www.czasopisma.pan.pl PAN www.journals.pan.pl

JOURNAL OF WATER AND LAND DEVELOPMENT

e-ISSN 2083-4535

Polish Academy of Sciences (PAN) Institute of Technology and Life Sciences – National Research Institute (ITP – PIB)

JOURNAL OF WATER AND LAND DEVELOPMENT
DOI: 10.24425/jwld.2021.139009
2021, No. 51 (X–XII): 11–16

Application of the Triple Diagram Method in forecasting lake water level, on the example of Lake Charzykowskie

Adam Piasecki[1)] ✉, Wojciech T. Witkowski[2)]

[1)] Nicolaus Copernicus University, Faculty of Earth Sciences and Spatial Management, ul. Lwowska 1, 87-100, Toruń, Poland

[2)] AGH University of Science and Technology, Faculty of Mine Surveying and Environmental Engineering, Krakow, Poland

RECEIVED 5.12.2019 REVIEWED 11.12.2020 ACCEPTED 04.02.2021

Abstract: The work focused on forecasting changes in lake water level. The study employed the Triple Diagram Method (TDM) using geostatistical tools. TDM estimates the value by information from an earlier two periods of observation, refers as lags. The best results were obtained for data with an average a 1-week lag. At the significance level of 1σ, a the forecast error of ±2 cm was obtained. Using separate data for warm and cold months did not improve the efficiency of TDM. At the same time, analysis of observations from warm and cold months explained trends visible in the distribution of year-round data. The methodology, built on case study and proposed evaluation criteria, may function as a universal solution. The proposed methodology can be used to effectively manage water-level fluctuations both in postglacial lakes and in any case of water-level fluctuation.

Keywords: geostatistics, kriging, lake, Lake Charzykowskie, Triple Diagram Method (TDM), water level forecasting, water resources

INTRODUCTION

Lakes are environmental features that are highly sensitive to changes in natural and anthropogenic conditions. Thus, they can indirectly provide information about changes in the natural environment of a given region. Therefore, systematic observations of lakes have been conducted around the world for many years, and water level is one of the basic parameters measured. The extent and direction of water level fluctuation (WLF) of lakes displays certain regularities and variability in the annual cycle. WLF is influenced by many factors, including climatic factors (precipitation, evaporation), lake basin depth, the lake's relationship with groundwaters, the nature of flow capacity, the surface area of the lake and its catchment area, and anthropopressure [Plewa *et al.* 2017]. In Poland, the highest water levels are usually seen in spring and summer, and the lowest in autumn and winter [Bajkiewicz-Grabowska 2005].

WLF analysis of lakes is frequently discussed in the literature. In recent years, WLF modelling and forecasting approaches have gained most in popularity. For this purpose, advanced mathematical methods are used, such as: Artificial Neural Networks [Altunkaynak 2007; Piasecki *et al.* 2018], Support Vector Regression [Buyukyildiz *et al.* 2014; Noury *et al.* 2014], Artificial Network Fuzzy Interference Systems [Kisi *et al.* 2012; Sanikhani *et al.* 2015] and Echo State Networks [Abrahart *et al.* 2012; Coulibaly 2010]. Altunkaynak *et al.* [2003] presented an interesting alternative WLF analysis method called the Triple Diagram Method (TDM). This method has broad application potential that extends beyond purely hydrological analyses [Özger, Şen 2007; Piasecki *et al.* 2019]. To date, no attempt has been made to verify the effectiveness of TDM on lakes other than Lake Van.

The present study innovatively sought to evaluate whether: (1) it is possible to successfully forecast the water level of small glacial lakes using TDM; (2) separate forecasting of water levels in the cool and warm half-year increases the effectiveness of TDM.

