Application of prognostic ensembles of global climate models to generate distributed estimates of water content in catchment basins

Valeriya Petruhina[1,*]

[1] Energy Systems Institute Siberian Branch of the Russian Academy of Sciences

Abstract. The problem of predicting climate change and its impact on humans is quite important and relevant in recent times. For a long time, mechanisms and methods for predicting the behavior of the climate in various regions and regions of our planet have been developed. Due to climate change, aggressive human impact on nature, and other various factors, the methods developed in the mid-twentieth century are becoming ineffective, and it is time-consuming but feasible to calculate using several methods. The article considers the technology of processing geoclimatic data, which is used to form spatially distributed predictive estimates of the state of the atmosphere.

Introduction

Assessing the water content in the basins of hydroelectric reservoirs is an important task. ESI SB RAS uses a whole range of means for processing prognostic ensembles. One such component is GeoGIPSAR [1], which uses the transformed data of the global CFSv2 system (Climate Forecast System Flux Products).

The second version of the system CFSv2[2] was developed at the NCEP Environmental Modeling Center. This is a fully connected model representing the interaction between the Earth's atmosphere, oceans, land and sea ice. This model has been operating in NCEP since March 2011 and provides prognostic data around the globe for general use. For long-term prognostic assessments, prognostic ensembles of up to 9 months are of greatest interest.

Table 1 shows the most used parameters for predicting geoclimatic data with atmospheric layers for isobaric surfaces (1000, 850, 500 hPa).

Table 1. System Settings CFSv2

Parameter designation	Parameter Name
chi200	Speed potential (200 hPa)
chi850	Speed potential (200 hPa)
cprat	Convective deposition rate [kg/m ^ 2/s]
csdlf	Downward long wave solar flux with clear sky [W/m ^ 2]
csdsf	Downward solar flow with clear sky [W/m ^ 2]
dlwsfc	Descending long wave rad. flow [W/m ^ 2]
dswsfc	A descending shortwave rad. flow [W/m ^ 2]
gflux	Soil heat flow [W/m ^ 2]
icecon	Ice cover [Proportion]
icethk	Ice thickness [m]
Lhtfl	Latent Heat Net Flow (Resultant Flow) [W/m ^ 2]
Nddsf	Near-IR-diffuse descending solar flux [W / m ^2]
Prate	Precipitation rate [kg/m ^ 2/s]
Pwat	Sedimentary water [kg/m ^ 2]. All atmosphere (considered as 1 layer)
Q2m	Specific humidity [kg/kg] at 2 m above ground
Shtfl	Sensitive Heat Net Flow [W/m ^ 2]
Snohf	Heat flow of phase transition of snow [W/m ^ 2]
Soilm1	Soil humidity [kg/m ^ 2] 0-2 s underground
Soilm2	Soil moisture [kg/m ^ 2]
Tmax	Maximum temperature [^{0}C] at 2 m above ground
Tmin	Minimum temperature [^{0}C] at 2 m above ground
Ulwsfc	Rising longwave rad. flow [W/m ^ 2] on the surface
Ulwtoa	Rising longwave rad. flow [W/m ^ 2] in upper atmosphere
Uswsfc	Rising shortwave rad. flow [W/m ^ 2] on the surface
Uswtoa	Rising shortwave rad. flow [W/m ^ 2] in upper atmosphere
Vddsf	Visible diffuse descending solar flow [W/m ^ 2]
Weasd	Water equivalent of accumulated snow depth [kg/m ^ 2]

Methods and Techniques

---

* Corresponding author: lera.petruhina@mail.ru

To create a forecast of environmental factors, an integrated approach is used, which includes Internet data monitoring. Internet monitoring performs the following tasks:

a) collection of actual global indicators;

b) collection of point indicators;

c) collecting data from prognostic ensembles formed by global climate models.

For the efficient processing of many prognostic ensembles in the GeoGIPSAR system, a specialized CFS format has been developed according to the aggregated values of various indicators per day. The most important parameters for assessing water content are precipitation intensity, air temperature, pressure, geopotential [3].

