International Letters of Natural Sciences
ISSN: 2300-9675, Vol. 82, pp 13-19
doi:10.18052/www.scipress.com/ILNS.82.13
CC BY 4.0. Published by SciPress Ltd, Switzerland, 2021

Submitted: 2020-08-04
Revised: 2021-01-14
Accepted: 2021-01-27
Online: 2021-04-28

Seasonal and Annual Probabilistic Forecasting of Water Levels in Large Lakes (Case Study of the Ladoga lake)

Natalia V. Myakisheva[1], Ekaterina V. Gaidukova[1,a*], Sergei V. Shanochkin[1], Anna A. Batmazova[1]

[1]Department of Engineering Hydrology, Russian State Hydrometeorological University (RSHU), Saint-Petersburg, Russia

[a]oderiut@mail.ru

Keywords: water resource, water level, large lake, seasonal fluctuation, statistical forecasting, ARIMA-model.

Abstract. The production functions of water-dependent sectors of the economy can include the water level in the lake as a natural resource. This characteristic must be able to reliably predict for the effective functioning of sectors of the economy. In the article the main attention is paid to the methods of forecasting based on the extrapolation of natural variations of the large lakes water level. As an example, Lake Ladoga is considered. In this paper, it is assumed that the level varies accordingly to a stochastic multi-cycle process with principal energy-containing zones in frequency bands associated with seasonal and multi-annual variations. Hence, the multi-year monthly and yearly averaged time series are represented by the ARIMA (auto-regression integrated moving average) processes. Forecasts are generated by using of the seasonal ARIMA-models, which take into account not only the seasonal but also the evolution non-stationarity. To compare the forecasts and the actual values, the relative errors are computed. It is shown that implementation of the models mainly allows receiving good and excellent forecasts.

Subject Classification Numbers: UDC 556.555.2.06(4)

Introduction

Many sectors of the economy, whose activities are closely related to the use of water resources, require hydrological forecasts. A reliable forecast allows you to optimize economic activities taking into account the needs of production (see, eg [1, 2]).

A human has many consumer interests in water, and this interest may be generalized in the language of the economy. If water is regarded as a natural resource, then hydrological variables in economic models may act either as parameters (then hydrology is an external reality for the economy) or on the contrary they can be influenced by economic variables (then we are dealing with an expanded subject area combining hydrology and economy).

Fig. 1 shows production and technological interpretation of the economy. The production characterizes the following aspects: labor (L); instrument of labor or fixed capital stock (K); subject of labor ($\widetilde{W} = W^S + W$), comprising natural resources WS and subject of labor returned to the production as a part of gross product X.

