Managing Intellectual Capital and Innovation for Sustainable and Inclusive Society
27–29 May 2015
Bari • Italy
Management, Knowledge and Learning
Joint International Conference 2015
Technology, Innovation and Industrial Management

APPLICATION OF INTELLECTUAL DSS TO MEDIUM-TERM FORECASTING OF THE SEA ICE AREA IN THE NORTHERN HEMISPHERE

Oleg Bukharov
National Research University - Higher School of Economics, Russia
oleg_bukh.box@mail.ru

Dmitry Bogolyubov
National Research University - Higher School of Economics, Russia
bogolub@mail.ru

Valery Fedorov
Moscow State University, Russia
fedorov.msu@mail.ru

Pavel Grebennikov
Moscow State University, Russia

Abstract:
The paper is devoted to the description of a new multi-purpose intellectual decision support system. We present the algorithms used and the results achieved in applying the system to analyzing and forecasting the sea ice area in the Northern Hemisphere. The impact of solar radiation on the changes in the sea ice area was confirmed. Application of interval neural nets to medium-term forecasting of sea ice area changes was justified.

Keywords: decision support system, DSS, sea ice, solar radiation, neural network, genetic algorithm

1. INTRODUCTION

Climate is one of the most important characteristics of natural environment. Thus, the tasks of exploring the global climate change and factors affecting it are essential for natural science. Sea and mountain ice as natural environmental element depends on climatic changes. Revealing the causes of climatic changes is the main task of climate research and forecasting (Budyiko, 1974).

Sea ice extent is constantly changing. Today's sea ice area could be slightly different from yesterday's. If the time interval between ice area measurements is short, e.g. one day, it is next to impossible to make a meaningful forecast for a few years ahead even based on the bulk of historical data regarding ice area and the effect of environmental factors. To make a meaningful medium-term forecast, it is necessary to decrease the number of values to be analyzed and forecast. As Northern Hemisphere ice tends to melt is summer and grow in winter, the annual area measurements may be presented as a single interval value comprising a continuum from the minimum to the maximum area. Applying interval arithmetic to analyze and forecast the ice area, we can replace an annual range of measurements with a single interval value.

Our decision support system combining interval neural nets and a genetic algorithm was used as a forecasting tool evaluating dependences between the ice area and the environmental factors.

2. THE STRUCTURE OF THE DEVELOPED DSS

2.1. Application of the Neural Nets

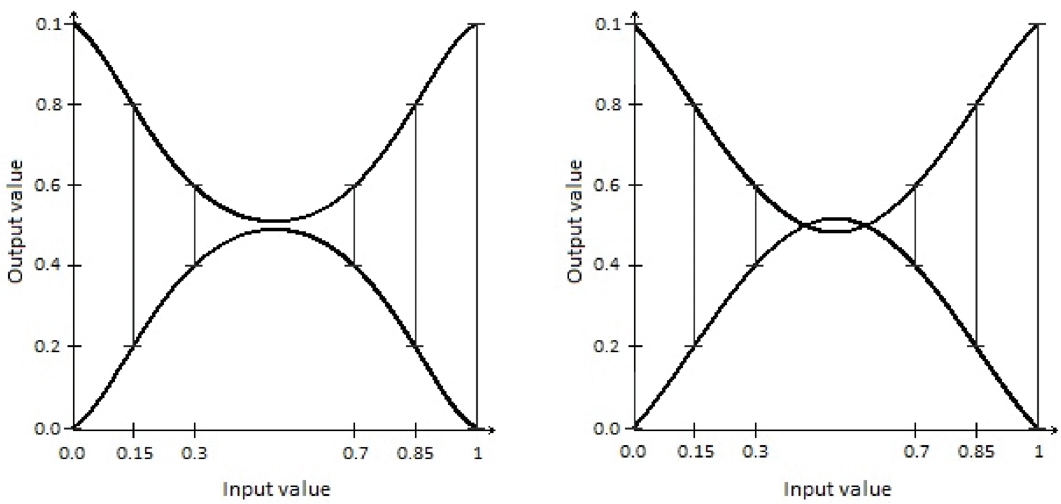
According to A.N. Gorban (Gorban, 1998), any continuous multivariable function can be approximated with preassigned precision by a neural net if the net has a continuous one-variable activation function.
Being a model of complicated multidimensional nonlinear regression, the neural network is more accurate than the above-mentioned methods and has a number of other advantages (Tsaregorodtsev, 2014):