To generate predictive distributions of the selected parameter for a particular region, absolute and relative indicators are used in the form:

$$P(k, t) = \{\bar{p}_{ij}, i = 1, \dots, N_x, j = 1, \dots, N_y\} \quad (1)$$

$$\bar{p}_{ij}(k) = \frac{1}{T}\sum_{t=1}^{T} \bar{p}_{ij}(k, t) \text{ for absolute values} \quad (2)$$

$$\bar{P}_{ij}(k) = \frac{\sum_{k=1}^{K} c_k p_{ij}(k)}{\sum_{k=1}^{K} c_k} \qquad 0 \leq c_k \leq 1 \qquad \text{for}$$

relative values [4]. $\qquad (3)$

The collection of actual global indicators allows us to see the dynamics of changes in environmental factors on a global scale, which contributes to the creation of accurate statistics on indicators distributed across the global grid (in terms of height, speed, etc.). The collection of point indicators is aimed at selecting data in a circle of specific coordinates that you need to specify.

Based on the processing of prognostic ensembles of meteorological indicators in the considered region, estimates of tributaries to the cascade reservoirs are formed in the form of ranges of probabilistic distributions. Final prognostic scenarios are determined by an automated procedure including components of search for analogue years, cut-off of unlikely events, processing of regression relationships of meteorological indicators with river expenditure and refinement of their boundaries based on expert estimates generated by other models [5].

To generate geoclimatic maps during data analysis, a set of programs for Internet monitoring and analysis of data in the language LuaIsem[6] using Gnuplot graphics system has been developed [7]:

1. Reading files from the Internet *.nc

2. updating the local database Gi3 by new files from the database NC2;

3. on-line generation of maps with pressure and temperature anomalies [8].

Global Model Predictive Allocation Program Brief

The formation of prognostic maps is one of the most important stages aimed at assessing water content. This step is performed by automated programs involving parameter prediction. Figure 1 shows how geoclimatic mapping programs work. The program receives the following data at the input:

a) CFS files generated from NC files obtained by Internet monitoring of system data CFSv2;

b) prognostic indicator for which geoclimatic maps are to be constructed;

c) input dates for which it is necessary to obtain a forecast for certain indicators;

d) the number of ensembles, according to which the selection of current data takes place, on the basis of which a forecast is made in the future;

e) coordinates of the prediction area, with the help of which a visual representation of the output data is provided.

When generating output data, the program is guided by data interpolation parameters, forecasting methods and selection methods of prognostic ensembles. Components of prognostic maps generation work on generated Lua-scripts [9], as well as on GeoGIPSAR system.

At the output, the program provides up-to-date predictive maps of absolute and relative indicators of various geoclimatic parameters, graphs of dynamics of change of indicators, as well as plots of probabilistic distribution.
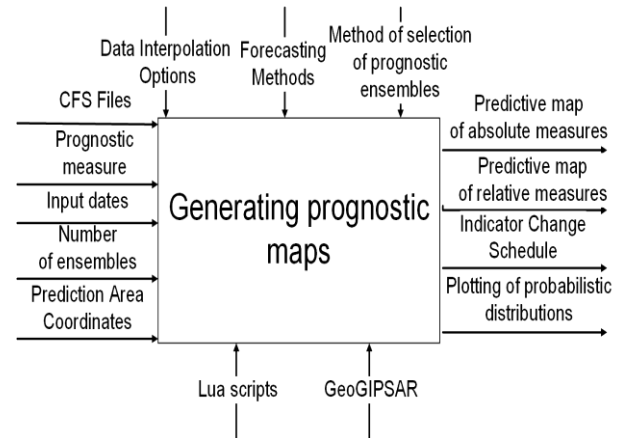

Fig. 1. Diagram of geoclimatic map generation program

Application Examples

Figure 2 shows how accurately the predictive map is formed. The predictive score is presented as the most successful result of the script for constructing predictive maps. For comparison, (a) an analytical map based on actual data and (b) a predictive map based on the global climate system are given - for January (01-31) 01.2021. For the accuracy of the study, the input dates were taken equally relative to the prediction month.

The date of the initial forecast ensemble was taken 01.12.2020. The 15.12.2020 date was selected to select the averages. The number of ensembles was chosen 90, and to calculate the average value - 30.