STUDY METHODS

STUDY AREA

The research object chosen to test TDM is Lake Charzykowskie. The lake is located in northern Poland in the Pomeranian Lakeland. The area of the lake is 13.36 km^2, and its volume is 134.533 mln m^3 [CHOIŃSKI 2006] According to PASIERBSKI [1975], in the pre-boreal period Lake Charzykowskie was almost three times its current size. The lake's water level was about seven metres higher than at present (121.0 m a.s.l.). The lowering of the water level was a consequence of natural processes related to the melting of ice blocks and subsidence of the bed [PASIERBSKI 1975]. The lake basin formed in a glacial trough, which is reflected in the elongation index of its bottom exceeding 4.1. The maximum depth is 30.5 m and the average depth is 9.8 m [CHOIŃSKI 2006]. It is a through-flow lake and the entire water body is theoretically refreshed every 230 days. It is fed by four main tributaries, of which the Brda River is the most significant (Fig. 1). The average Brda flow rate before its mouth to Lake Charzykowskie is 5.54 $m^3 \cdot s^{-1}$ [BARAŃCZUK, BOROWIAK 2010]. The lake's catchment area slightly exceeds 920 km^2. Forests (mainly coniferous) cover over 81.5% of the catchment area, and agricultural land only 14.5%. The lake is used for recreation and tourism, and the largest settlement on its shores is the village of Charzykowy. The lake and a large part of its catchment belongs to the protected area of Zaborski Landscape Park.

The water level in Lake Charzykowskie in 1960–2015 was moderately variable. The amplitude of changes in water level for the period was 87 cm, with the lowest values (0 cm) in February and April of 1996, and the highest (87 cm) in September of 1980 and 2001 (Fig. 2). Day-to-day water level fluctuations ranged from 0 to 12 cm, while weekly fluctuations ranged from 0 to 24.14 cm,

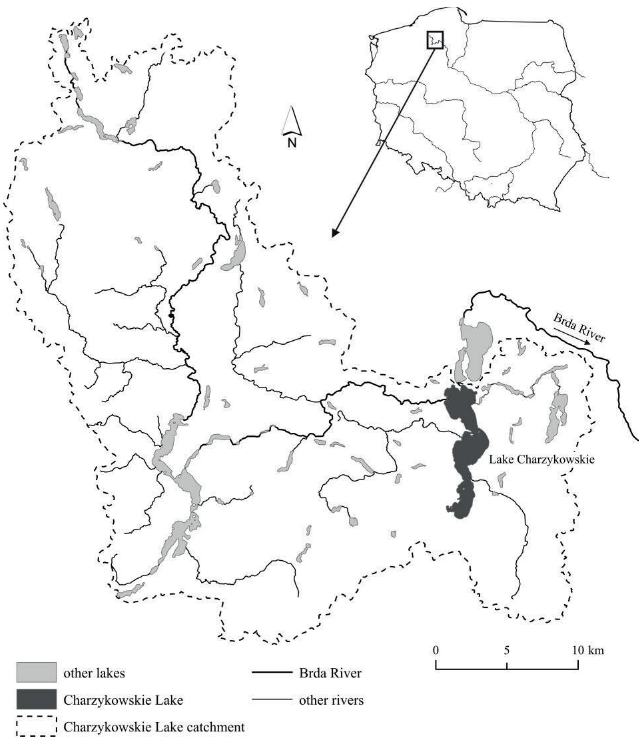

Fig. 1. Drainage basin of the Lake Charzykowskie; source: own elaboration

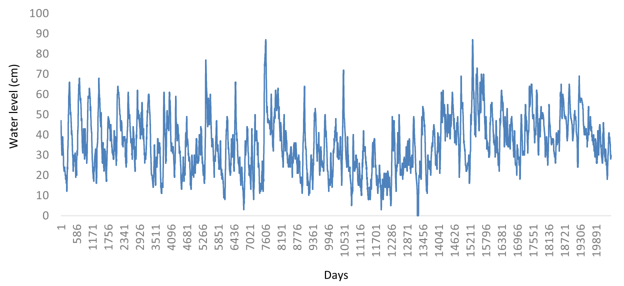

Fig. 2. Time serie of Lake Charzykowskie water level; source: own study

and monthly fluctuations from 0 to 40.5 cm. Lake Charzykowskie has an annual variability of water level that is typical for most lakes in Northern Poland. During the year, there are usually two extremes (maximum and minimum water levels), between which the water level steadily increases (from spring to autumn) and then decreases (from late autumn to spring). The highest values occur in September and October, and the lowest in May and June.

