DIAGNOSIS OF LONG–TERM OSCILLATIONS AND POSSIBILITY OF FORECASTING TRENDS OF THE CASPIAN SEA LEVEL FROM THERMAL AND PRECIPITATION FORMING PROPERTIES OF SYNOPTIC PROCESSES

O.V. Zhiltsova and A.A. Isaev

M.V. Lomonosov State University, Moscow

Received November 15, 1993

Feasibility is considered of forecasting the level of the Caspian Sea from thermal and precipitation forming properties of synoptic processes. To construct a statistical model a combined technique is used of increments and of group account of arguments. According to computational results further increase of the level of the Caspian Sea is to be expected to –26.93 m *in 1993,* –26.76 m *in 1994, and* –26.64 m *in 1995.*

Noticeable oscillations of the level of the Caspian Sea have long been drawing the attention of researchers. That problem has become vitally important now, when further rise of the level of the Caspian Sea threatens its coastal cities with catastrophic consequences. Three basic hypotheses try to explain the causes of those changes[9]: the stochastic, which treats such oscillations as a result of random variations of the components of moisture budget; the geological, which explains variations in that level by the tectonic activity in the Caspian area; the climatic, according to which oscillations of the level of the Caspian Sea reflect large–scale hydrometeorological processes, which take place no only in the basin of the Caspian Sea itself, but far outside it as well. The aim of the present publication is to discuss different aspect of the latter hypothesis.

We investigated the relation between the level of the Caspian Sea and the atmospheric processes. The classification used was developed by Klimenko[7*)] for the first natural synoptic region. Basied on the results from Refs. 5 and 6 all the synoptic processes were divided into dry and rainy, warm and cold in depending on their thermal and precipitation forming properties.

Analysis of the balance between the rainy and the dry synoptic processes indicates that quite a clear conjunction is traces between the trends in those characteristics and oscillations of the sea level.

Starting from the turn of the century and up to the 1970's the dry atmospheric processes had generally and stable prevailed over the European Russia, while the rainy ones were in apparent deficit (see Table I). During the early 1970's the tendency sharply reversed. We may trace a similar behavior in the level of the Caspian Sea: from the beginning of the current century and up to 1977 practically continuous drop of the sea level took place , which reversed to a sharp increase starting from 1978. Thus, the increase/decrease of the level of the Caspian Sea is controlled by the particular features of the synoptic processes over the European Russia.

*) According to Klimenko's classification all the synoptic processes are divided into the processes of the cold and the warm seasons.

TABLE I Number of years with predominant rainy synoptic processes ratioed to the number of years deficient in such processes and cold seasons separately), according to phases of oscillations of the level of the Caspian Sea*

Oscillations periods.years	Cold season	Warm season	Annually	Behavior of the Caspian Sea level
1901 – 1909	0.33	0.80	0.80	Stabilization
1909 –1932	0.53	1.56	1.09	–"–
1933 – 1940	1.00	0.50	0.30	Drop
1941 –1948	0.33	0.60	0.60	Stabilization
1948 – 1956	0.50	0.60	0.33	Drop
1956 –1973	0.56	0.64	0.75	Stabilization
1974 –1977	0.33	4.00	0.67	Drop
1977 – 1991	4.00	1.50	6.50	Growth

*) In Klimenko's classification all the synoptic processes are divided into the processes of the warm and the cold periods (seasons).

Since one can hardly expect to arrive at a full understanding of mechanisms, through which various causes affect the level of the Caspian Sea, it is reasonable to transform the initial data into other easier predictable sequences, and to construct various statistical or physico–statistical models, which would predict the future level of the Caspian Sea without explicating fully the mechanism of interaction. The basis for developing such forecasts lies in the statement that the temporal sequences of hydrometeorological processes contain all the basic information on the causes of their development.

To construct such a model we used the prognostic properties of the sets of increments to be described below, which were combined with the method of clustering of arguments (MCA). The essence of this method, which may be found in more detail in Ref. 1, is as follows. Let the initial data present the stationary stochastic sequences, while their function of cohesiveness is given by the Kolmogorov–Gabor polynomial. According to the technique the solution for coefficients of the polynomial should then be sought from a system of normal Gaussian equations.

0235-6880/94/02 136-06 $02.00

Methodologically such a solution is reduced to recurrently solving several systems of normal equations, formed for each pair of arguments and for the new auxiliary values. It results in a sharp increase of the solution accuracy with simultaneous reduction of the amount of calculations. The increase in accuracy is due, on the one hand, to an hierarchical structure of the MCA algorithm and, on the other hand, to a possibility of selecting "useful" information, which is separated from the "deleterious" data by thresholds, set at each step of solving the elementary system of normal equations. If it would be possible to solve the complete system of normal equations for all the arguments simultaneously, the accuracy of the presentation would appear to be insufficient because of the limited number of the samples and nonstationary character of deviations.