In the upper part (88,52; 113,61), the prognostic map differs from the actual on average on the 2-3^0C. In the lower part (88,46; 113,52), deviations in 2-3^0C are observed. But even with such deviations, the temperature tendency can be seen with sufficient accuracy.
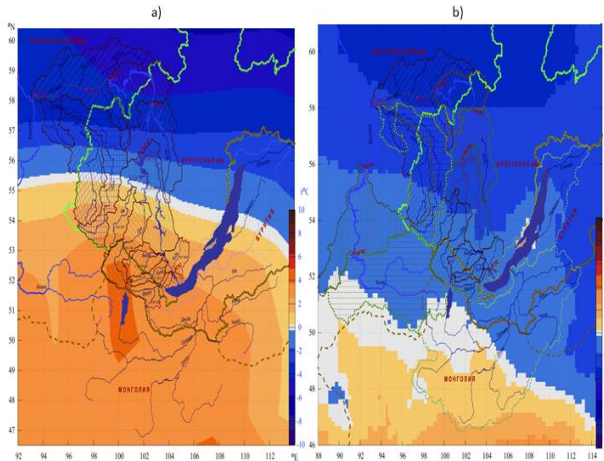
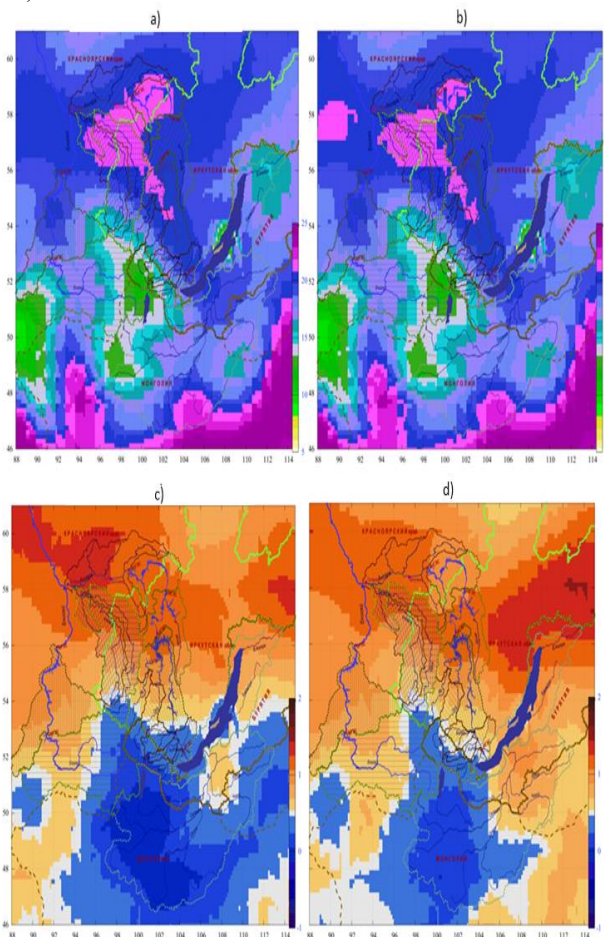**Fig. 2.** Comparison of actual and forecast data for period (01-31) 01.2021

Fig. 3. Geoclimatic temperature maps for July 2021: for absolute indicators for 15 ensembles from 08.03.2021 (a), for 10 ensembles from 11.03.2021 (b); for relative indicators for 15 ensembles from 21.03.2021 (c), for 10 ensembles from 26.03.2021 (d).

At the moment, one of the most important periods for forecasting is July 2021, the results of the forecast of which could be seen already in March 2021. The figures 3,4 show forecasts of temperature and precipitation indicators for the period (01-30) 07.2021. You can see that when the number of ensembles changes, the forecast changes. The ensembles have the greatest influence on the relative indicator in the forecast. At the same time, the analytical indicator remains almost unchanged.