METHODS

The study used a geostatistical data modelling technique. One of its main benefits is its ability to describe a natural phenomenon using a system of two different variables. A basic tool of geostatistics is the semi-variance function $\gamma(h)$ expressed by Equation (1).

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (z_i - z_{i+h})^2 \qquad (1)$$

This function calculates differences in values (z_i, z_{i+h}) between points in space. The calculations analyse all pairs of points $N(h)$ separated from each other by the vector h. Semi-variance analysis allows kriging equations to be created and system variance to be minimised. This process can be represented in matrix notation in the form of Equation (2), where γ_{ij} is the semi-variance between two known points in space i and j; γ_{io} is the semi-variance between the known point i and the modelled value at the new point o; the value is the Lagrange constant of the kriging system of equations.

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} & 1 \\ \gamma_{21} & \gamma_{22} & & \gamma_{2n} & 1 \\ \vdots & & \ddots & \vdots & \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{nn} & 1 \\ 1 & 1 & & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{1o} \\ \gamma_{2o} \\ \vdots \\ \gamma_{no} \\ 1 \end{bmatrix} \qquad (2)$$

The solution to the presented system of equations can be used to determine the weight of the w_i values assigned to each point in space. By linear combination of the weight values and the observed z_i value of the point, the value of the new point o in space is determined by Equation (3).

$$z_o = \sum_{i=1}^{N} w_i z_i \qquad (3)$$

Geostatic modelling was conducted on the empirical data, which were diurnal water levels in Lake Charzykowskie for 1960–2015 provided by the Institute of Meteorology and Water Management – National Research Institute of Poland. The set of 20,454 observations was divided into two datasets. Geostatistical modelling was performed on the set of 18,627 observations for 1960–2010. The effectiveness of the created models was verified against the set of observations from the period 2011–2015 with a total of 1,827 data points.

The TDM method [Altunkaynak 2007] that was employed uses geostatistical modelling to describe a phenomenon. In TDM, information from any given point in time is described in terms of information from an earlier period, referred to as a "lag". The observation (Lag0) is assigned information about the phenomenon with a single lag (Lag1) and a double lag (Lag2). For a lag of one week example, observation of the water level on February 28, 2015 (Lag0) is described in terms of two lagged observations, from February 21, 2015 (Lag1) and February 14, 2015 (Lag2). The Lag0 variable can therefore be expressed in the distribution of Lag1 and Lag2 observations forming a TDM.

TDM studies were performed with a lag of one day, one week and two weeks and two variants of a 1-month lag. In the first 1-month lag variant, the observations were taken directly from measurements, and in the second the average value for a given period was determined for the lag. This approach aimed to minimise the impact of outliers. Additional analyses were performed using data divided into warm months (V–X) and cold (I–IV and XI–XII).

In order to compare incremental values of modelling results (x_p) with incremental values of observations from 2011–2015 (x_o), relative prediction errors were determined (δ_i). These errors were determined in relation to the maximum predicted value (max $\{x_{oi}\}$) from the dataset (4). This allowed for comparison between results obtained from different modelling variants with different lags. For the errors thus obtained (δ_i) their distributions were analysed and standard deviations (σ_{δ_i}) of errors were determined. The results were compared against the semi-variance values obtained from geostatistical modelling.

$$\delta_i = \frac{|x_o - x_p|}{\max\{x_{oi}\}} 100\% \tag{4}$$

RESULTS AND DISCUSSION

The first stage in TDM work was to analyse empirical semi-variograms (Fig. 3) in order to achieve theoretical model fit. The power function was identified as the best theoretical model for all analysed examples. As the lag value increases in successive data sets, the increasing influence of the "nugget effect" is seen (γ_{NE}). For data with a one-day lag, this value is the smallest. When a 1-month lag is assumed in calculations, the largest value of unexplained variance of the model is visible, taking into account observations from the whole year. Employing averaged lag values in calculations reduces the γ_{NE} value in each case. When using data for cold months, the γ_{NE} is less than for the model built for year-round observations. The highest unexplained variance was found for observations from warm months. Due to the climatic conditions of Central and Eastern Europe, changes in the water level of lakes are usually slight in the cold months. In these months, the water resources of lakes and drainage basins are replenished due to lower evaporation. This is despite sums of precipitation being lower than in the warmer months. In the cold