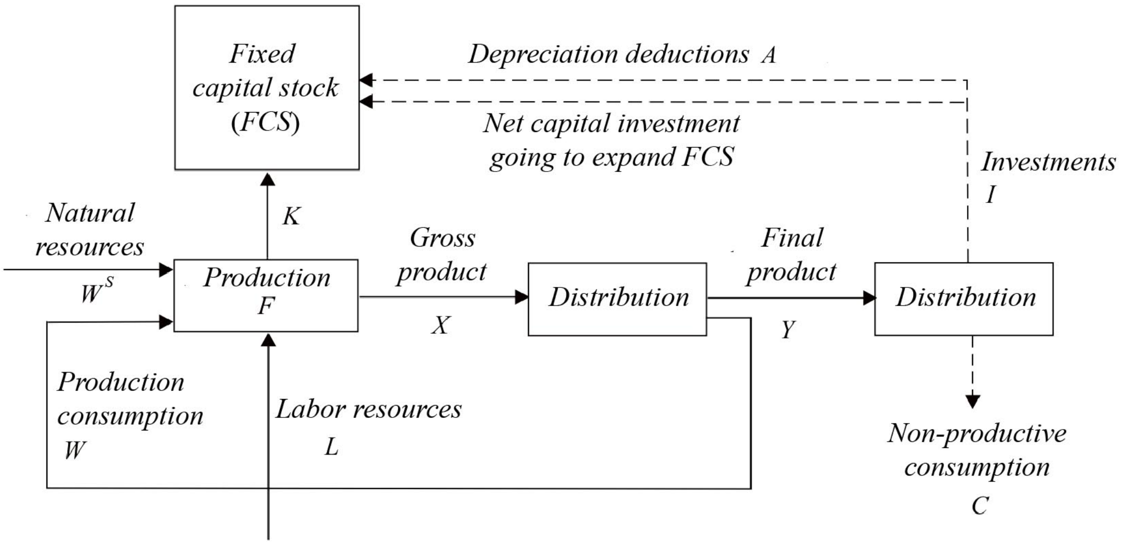


Figure 1. Interrelation of the production elements.

In this situation the condition of economy is determined by the capital-labor ratio *k* (*K*/*L*), and its management – by labor performance *x* (*X*/*L*) and a share of non-production consumption *u* (*C*/*Y*). With the use of theory of sufficient conditions of optimality, we may show that optimum labor performance $\bar{x}$ is ensured at $\bar{x} = f\,(k,\ t)$. The production function $f\,(K,\ L,\ t)$ reflects a relation between production factors and its result. Parameter *t* may be used to account for the influence of external factors, including scientific-technical progress and variability of the natural resources, on the economy (its model).

Further we show how hydrological consequences of the climate change in economic calculations may be taken into account (see, eg [3]).

The purpose of the study is to consider methods and mathematical models of probabilistic forecasting, taking into account possible climatic changes, the water level in large lakes in Europe, using the example of Lake Ladoga, as a natural factor included in production functions.

Level regime of lakes is formed under impact of active and adaptive factors, which, in their turn, are influenced by modern climate and anthropogenic pressure. Climatic factors are usually considered as active, while factors of the underlying surface are adaptive [4]. Various combinations of the climate signal and the underlying surface response, in particular, define certain differences in patterns of seasonal [5, 6] and multi-years fluctuations of the lakes water level.

In this paper, the Ladoga Lake is considered as an example of a large European lake [7]. The seasonal cycle of its water level is quite smooth, while its multi-years fluctuations are basically represented by the overlapped low-frequency oscillation and slightly adjusted characteristic elements of the seasonal cycle. Previously performed researches [8, 9] showed that contribution of the seasonal fluctuations dispersion into the total dispersion is significantly less than that of the low-frequency oscillation.

Materials and Methods

Application of the autoregressive-integrated-moving-average (ARIMA) methods [10] for the observation time series analysis opens wide perspectives for statistical forecasting of the large lakes water stage.

In our research, the Ladoga Lake was considered. Monthly averages of the water level observed at the gauge in Syas'skie Ryadki in the period from 1881 to 2004 were used to simulate their seasonal fluctuations in the typical (in terms of water content) years (1923–1925, 1986, 1990, 1993, 1995, 1998, 1999 и 2003). Quantile analysis of the yearly averages has shown that, in 1923–1925, 1986, 1990, 1993 and 1995, the Ladoga Lake stages were above the multi-years median. Meanwhile in 1924, a historical maximum (609 cm) was recorded. In 1998, 1999 and 2003, average water stages were below the median; in 2003, the lowest stage for the considered period was recorded (381 cm), which is the ranked #2 since 1940 (361 cm).

Statistical generation of forecasts was performed in two ways. The first way was based on the seasonal ARIMA-model (p, d, q) (P_s, d_s, Q_s) (where p denotes the autoregression parameter, d denotes an order of the residual sampling, q denotes the moving average's parameter, P_s denotes the seasonal parameter of autoregression, d_s denotes the seasonal residual, Q_s denotes the seasonal parameter of the moving average). This model takes into account both evolutional and seasonal non-stationarity. The water stages were forecast with 12-monthly lead time (January through December). The model included two parameters of the moving average (i.e. regular and seasonal parameters). The initial time series was transformed three times: (1) when it was logarithmed, (2) when residuals were taken with lag equaling to 1 (to remove the evolution nonstationarity), and (3) when residuals were taken with lag equaling to 12 (to remove the seasonal nonstationarity).