1. Possibility to work with non-informative noise input signals: the neural network can reject them as useless for the task solution;
2. Possibility to work with polytypic information (continuous and discrete, qualitative and quantitative data types), which is considered to be a difficult task for statistical methods;
3. Given several outputs, the neural network can solve a number of problems simultaneously;
4. There are algorithms for inverse task solution with a neural network trained to resolve a specific task. For example, it is possible to connect new neural network inputs with the outputs of a current neural network and train the new network to produce the previous network inputs as its own outputs;
5. A neural network has fewer requirements for the qualification of its user compared to complicated statistical models capable of obtaining similar results;
6. Having initially set synaptic weights of a neural network, it is possible to recreate and check the suggested statistical models as well as improve them by network training (Haykin, 2006).

2.2. Interval neural net

When forecasting the intervals of values, it is necessary to use interval neural networks. As professor Ishibuchi visually illustrated in his paper (Ishibuchi, 1991), the application of two standard neural networks for this purpose may cause forecast errors when the predicted value of the upper limit of the interval is lower than that of the lower one.

Figure 1: The left graph shows the result of an interval neural network, the right graph shows the result of two standard neural networks. The vertical lines represent the output intervals of training sets.



An interval neural network is a system of interconnected and interacting interval neurons. Inputs and outputs of interval neurons are intervals, each of them being a continuum set of values between the two limit values.

2.3. Application of genetic algorithms

Putting all known parameters into the inputs of the neural network will considerably slow down its operation and increase the probability of finding non-existent dependences. On the other hand, it is possible to overlook some of the parameters which really impact the predicted value, putting in only those which seem to be the most significant.
We have coped with this problem by using a genetic algorithm. The table below compares a genetic algorithm with some other methods applied to find the most significant parameters:

Table 1: Comparative characteristics of the methods

Method	Linear dependency detection	Functional dependency detection	Process of obtaining the first significant results
Successive increase in the number of significant parameters	Yes	Yes	Slow
Correlation analysis	Yes	No	Fast
Genetic algorithm	Yes	Yes	Fast

Genetic algorithms are adaptive search methods which are currently used to solve functional optimization tasks. They are based on the genetic processes of biological organisms: biological populations develop throughout several generations in accordance with the laws of natural selection and the "survival of the fittest" principle discovered by Charles Darwin. Imitating this process, genetic algorithms are capable of "developing" solutions to real tasks if the latter are coded appropriately.

2.4. The Algorithm of the Developed System

Prior to starting the search algorithm for optimal neural network to forecast a required parameter, the user is able to specify the following settings:

1. The name and path of the CSV file containing the time series for network learning or forecasting;
2. The input column numbers and the predicted value column number;
3. The input amount range;

4. The time window size range for forecasting inputs (the quantity of successive time series values to be used as a forecasting basis);
5. The time window size for the value to be predicted (the quantity of successive values to be predicted);
6. The time lag between the input window and the window of predicted values (can be negative);
7. The upper and lower limits of the values presented (for normalization);
8. The genetic algorithm population size;
9. The number of hidden neurons in the neural networks;
10. The number of neural network learning cycles;
11. The number of neural networks learning with the same set of inputs;
12. The maximal number of iterations.

The system algorithm:

1. The first step of the search algorithm is the creation of initial generations. Random sets of input parameters are generated. Each set is limited to a predefined input amount and time window size range.
2. The next step is to train a set of neural networks for each set of parameters. Learning is carried out using CUDA for parallel computing on GPU, which significantly reduces the algorithm's operating time.
3. If there is a network working better than the others (i.e. with a smaller error) among the trained networks, the parameters of this network are saved in the high-quality network pool. As soon as the first network is in the pool, the forecasting algorithm can work simultaneously with the search algorithm, choosing a network with the smallest error.
4. If the population has not converged or the maximum number of iterations has not been exceeded, the algorithm identifies sets of the best individuals with the smallest forecasting error.
5. A new generation is produced by crossover and mutations of the selected best sets.
6. The algorithm is completed if the maximum number of iterations has been exceeded or in the case of population convergence. Otherwise, the algorithm is repeated from step 2.