Apart from the MCA in this paper we accounted for the methodology of extra–long–term forecasting to predict future levels of the Caspian Sea. The experience shows that forecasts may be more effective in case when one uses temporal sets of predictors already transformed with respect to initial observational data into increments of different orders:[3,4]

$$\Delta_i^1 = x_i - x_{i-1}\ ,\ \Delta_i^2 = \Delta_i^1 - \Delta_{i-1}^1\ ,\ \ldots\ ,\ \Delta_n^k = \Delta_n^{k-1} - \Delta_{n-1}^{k-1}\ ,$$

where x_1, x_2, ... x_i are the values from some temporal sequence; n is the length of the initial series; Δ_n^k is the increment of the kth order.

Sets of increments of different orders display various helpful features.[3] For example, one of the most important properties of individual increments is that each increment of the kth order may be treated as the acceleration of processes of the order $k - 2$. As compared to the initial values x_i, these latter ones are much more sensitive to accelerations of various external forcings,[3] the effect of such forcings to be first felt in the increments. Thus one may state that the increments of various orders are easier predictable than the values of hydrometeorological elements themselves, so they should be used for the purposes of long–term forecasting.

Knowing the forecasted increment of some order for the year $n + 1$, one may after rather simple mathematical transformations forecast the values of temporal sequence for that year.

We used the number of rainy, dry, warm, and cold synoptic processes for the warm and the cold seasons and their increments of various orders as the initial predictors. Initially we had 8×30×11 = 2 640 predictors, where 8 stands for 4 types of synoptic processes (warm, cold, rainy and dry) taken for the warm and the cold seasons, 30 is the temporal lag up to 30 years (which accounts for 30 years prehistory), 11 is the number of increments in every predictor (down to the 11th order). For the predicant (Y) we chose the first increments of the level of the Caspian Sea. This choice is explained, first, by the smoothing of trends and their effects, and, second, by lower errors, which they yield when retrieving the level of the Caspian Sea in contract to increments of higher orders.

The MCA procedure envisages several stages in selecting the predictors. After the first stage we left 8 predictors (see Table II): X_1 is the first order increment of the number of dry synoptic processes during the warm season with a time lag of 3 years: X_2 is the first order increment of the number of warm synoptic processes during the warm season with a time lag of 3 years; X_3 is the first order increment of the number of rainy synoptic processes during the warm season with a time lag of 3 years; X_4 is the

TABLE II Intermediate calculations(1)

Increments	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	r	S
1st stage of selection																				
(X_1)	2.37	–0.98	–0.26	–1.94	1.65	–0.98	0.45	0.69	0.93	–0.50	0.69	0.69	–1.46	1.93	0.69	–0.98	–0.26	1.41	–52	
(X_2)	2.00	–1.74	0.32	–1.93	0.32	0.32	2.00	0.50	2.00	–0.81	0.13	1.25	–1.74	0.88	1.81	–0.06	–0.43	0.69	–65	
(X_3)	–1.41	2.03	1.05	2.03	–0.42	2.03	0.56	0.56	–0.42	1.05	0.07	–1.16	0.80	–0.18	0.80	2.52	1.79	–0.18	51	
(X_4)	–2.27	0.88	–0.69	–0.02	–0.02	0.66	–1.37	–0.69	–1.59	0.21	2.46	0.21	–0.47	0.67	–0.69	–1.59	–0.24	–1.37	51	
(X_5)	2.74	–0.08	1.04	0.86	0.48	–0.08	1.61	1.05	2.55	0.67	–1.21	0.86	0.67	–0.47	0.86	1.61	0.67	3.11	–47	
(X_6)	1.18	–0.32	0.48	–1.66	0.87	–0.08	–1.43	1.58	–1.03	1.35	–1.82	0.79	–0.08	0.48	–1.19	1.50	–1.03	0.00	–43	
(X_7)	–1.07	0.53	–0.27	1.24	–0.89	–0.45	1.33	–0.89	0.45	–1.60	1.87	–0.36	–0.36	–0.89	1.69	–1.43	1.25	–0.27	40	
(X_8)	–0.88	1.56	–1.95	1.71	–0.67	–0.77	1.88	–2.28	2.13	–1.76	1.49	–1.48	1.57	–1.43	0.88	–0.03	–0.80	1.14	–52	
2nd stage of selection																				
Y_{1-2}	0.06	0.20	0.13	0.21	0.12	0.13	0.06	0.12	0.06	0.17	0.13	0.09	0.20	0.10	0.07	0.14	0.15	0.11	65	0.063
Y_{3-4}	0.02	0.20	0.13	0.17	0.11	0.19	0.10	0.12	0.07	0.15	0.18	0.09	0.13	0.10	0.12	0.15	0.16	0.08	62	0.065
Y_{5-6}	0.03	0.16	0.11	0.17	0.11	0.14	0.13	0.08	0.09	0.10	0.24	0.10	0.13	0.12	0.15	0.06	0.16	0.05	60	0.067
Y_{7-8}	0.09	0.14	0.11	0.16	0.10	0.11	0.16	0.09	0.14	0.08	0.18	0.11	0.12	0.10	0.17	0.09	0.15	0.12	40	0.076
3rd stage of selection																				
$Y_{3-4-5-6}$	0.01	0.19	0.12	0.18	0.11	0.18	0.11	0.09	0.07	0.13	0.22	0.09	0.13	0.11	0.14	0.11	0.16	0.05	65	0.063
Y	0.04	0.30	0.15	0.28	0.09	0.06	0.05	0.04	0.07	0.11	0.18	0.01	0.10	0.17	0.13					
Y'	0.01	0.21	0.12	0.21	0.11	0.16	0.08	0.10	0.05	0.15	0.18	0.08	0.18	0.10	0.10	0.12	0.17	0.08	72	0.057