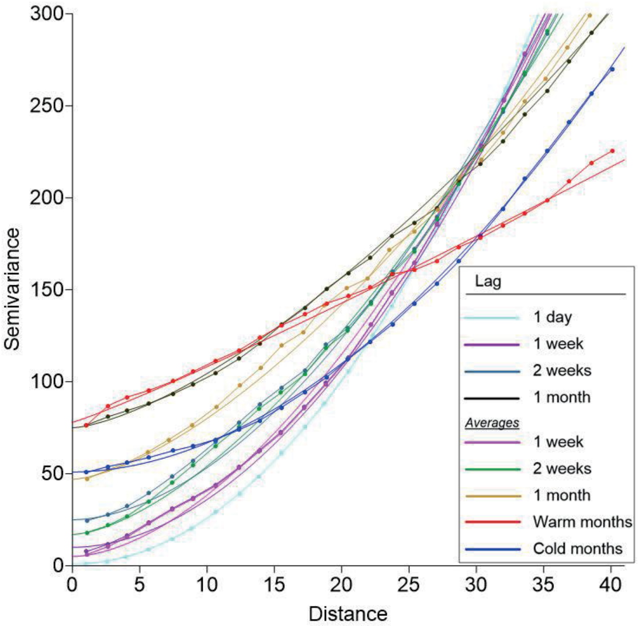

Fig. 3. Semi-variograms for lags of one day, one week, two weeks and one month and in corresponding variants with observations comprising averages for one week, two weeks, one month and also cold and warm months; source: own study

months, precipitation often occurs as snow, which, combined with negative air temperatures, means that the water is temporarily stored. Thus, also, the lower the air temperatures in colder months, the lower the "mobility" of water, which also directly results in lake water levels varying only slightly.

Using the created theoretical semi-variograms, geostatistical models of water level were constructed for various lag values. The forecast results are presented as distributions of values in a Lag1 and Lag2 system for selected periods only. The axes show water levels at selected lags. The geostatistical models are presented as rasters. For the test sets, the difference between the water level reading and the forecast result as per the triple diagram was determined. The scale of errors is presented in a colour scale.

For the 1-day lag, the forecast has a linear relationship in the Lag1–Lag2 system (Fig. 4). The observed water level is closely correlated with the water level from one and two days previously.

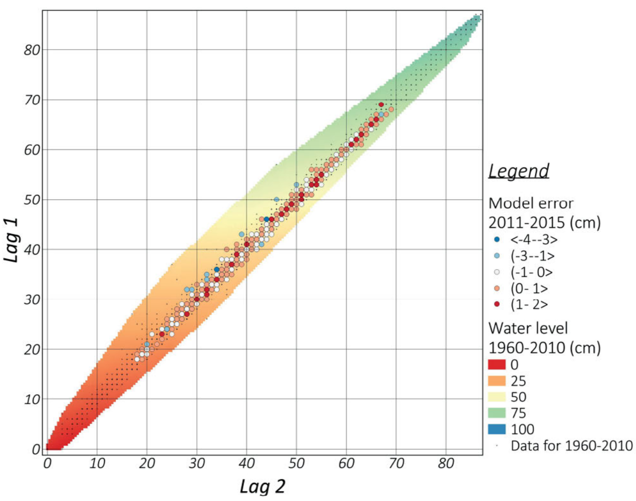

Fig. 4. Forecast result using Triple Diagram Method for a 1-day lag; source: own study

In this case, the errors in forecasts made on the test set ranged from –4 cm to +2 cm. It should be noted that, in practice, a one-day lag may be too short for planning and managing water resources.

Models with one- and two-week lags had weaker forecast results, with values in the ranges from –13 to +11 cm and from –25 to +18 cm, respectively. When averaged reading values were used, these ranges were several cm narrower. These observations confirm the conclusions obtained from when the empirical semi-variograms were analysed.