According to the forecast in July in the areas of tributaries of Lake Baikal, the temperature indicator indicates $+ 20^0C$, which is the norm for this period, the intensity of precipitation ranges from 13 to 17 mm, exceeding the norm by 3-5 units.
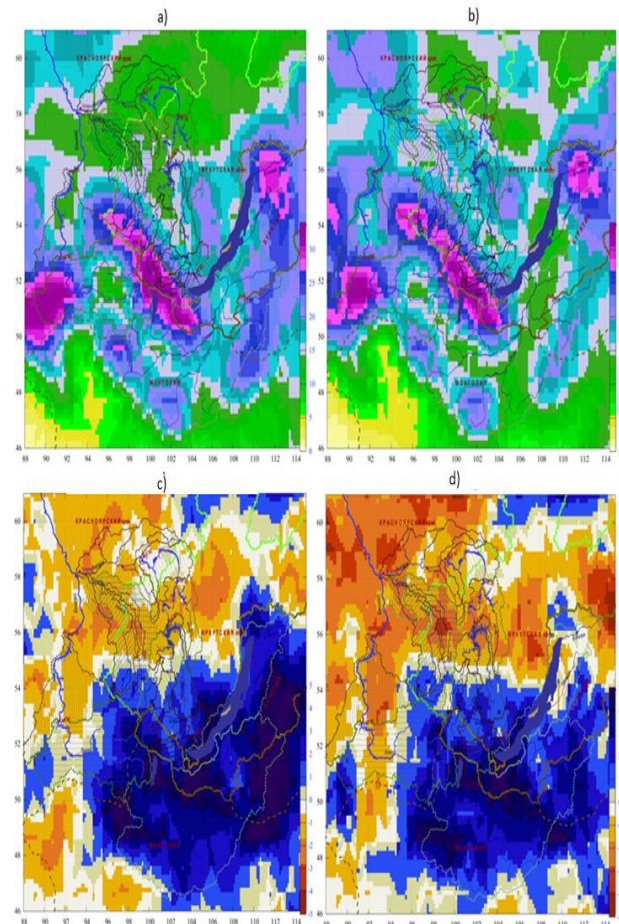

Fig. 4. Geoclimatic map of the precipitation regime for July 2021: an absolute indicator of precipitation intensity for 15 ensembles from 08.03.2021 (a), 10 ensembles from 05.03.2021 (b); for the relative values of 15 ensembles from 21.03.2021 (c), 10 ensembles from 26.03.2021 (d).

Conclusion

Specialized components in ESI SB RAS are most effectively used to predict geoclimatic data, assess water content in basins of hydroelectric reservoirs, and determine predictive scenarios of water inflows to hydroelectric reservoirs.

It is planned to introduce new important parameters into the system for the most accurate forecasting. Increasing the number of parameters in the system cannot always give a positive result. You need to select data that correlates with the predicted parameter. Thus, ice cover, sedimentary water, and soil moisture may be the most suitable parameters for assessing water content. These parameters are recommended to be included in the prediction system to obtain improved prognostic data results.

References

1. Abasov N.V., Berezhnykh T.V., Nikitin V.M., Osipchuk E.N. Methodology for the formation of long-term predictive scenarios of water content and temperature regimes//Collection of reports of the international conf. "Energy of the 21st century: Sustainable development and intellectual management" (September 7-11, 2020), Irkutsk: ESI SB RAS, 2020. Page 100-101.

2. Suranjana Saha, Shrinivas Moorthi, Hua-Lu Pan, Xingren Wu, Jiande Wang, Sudhir Nadiga, Patrick Tripp, Robert Kistler, John Woollen, David Behringer, Haixia Liu, Diane Stokes, Robert Grumbine, George Gayno, Jun Wang, Yu-Tai Hou, SS.SL., 2020. The NCEP Climate Forecast System Reanalysis. Amer. Meteor. Soc., pp: 44.

3. Abasov N.V., Vetrova V.V., Technology of geoclimatic data processing//Proceedings of the XII Baikal All-Russian Conf. "Information and mathematical technologies in science and management." Part II. - Irkutsk: ESI SB RAS, 2007. Page 85-91.

4. Abasov N.V., Berezhnykh T.V., Vetrova V.V. Long-term forecasting of hydropower potential of the HPP cascade in climate change//Izvestia RAS, Energy, 2012, No. 1, S. 49-57.

5. Abasov N.V., Berezhnykh T.V., Reznikov A.P. Long-term forecasting of environmental factors in the energy sector/Systemic research of energy problems. - Novosibirsk: Science. Sib. ed. company RAS, 2000, -S.415-429.

6. Roberto Ierusalimschy, RI., 2016. Programming in Lua forth edition. Lua.org, pp: 310.

7. Schneivais, A.B., 2006. Bases GNUPLOT'a: Tutorial. St. Petersburg State University, pp: 59.

8. Abasov N.V., On the development of the predictive system of GIPSAR based on Lua-components//Tp. XVI Baikal Vseros. conf. "Information and mathematical technologies in science and management." T.3. - Irkutsk: ESI SB RAS, 2011. Page 212-217.

9. Roberto Ierusalimschy, RI., 2013. Programming in Lua Third Edition. Rio de Janeiro: Lua.org, pp: 366.