first order increment of the number of warm synoptic processes during the cold season with a time lag of 3 years; X_5 is the first order increment of the number of cold synoptic processes during the cold season with a time lag of 3 years; X_6 is the first order increment of the number of dry synoptic processes during the cold season with a time lag 3 years; X_7 is the third order increment of the number of rainy synoptic processes during the cold season with a time lag of 3 years; X_8 is the tenth order increment of the number of cold synoptic processes during the warm season with a time lag of 3 years. All the above had to satisfy the following conditions:

1) the presence of a high correlation with the forecasted series ($r > 40\%$);

2) mutual independence;

3) possibility of forecasting the level of the Caspian Sea from the used set of the increments within the nearest three years (i.e., the time lag should not exceed two years);

At the second stage we filtered out predictors setting the limit correlation at $r > 60\%$ and the standard forecast error of the level at < 0.07 m. Within these limitations the latter two criteria (Y'_{7-8}) failed to satisfy them: the coefficient of correlation with the forecasted series appeared to be less than the necessary 60%, and the standard error exceeded 0.07 m.

At the last stage of selection of the informative data the correlation coefficient reached 72%, while the standard error decreased to –0.057 m. As a result the final equations used to forecast the first order increments of the level of the Caspian Sea (Y') took the form:

$$Y' = -0.029\,096 + 0.622\,836\, Y_{1-2} + 0.610\,935\, Y_{3-4-5-6}\,, \quad (1)$$

where $Y_{1-2} = 0.138\,053 - 0.001\,136\, X_1 - 0.037\,287\, X_2$, $Y_{3\text{-}4\text{-}5\text{-}6} = -0.018\,972 + 0.650\,619\, Y_{3\text{-}4} + 0.503\,278\, Y_{5\text{-}6}$, and $Y_{3-4} = 0.117\,313 + 0.027\,45\, X_3 + 0.025\,297\, X_4$, $Y_{5\text{-}6} = 0.153\,319 - 0.034\,388\, X_5 - 0.025\,336\, X_6$, $X_1, X_2, \ldots, X_6$ are the predictors.

The prognostic relation, from which the level of the Caspian Sea (Y'') is retrieved, is

$$Y'' = \Delta^1_{n-1} + Y - 27.17\,, \quad (2)$$

where Δ^1_{n-1} is the first order increment of the level of the Caspian Sea for the year preceding the forecasted; Y' is the prognostic value of the first order increment of the level of the Caspian Sea for the year of the forecast, as yielded by relation (1).

The following conclusions may be drwan from considering the described statistical model. Further increase of the level of the Caspian Sea is to be expected. In 1993 it will exceed the average level of 1992 by 12 cm, to reach –27.05 m; in 1994 – by 17 cm the level of 1993, to reach –26.76 m; in 1995 – by 8 cm the level of 1994, to reach –26.64 m. The standard forecast error is 0.06 m.