Models developed using 1-month lag data had forecast errors ranging from –37 cm to +17 cm (Fig. 5) and from –33 cm to +19 cm for the averaged readings (Fig. 6). On both charts, two groups of points clearly stand out. The bottom of the graph shows observations from the test set where the model underestimated values. The middle part shows points with overestimated values. The situation is similar on the diagram for the model with 1-month-average readings. Using data with a 1-month lag decreased the quality of results, but may at the same time be a compromise between the tool's accuracy and its practical applicability.

In the model built using cold month data, forecasting errors ranging from –21 cm to +17 cm were obtained (Fig. 7). For the warm month data, the results were worse, ranging from –28 cm to +25 cm (Fig. 8).

Analysing the results of forecasts using observations from warm and cold months showed a shift of both graphs relative to each other (Figs. 7, 8); the results from cold months are distributed across lower Lag1 values than those from the warmer months. Similar patterns are seen in the diagrams for year-round observations divided into 1-month periods (Figs. 5, 6). Two groups of points stand out in the forecasting results for the 1-month lag. The first group of underestimated values can be seen at the bottom of the charts (in blue). The second group, representing overestimated values, appears in the region of larger Lag1 values (in red). Thus, the distribution of data in the TDM broken down into cold months (Fig. 7) and warm (Fig. 8) may explained the previous concentrations of data in Figures 5 and 6.

The information about the variability of lake water level in the analysed period (2011–2015) allow an optimal solution to be selected. The error was calculated using the daily, weekly and monthly variability of the water table (Tab. 1). The variability of

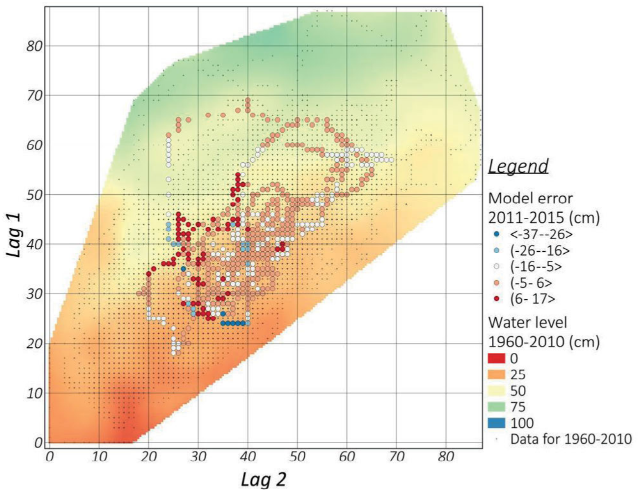

Fig. 5. Forecast result using Triple Diagram Method for a 1-month lag; source: own study

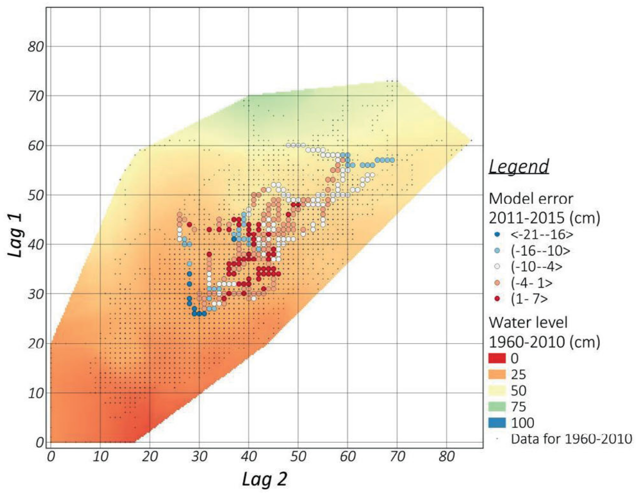

Fig. 7. Forecast result using Triple Diagram Method for cold month observations; source: own study

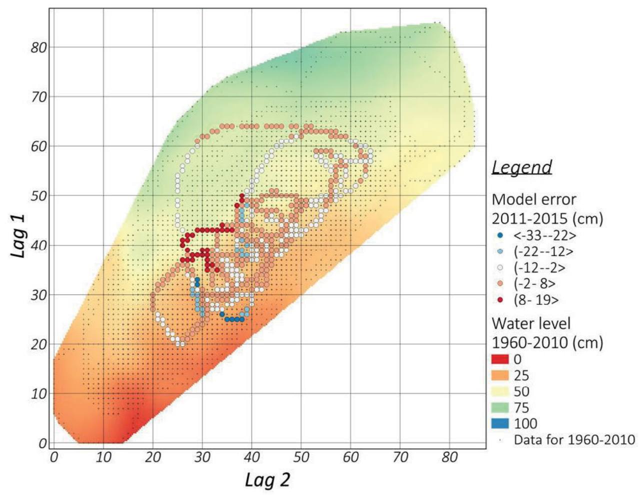

Fig. 6. Forecast result using Triple Diagram Method for 1-month averages of observations; source: own study

