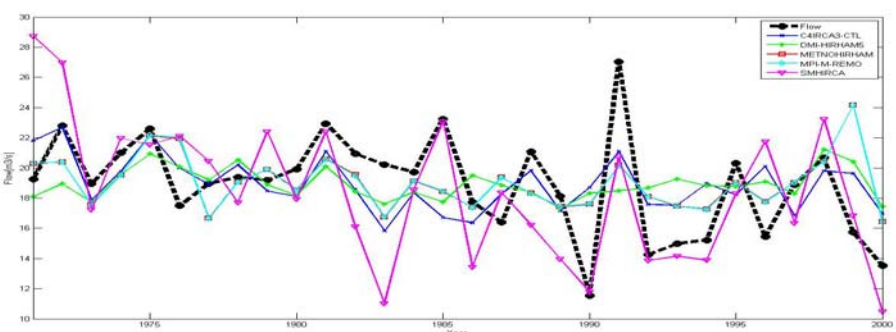

Fig. 3. Cross-validation – nearest point data

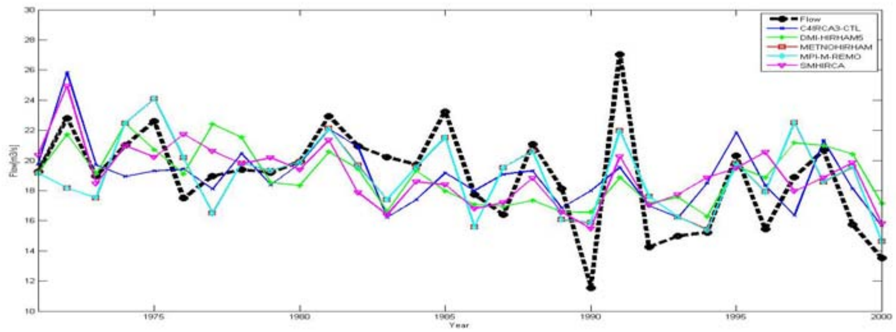

Fig. 4. Cross-validation – highest correlation point data

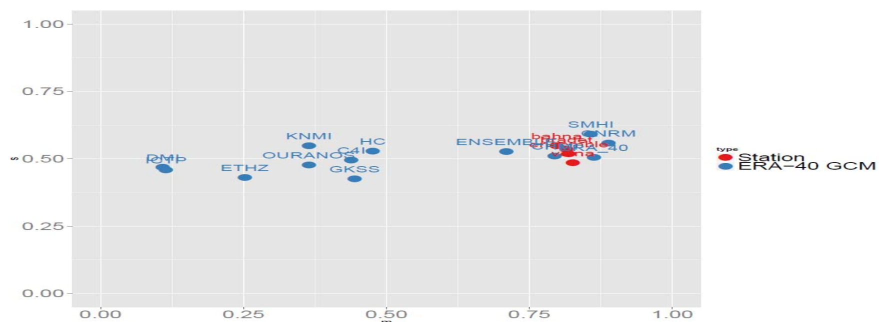

Fig. 5. R^2 values

To be able to compare these results with the real data, a flow time series was build and then created histograms for each model, figure 6.

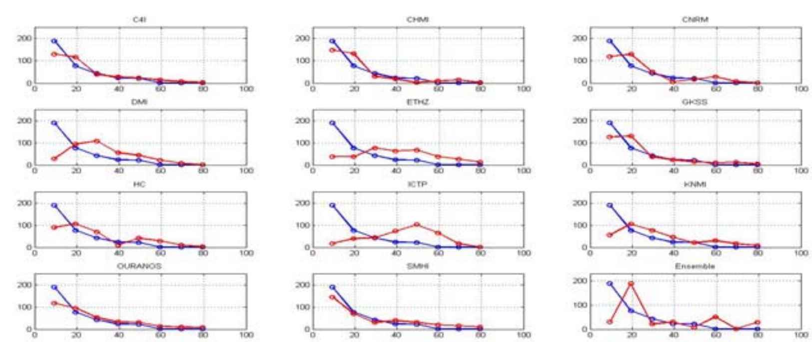

Fig. 6. Histogram compare

In this way we could assess this approach and could establish for the next step that the data from the weather stations is very reliable related to the other data that we have and we can use it in the further step.

6. Conclusions

As it was said the monthly assay analyzes the impact of climate changes on monthly averages flow values. In figure 7 it is presented a comparison between the current values of monthly mean values and the predicted ones from the GCM models. The blue color represents the mean flow value and with red is represented the standard deviation value.

Almost every GCM model gives as a result a reduction in flow values, especially in the summer months, June, July and August. For the July month is observed the biggest difference.