An increase of the number of informative predictors will apparently lead to higher accuracy of such forecasts. We also tried to use, apart from the number of synoptic processes and their increments of various orders, the monthly average air temperatures for the 5×5 ° grid, mainly for gridpoints in the Volga River basin (data were taken from 1891 to 1990, see Fig. 1).

We constructed a statistical model, similar to the one described above, in which we used, besides the predictors suggested above, also the monthly mean air temperatures at 16 gridpoins and their increments up to the 11th order. The effect of 30 years prehistory was also accounted for, a separte dependence was constructed for each 15 years (1935–1950, 1950–1965, 1965–1980, 1975–1990).

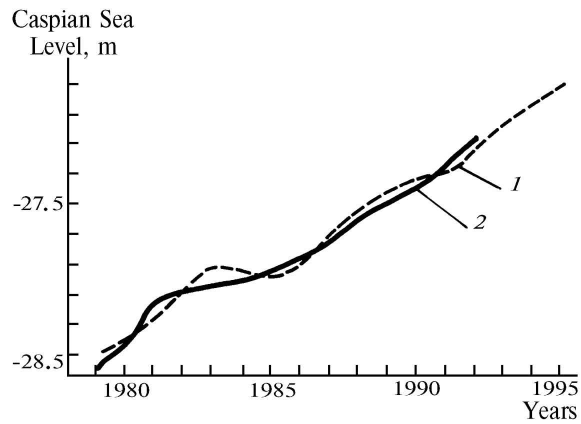

Fig. 1. The forecasted (1) and actual (2) levels of the Caspian Sea.

Bringing in these new predictors we could introduce stricter criteria at each stage of the selection. Those predictors, which had their correlation coefficients r in excess of 60 % passed the first selection stage; the second – those having $r > 55–71$ % (depending on the 15–year span considered).

The quality of the forecast involving additional predictors is much better (cf. Fig. 2 and Table III): The standard error reduced to 0.02–0.04 m, and the coefficient of correlation between the retrieved and the actual values increased to 81–97 %.

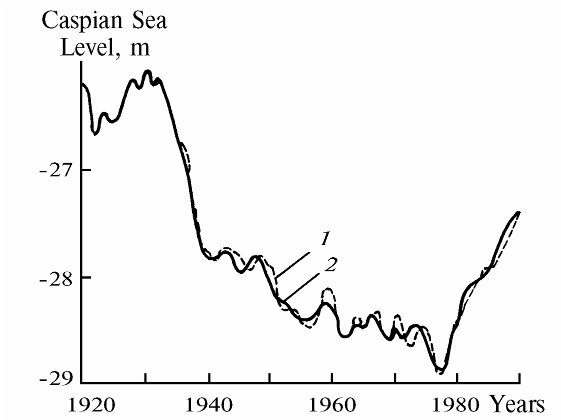

Fig. 2. The forecasted (1) and actual (2) estimates of the level of the Caspian Sea.

TABLE III Root–mean–square error σ of the retreived levels of the Caspian Sea and coefficient r of their correlations with the actual data

Years	σ	r
1935 – 1950	0.02	0.97
1950 –1965	0.02	0.85
1965 –1980	0.04	0.81
1975 –1990	0.02	0.85

Thus the described forecasting technique makes it possible to recover knee–bends in the trend of the level of the Caspian Sea. The forecast results may be improved by introducing new predictors into the model.

REFERENCES

1. A.G. Ivakhenko and V.G. Lapa, *Forecasting Stochastic Processes* (Naukova Dumka. Kiev, 1971).
2. A.A. Isaev, L.V. Klimenko, V.B. Sundareev , and N.N. Borisova in: *Atmospheric Circulation and Climate Oscillations* (Kazan University Publishing House, Kazan, 1989) p. 66–74.
3. A.A. Isaev, Tr. VNIIGMI–MCD, No. **140,** 104–143 (1987).
4. A.A. Isaev, and M.A. Petrosyants, ed., *Climatic Resources and their Applications* (State University, Moscow).
5. L.V. Klimenko, Tr. Mos. CGNS,No. 3, 57–62 (1989).
6. L.V. Klimenko, ibid., 51–66 (1989).
7. L.V. Klimenko, *Synoptic–Climatic Typification of Atmospheric Process and their Catalogue*. (State University, Moscow, 1976).
8. Proc. All–Union Conf. Problems of the Caspian Sea. June 3–5, 1991 (Gurjev, 1991).
9. Technical and Economical Foundation for the Protection of Objects of the National Economy and of the Coastal Habitants of the Caspian Sea within the Russian Federation. Report of the VNNPK "Kaspii" at the Roskomvod of the Ministry of Ecology of the Russian Federation (1992).