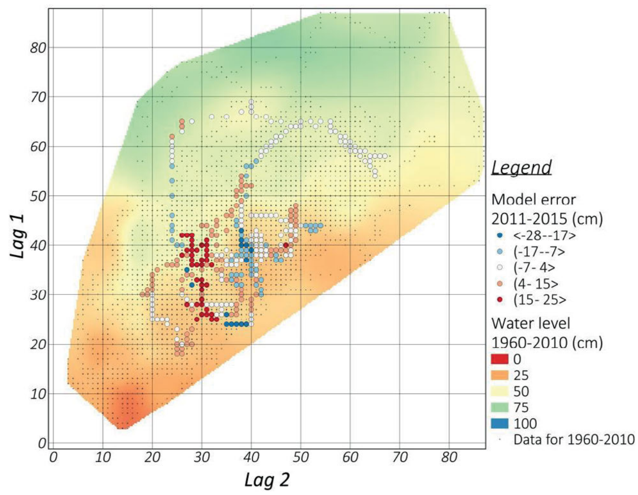

Fig. 8. Forecast result using Triple Diagram Method for warm month observations; source: own study

Table 1. Calculated characteristics for individual lags considering in study

Characteristics	Values for lags ID								
	1 1 day	2 1 week	3 1 week (avg)	4 2 weeks	5 2 weeks (avg)	6 2 weeks (avg)	7 2 weeks (avg)	8 2 weeks (avg)	9 2 weeks (avg)
max$\{x_{oi}\}$(cm)	4	16	15	28	22	39	32	20	39
δ_i%	–2.1	–0.7	0.5	3.0	4.0	7.2	6.6	3.4	–4.4
$\sigma_{\delta i}$ (%)	16.6	14.6	13.5	14.5	16.6	19.3	21.3	25.1	27.8
min $\{\delta_i\}$(%)	–59.5	–66.7	–56.0	–63.6	–70.5	–43.3	–58.8	–33.5	–64.6
max $\{\delta_i\}$(%)	96.7	83.2	63.3	88.6	80.9	95.4	102.5	106.0	71.5

Source: own study.

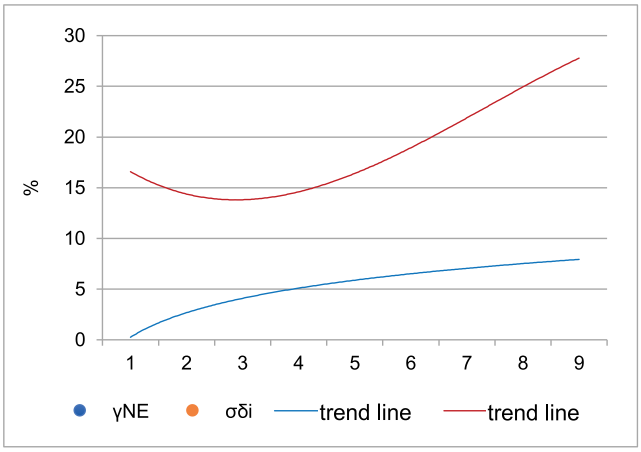

Fig. 9. Values of "nugget effect" from modelling of semi-variograms (γ_{NE}) in relation to the standard deviation of relative errors () with fitted trend functions; 1–10 at horizontal axis corresponds to ID numbers from Table 1; source: own study