The second way was based on presenting of the monthly averages through a vector, whose components were defined by the annual consequences of values estimated for every month of the year T:

$$H_t = \{H_1, t, \ldots, H_{12}, t\},\ T = \{H_i, t\},\ i = 1, \ldots, 12, \tag{1}$$

where t denotes corresponding time series of monthly values, which are considered as a stationary and stationary correlated stochastic process.

Expression of the modelled structure of every vector's component via the first order autoregressive process (AR(1)) or the first order moving average process (MA(1)) allows statistical prediction of the lake stage with a 1-year lead time. The forecast monthly values taken within the 1-yearly period for many years allow obtaining the seasonal stage fluctuation forecast.

To compare the forecast and actual values, relative errors were computed:

$$\delta = \frac{H_a - H_f}{H_a} \cdot 100\,\%. \tag{2}$$

To assess the forecasts' success rate, they were compared with actual values by using an additional criterion [11]:

$$\delta_{add} = \pm\, 0.674\sigma, \tag{3}$$

where

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}\left(H_{i,a} - \overline{H}_a\right)^2}{(n-1)}}, \tag{4}$$

$H_{i,a}$ denotes the actual water stage value, $\overline{H}_a$ denotes the multi-year mean of the water stage, n denotes the length of the time series). Generated forecasts were assessed as 'excellent' if $\delta < 0.3\delta_{add}$, 'good' if $\delta = (0.3 \div 0.6)\delta_{add}$, and 'acceptable' if $\delta = (0.6 \div 1.0)\delta_{add}$.

In addition, efficiency of the forecasting methodology was assessed by using the $\frac{S}{\sigma_\Delta}$ ratio (where S denotes the root mean square error (RMSE) of forecasts issued for the validation period,

$$S = \sqrt{\frac{\sum_{i=1}^{n}\left(H_{i,f} - H_{i,a}\right)^2}{n-1}}, \tag{5}$$

where $H_{i,f}$ denotes the forecast stage values, $H_{i,a}$ denotes the actual stage values, n denotes a number of forecasts issued within the validation period; σ_Δ denotes RMSE of the inertial forecasts for the same period,

$$\sigma_\Delta = \sqrt{\frac{\sum\limits_{i=1}^{i=n-\tau}\left(\Delta_i - \overline{\Delta}\right)^2}{n-\tau-1}}, \tag{6}$$

$\Delta_i = H_{i+\tau} - H_i$, $\overline{\Delta} = \sum\limits_i \Delta_i \big/ (n-\tau-1)$, where τ denotes the forecast lead time). An assessed methodology is effective, if $\frac{S}{\sigma_\Delta} \le 0.70$ at $n < 15$, $\frac{S}{\sigma_\Delta} \le 0.75$ at $15 \le n < 25$, or $\frac{S}{\sigma_\Delta} \le 0.80$ at $n \ge 25$.

Results

Results of forecasting of the Ladoga Lake seasonal water stage fluctuations in the period from 1923 to 1925 performed by using the seasonal ARIMA-model are presented in Fig. 2 and Table 1. Intercomparison of the actual and modelled values computed for 1923–1925 showed their quite good coherence, which was especially nice in the period from June to September, when the difference between the actual and modelled values did not exceed 1–3 % (e.g. in 1923 and 1925). In the periods from April to May and from November to December, the difference increased up to 12 % – this is due to the larger natural fluctuations during these periods (spring and autumn floods). In accordance with the mentioned criterion δ, statistical forecasts issued for 1923 and 1925 can be assessed as ‘good’, while hindcasts of stages in 1924 can be assessed as ‘acceptable’. Beside that, the forecasting methodology is effective ($\frac{S}{\sigma_\Delta} = 0.8 \div 0.92$).

In 1986 and 1993, the AR(1) model performance was found better, while in 1990 and 1995 years the MA(1) succeeded. In these years, hindcasts were ‘excellent’.