There are also models like SMHI-BCM and KNMI-MIROC where the difference is very small and where practically the mean monthly values for flow are predicted to be practically the same as in the present time. The most curious result is given by the DMI model where we can observe a big negative difference for July and August but for September and October the difference is positive and significant.

As a final remark we can say that it is expected a flow reduction in the period 2071-2099, with the biggest reduction in the summer months.

To be able to give a response in the energy analyze that was the original problem of this study; with the mean monthly values of flow it was possible to compute a potential energy production. This comes to complete the previous chart. In figure 8 is presented, for each model the relative percentage value related to the energy production with the original flow data from 1970-2000.

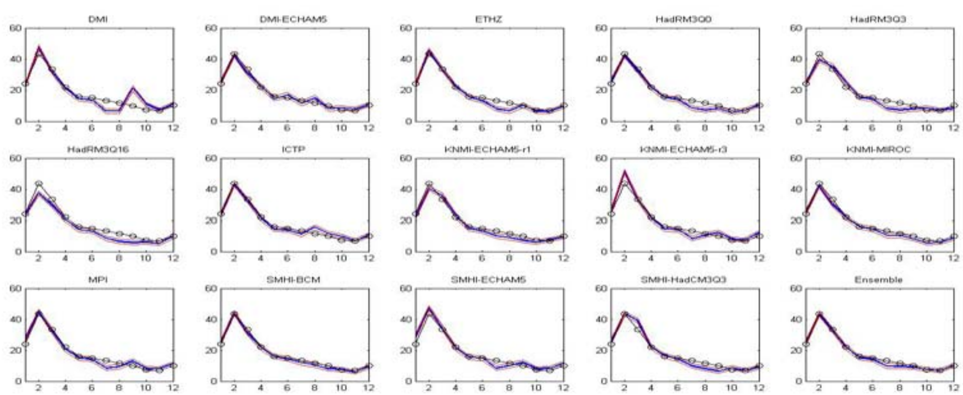

Fig. 7. Mean monthly flow values

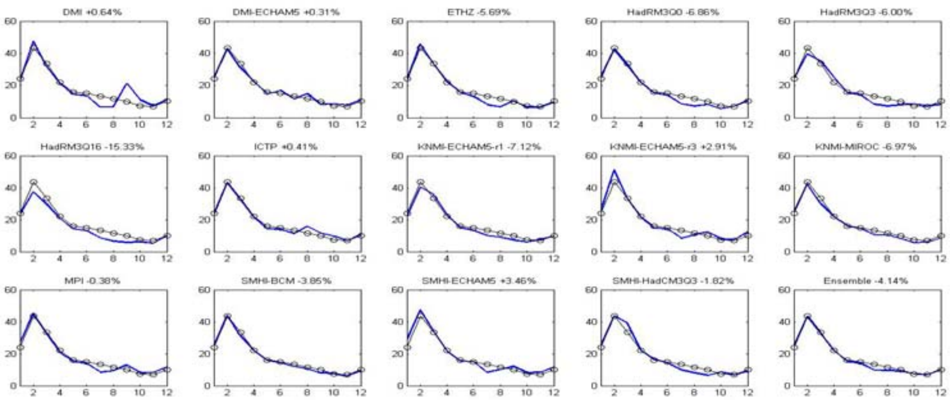

Fig. 8. Energy production

Also it can be observed that the majority of the models gave as a prediction that the potential energy production will decrease in the future. The model HadRM3Q16 gave as a result a reduction of 15.33% but also SMHI-ECHAM5 predicted a 3.46% increase in production. The Ensemble gave us the most reliable prediction, a reduction of 4.14%.

A final conclusion could be that it will be a reduction in potential energy production from the HPP Vidraru in the 2071-2099 period.

This analysis, despite of the problem in real data, is a trusty one and the result are one which must be taken in consideration. The used models, ERA-40 and GCM are made in specialized institutes and are used in the entire world for any type of analyze.

Acknowledgments

The work has been funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/134398.

R E F E R E N C E S

[1]. *R. Schaeffer, et al*. "Energy sector vulnerability to climate change: A review." Energy 38.1 (2012): 1-12.
[2]. *N. W*. Arnell "Climate change and global water resources." Global environmental change " (1999): S31-S49.
[3]. *E. Parson*, *et al*. "Global-change scenarios: their development and use. "US Department of Energy Publications (2007): 7.
[4]. *A. Frossard Pereira de Lucena*, *et al*. "The vulnerability of renewable energy to climate change in Brazil." Energy Policy 37.3 (2009): 879-889.
[5]. *I. H. Witten, and Eibe Frank.* Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2005.