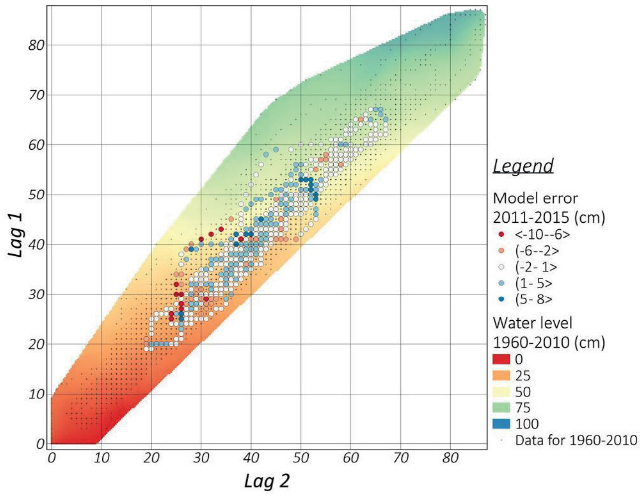

Fig. 10. Forecast result using Triple Diagram Method for 1-week averages of observations; source: own study

observations increases as lag values increase, which is to be expected due to the nature of the phenomenon. The mean relative error values in each set were determined (Tab. 1). These values range around zero, relative to the maximum and minimum values in sets, indicating the lack of a trend in errors. The values were compared with the "nugget effect" (γ_{NE}) values obtained from semi-variogram modelling (Fig. 9). As lag values increase, so too do both and γ_{NE}. The γ_{NE} values show a linear trend. An increase in lag results in a decrease in data correlation and thus an increase in semi-variance. The values were determined from the model tested on the observation set. They show a non-linear trend that is close to a quadratic function. The lowest value of was obtained for the 1-week average observations. For this model, the γ_{NE} value is one of the lowest, which indicates a high degree of correlation in observations.

Analysis of the results revealed the optimal model, which was built on 1-week lag average observations. On the model the forecast errors for observations from 2011–2015 was presented (Fig. 10). They are in the range from –10 cm to +8 cm with maximum increments of up to ±15 cm. Furthermore, at a confidence level of 1σ, the difference between observation and forecast does not exceed ±2.0 cm.

CONCLUSIONS

The analysis carried out in the work showed that it is possible to effectively forecast the water level of small glacial lakes using Triple Diagram Method (TDM). The best results for the Lake Charzykowskie were obtained for data with 1-week averages lag. Forecast errors in the range –10 cm to +8 cm were obtained. At level 1σ, the error values do not exceed ±2.0 cm. In addition, TDM was found to be not more effective when water levels for warm and cold months were forecast separately. The model built for the 1-day cold-month data did not show a linear correlation between Lag1 and Lag2 on the TDM chart, in contrast to the 1-day lag data, which did. Data variability in the Lag1 and Lag2 system was greater than for the 1-day lag data. Nevertheless, the using the 1-month lag observations divided to the cold and warm seasons did not improve the effectiveness of TDM modelling. The proposed criterion for comparing results using relative error showed non-linear characteristics with an indication of the model built on average 1-week lag observations. This model is a compromise between data variation over time (lag values applied in TDM) and achievable forecast accuracy.

The results show the great potential of the proposed methodology for forecasting the water level of small glacial lakes.

However, more research is needed, covering more lakes. If the effectiveness of the proposed methodology is confirmed, it can be practically applied as a tool for use in managing lake water resources. It may be particularly helpful in managing water level fluctuations in lakes with hydroelectric power stations or those used for fishing purposes. To some extent, it can also be used in extraordinary flood-related situations.

The time interval used in TDM construction can be selected depending on the assumed goals and the required forecast accuracy. Thus, also, although the best results were obtained for average 1-week.

Further studies are planned on the use of TDM in forecasting lake water levels. They will aim to verify the effectiveness of the method for lakes of different hydrological and morphological characteristics. In addition, it is also planned to compare the results of TDM forecasts against other geostatistical methods, e.g. co-kriging.

REFERENCES

ABRAHART R.J., MOUNT N.J., SHAMSELDIN A.Y. 2012. Neuroemulation: definition and key benefits for water resources research. Hydrological Sciences Journal. Vol. 57(3) p. 407–423. DOI 10.1080/02626667.2012.658401.