In 2003, statistical forecasts were obtained by using two approaches from three models (Fig. 3 and Table 2). Thus, forecasts issued by using ARIMA and MA(1) are ‘good’, while hindcasts issued by using AR(1) are ‘excellent’.

Table 1. Intercomparison of actual (H_a) and modelled (H_m) values of the Ladoga Lake water stage.

Month, year	**H_a, cm**	**H_m, cm**	**δ**	**Month, year**	**H_a, cm**	**H_m, cm**	**δ**
January, 1923	483	486	–1	July, 1924	659	565	14
February, 1923	482	483	0	August, 1924	640	556	13
March, 1923	480	492	–3	September, 1924	618	541	12
April, 1923	472	511	–8	October, 1924	599	531	11
May, 1923	488	543	–11	November, 1924	576	513	11
June, 1923	518	551	–6	December, 1924	556	505	9
July, 1923	528	546	–3	January, 1925	547	486	11
August, 1923	523	534	–2	February, 1925	532	484	9
September, 1923	518	520	0	March, 1925	539	492	9
October, 1923	520	512	2	April, 1925	547	511	7
November, 1923	536	495	8	May, 1925	462	543	3
December, 1923	558	489	12	June, 1925	553	551	0
January, 1924	566	487	14	July, 1925	541	546	–1
February, 1924	589	472	20	August, 1925	521	534	–2
March, 1924	586	492	16	September, 1925	503	520	–3
April, 1924	608	522	14	October, 1925	497	513	–3
May, 1924	646	557	14	November, 1925	494	495	0
June, 1924	664	568	14	December, 1925	486	489	–1

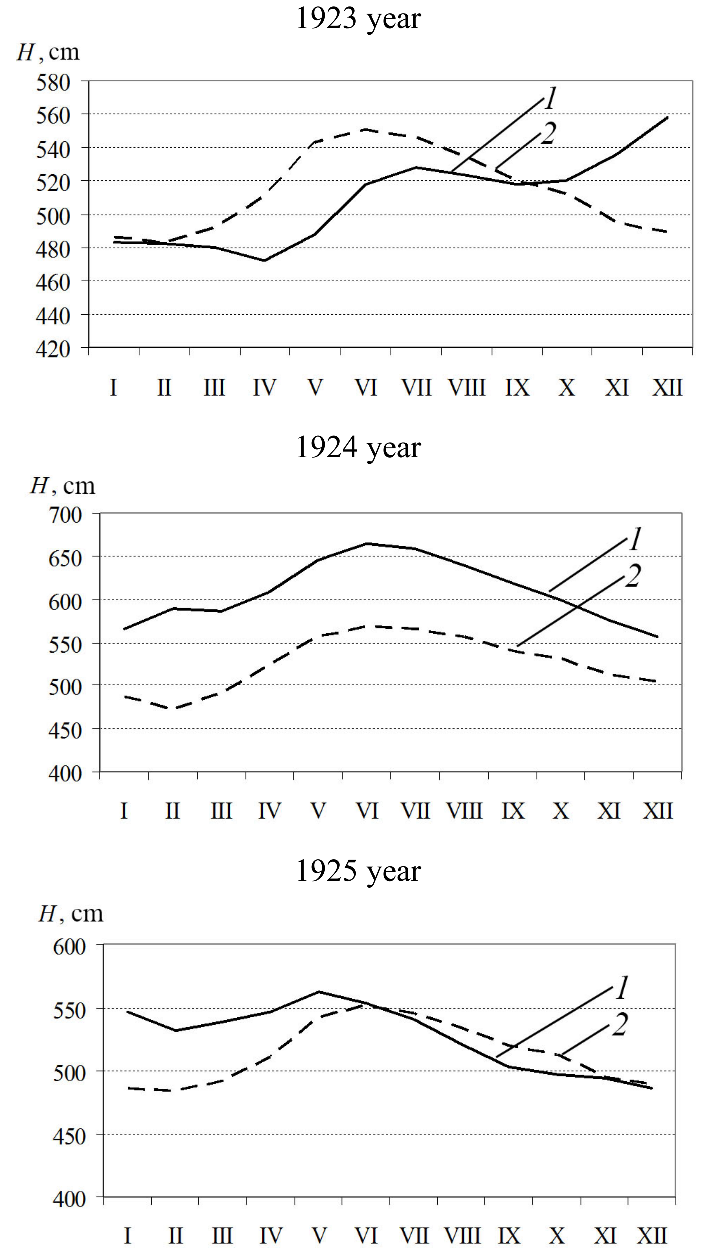
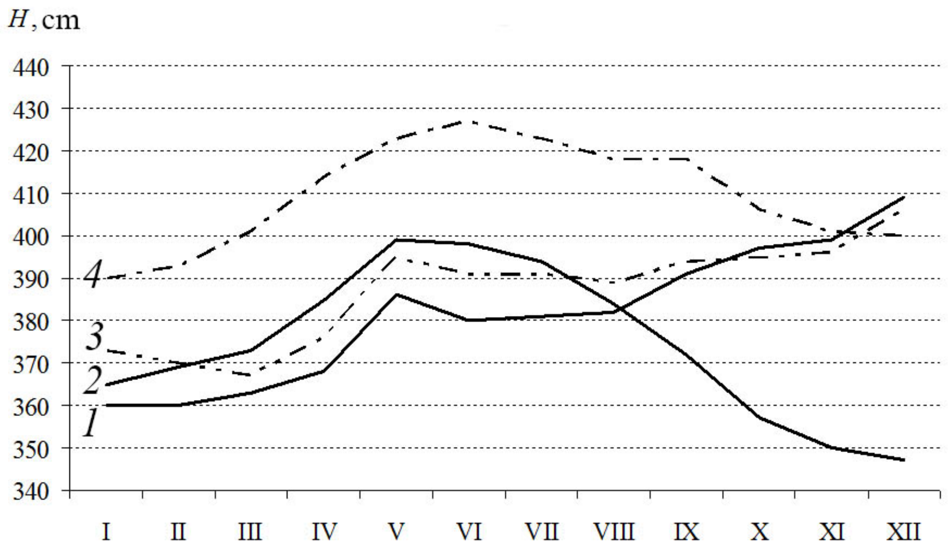