ALTUNKAYNAK A., ÖZGER M., SEN Z. 2003. Triple diagram model of level fluctuations in Lake Van, Turkey. Hydrology and Earth System Sciences. Vol. 7(2) p. 235-244. DOI 10.5194/hess-7-235-2003.

ALTUNKAYNAK A. 2007. Forecasting surface water level fluctuations of Lake Van by artificial neural networks. Water Resources Management. Vol. 21(2) p. 399–408. DOI 10.1007/s11269-006-9022-6.

BAJKIEWICZ-GRABOWSKA E. 2005. Jeziora. Zmiany stanów wody. W: Geografia fizyczna Polski [Lakes. Changes in water levels. In: Physical Geography of Poland]. Ed. A. Richling, K. Ostaszewska. Warszawa. Wydaw. Nauk. PWN p. 173–183.

BARAŃCZUK J., BOROWIAK D. 2010. Jezioro Charzykowskie. W: Atlas jezior Zaborskiego Parku Krajobrazowego [Lake Charzykowskie. In: Atlas of the Zaborski Landscape Park Lakes]. Ed. J. Barańczuk, D. Borowiak. Gdańsk, Charzykowy. Katedra Limnologii Uniwersytetu Gdańskiego; Pomorski Zespół Parków Krajobrazowych p. 32–41.

BUYUKYILDIZ M., TEZEL G., YILMAZ V. 2014. Estimation of the change in lake water level by artificial intelligence methods. Water Resources Management. Vol. 28(13) p. 4747–4763. DOI 10.1007/s11269-014-0773-1.

CHOIŃSKI A. 2006. Katalog jezior Polski [The catalogue of Polish lakes]. Poznań. Wydaw. Nauk. UAM. ISBN 8323217327 pp. 599.

COULIBALY P. 2010. Reservoir computing approach to Great Lakes water level forecasting. Journal of Hydrology. Vol. 381(1) p. 76–88. DOI 10.1016/j.jhydrol.2009.11.027.

KISI O., SHIRI J., NIKOOFAR B. 2012. Forecasting daily lake levels using artificial intelligence approaches. Computers & Geosciences. Vol. 41 p. 169–180. DOI 10.1016/j.cageo.2011.08.027.

NOURY M., SEDGHI H., BABAZEDEH H., FAHMI H. 2014. Urmia lake water level fluctuation hydro informatics modeling using support vector machine and conjunction of wavelet and neural network. Water Resources. Vol. 41(3) p. 261–269. DOI 10.1134/S0097807814030129.

ÖZGER M., ŞEN Z. 2007. Triple diagram method for the prediction of wave height and period. Ocean Engineering. Vol. 34(7) p. 1060–1068.

PASIERBSKI M. 1975. Uwagi o genezie niecki jeziora Charzykowskiego [Remarks on the genesis of the Charzykowskie lake basin]. Acta Universitatis Nicolai Copernici. Geografia. Vol. 11 p. 101–103.

PIASECKI A., JURASZ J., ADAMOWSKI J.F. 2018. Forecasting surface water-level fluctuations of a small glacial lake in Poland using a wavelet-based artificial intelligence method. Acta Geophysica. Vol. 66(5) p. 1093–1107. DOI 10.1007/s11600-018-0183-5.

PIASECKI A., JURASZ J., WITKOWSKI W. 2019. Application of triple diagram method in medium-term water consumption forecasting. In: Infrastructure and Environment. Ed. A. Krakowiak-Bal, M. Vaverkova. Springer p. 59–66.

PLEWA K., WRZESIŃSKI D., BACZYŃSKA A. 2017. Przestrzenne i czasowe zróżnicowanie amplitud stanów wody jezior w Polsce w latach 1981–2015 [Spatial and temporal differentiation of the amplitudes of lake water levels in Poland in the years 1981–2015]. Badania Fizjograficzne. Ser. A Geografia Fizyczna. Vol. 68 p. 115–126. DOI 10.14746/bfg.2017.8.8.

SANIKHANI H., KISI O., KIAFAR H., GHAVIDEL S. 2015. Comparison of different data-driven approaches for modeling lake level fluctuations: The case of Manyas and Tuz Lakes (Turkey). Water Resources Management. Vol. 29(5) p. 1557–1574. DOI 10.1007/s11269-014-0894-6.