Figure 2. Actual (Series 1) and modelled (Series 2) values of the seasonal fluctuations of the Ladoga Lake water stage in 1923–1925 years.

Table 2. Intercomparison of actual (H_a) and modelled (H_m) values of the water stage in 2003.

Month	**H_a, cm**	**H_m, cm**		
		1st way	**2nd way**	
		ARIMA	**AR(1)**	**MA(1)**
January	360	365	373	390
February	360	369	370	393
March	363	373	367	401
April	368	385	376	414
May	386	399	395	403
June	380	398	391	427
July	381	394	391	423
August	382	384	398	418
September	391	372	394	418
October	397	357	395	406
November	399	350	396	401
December	409	347	406	400



(1) observation, (2) ARIMA, (3) AR(1), (4) MA(1)

Figure 3. Actual and modelled values of the Ladoga Lake water stage in 2003.

Note that monthly correction of statistical forecast by using 12 regression equations is possible (Table 3). In this case, two factors (the actual average water level for the preceding month and the average level of the same month for the previous year) are used as predictors.

Table 3. Forecast correction in 1998 and 1999

Year	***H*, cm**	**Months**												**Average**
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
1998	H_a	417	422	428	447	466	471	477	483	485	476	780	485	461
	H_m	416	421	427	444	468	476	469	471	474	478	468	479	458
1999	H_a	491	493	492	510	528	523	509	485	463	436	419	411	480
	H_m	490	495	497	510	534	545	541	531	516	507	497	494	513

The obtained results once again confirm that the ARIMA-methods can be used as universal of prediction of the water stage fluctuations in the large lakes.

Conclusion

The use of lakes for economic purposes is very diverse. Lakes serve as sources of water supply and irrigation of fields. Large lakes, lake-river systems are mostly waterways. If dams are built on rivers flowing out of lakes, these objects become reservoirs for hydroelectric power plants. Lakes are also widely used for fishing and fish farming, as recreational facilities, cooling ponds for state district power stations and nuclear power plants, for sapropel mining, for balneology purposes, etc.

When planning economic activities, it is necessary to take into account future changes in water levels in the lake. For this, mathematical models are used.

On the example of the large Ladoga Lake, which is characterized by a slow water exchange, our research has shown that the seasonal ARIMA models can be successfully applied for the large lakes level statistical forecasting, which are described by seasonal fluctuations overlapped on the multi-years trend, while the AR(1) and MA(1) models can be used for modeling of multi-years consequences of the water level monthly values and, possibly, taking into account other components of the water balance [12, 13].

Conflict of Interest

The authors declare that there is no conflict of interest.

Acknowledgments

This study was supported by the grant issued by the Ministry of Science and Higher Education of the Russian Federation under the state contract No. FSZU-2020-0009 ‘Research of physical, chemical and biological processes in the atmosphere and hydrosphere under conditions of climate change and anthropogenic impacts’.

References

[1] V. Burlov, A. Grobitski, A. Grobitskaya, Construction management in terms of indicator of the successfully fulfilled production task, Magazine of Civil Engineering. 63(3) (2016) 77−91, doi:10.5862/MCE.63.5.

[2] V. Burlov, A. Andreev, F. Gomazov, Development of a model for the management of environmental safety of the region, taking into account of the GIS capacity (2018), doi:10.1051/matecconf/201819302038.

[3] E. Shevnina and et al., Assessment of extreme flood events in a changing climate for a long-term planning of socio-economic infrastructure in the Russian Arctic, Hydrology and Earth System Sciences, 21(5) (2017) 2559–2578, doi:10.5194/hess-21-2559-2017.

[4] A. Aksenov, A. Rybalko, M. Naumenko, Geomorphology of Lake Ladoga basin, Limnology and Freshwater Biology. 4 (2020) 492–494, doi:10.31951/2658-3518-2020-A-4-492.

[5] N. Filatov, L. Rukhovets, Ladoga Lake and Onego Lake (Lakes Ladozhskoye and Onezhskoye), Encyclopedia of Earth Sciences Series. 2012. 429–432.

[6] S. Karetnikov, M. Leppäranta, A. Montonen, A time series of over 100 years of ice seasons on Lake Ladoga, Journal of Great Lakes Research. 43(6) (2017) 979–988, doi: 10.1016/j.jglr.2017.08.010.

[7] Ladoga. Institute of Lakes Science RAS (2013), https://www.rgo.ru/sites/default/files/media/2013/08/ladoga.pdf.

[8] A.M. Doganovsky, N.V. Myakisheva, The water level of Lake Ladoga in different frequency ranges, University of Joensuu, Publications of Karelian Institute. 129 (2000) 374–377. (In Russian).

[9] A.M. Doganovsky, N.V. Myakisheva, Rejim napolneniya ozer raznykh klassov vodoobmena, VI Vsesojuzniy gidrologicheskiy sezd. Sektsiya 5 (2006) 137–140. (In Russian).

[10] Box and Jenkins, Time Series Analysis, Forecasting and Control (2013), doi:10.1057/9781137291264_6.

[11] Guide to Hydrological Practice. Water resources management and practical application of hydrological methods, WMO. Volume II (2012), http://www.whycos.org/hwrp/guide/russian/168_Vol_II_ru.pdf.

[12] V.V. Kovalenko, E.V. Gaidukova, The phenomenon of nonzero norm of long-term changes in the total water supply in river basins, American journal of environmental sciences. 11(2) (2015) 76–80, doi:10.3844/ajessp.2015.76.80.

[13] V.V. Kovalenko, E.V. Gaidukova, D.V. Chistyakov, A. Hamlili, Prognostic models of development of processes of catastrophic long-term annual river runoff formation, Russian Meteorology and Hydrology, 10 (2010) 695–699, doi:10.3103/S1068373910100079.