If the population has converged, the algorithm has got the optimal set of inputs and the corresponding neural network has been saved in the pool of high-quality networks. (Bukharov, 2013)

3. DETERMINING DEPENDENCES AND FORECASTING SEA ICE AREA

Obviously, solar radiation is the key factor influencing the climate and life on Earth as a whole. It is a prerequisite to all meteorological phenomena and processes in the atmosphere and on the Earth's surface (Voyeykov, 1948). Therefore, the variations of solar radiation reaching the upper limit of the atmosphere and its spatial (latitudinal) distribution are analyzed as the key factors influencing the ice indicators of global climate change.

The time series of the following annual indicators from 1870 to 2007 were offered as input data for the system:

- The sea ice area interval values (from minimum in the summer to maximum in the winter);
- The incoming annual solar radiation in the Northern Hemispheres;
- The incoming solar radiation saved up in the Northern Hemispheres in a year;
- The sequence number of the year.

When the system was applied to analyze and forecast the sea ice area, the networks using the incoming solar radiation as one of the input parameters were always in the pool of high-quality networks. It shows that this factor was always selected by the genetic algorithm as having a significant impact on the predicted value.

Comparison of the resulting forecasts with the direct multi-year target values known in advance confirms the forecast accuracy and sufficiency of historical knowledge of the sea ice area and incoming radiation calculation for adequate sea ice area forecasts.

Table 2: Forecasting quality depending on forecasting depth

Forecasting depth	Normalized mean squared error
1 year	0.00234333
2 years	0.00277834
3 years	0.00367193
4 years	0.00531844
5 years	0.00584955

The above results were achieved by giving the system the historical data on sea ice area and incoming solar radiation in 1870-1980. Test forecasting and forecast quality assessment were carried out based on the historical data from 1980 to 2007.

We used the following formula to calculate the normalized mean squared forecasting error for interval values:

$$Error = \frac{1}{N}\sum_{i=1}^{N}\frac{\left(\hat{x}_i^L - x_i^L\right)^2 + \left(\hat{x}_i^U - x_i^U\right)^2}{2\cdot\Delta x^2},$$

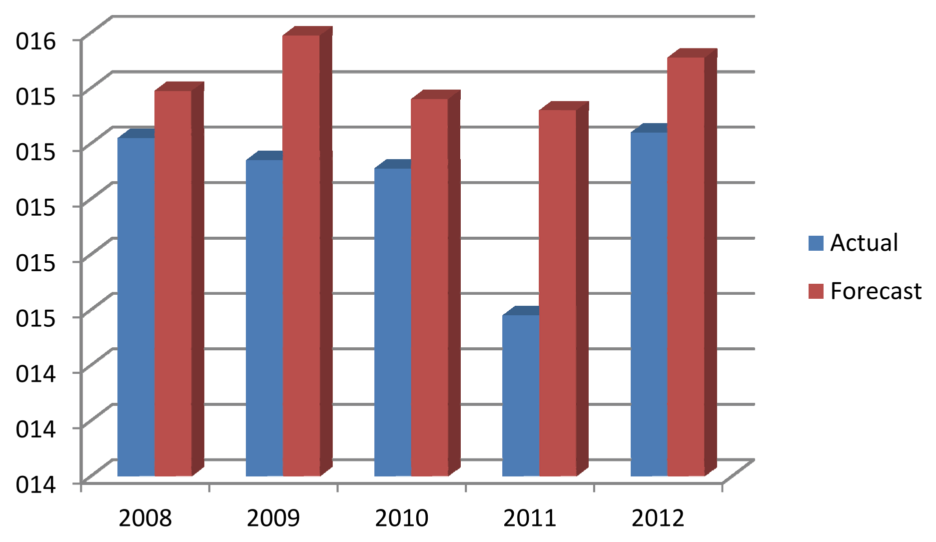
where $(\hat{x}^L;\hat{x}^U)$ is a forecasted interval value, $(x^L;x^U)$ is a real interval value, x^L is the lower limit of the interval, x^U is the upper limit of the interval, *i* is the sequential number of the forecast.

The upper limits of sea ice area for the period from 2008 to 2012 obtained by the system were compared with the sea ice area satellite measurements (Sea Ice Index, 2013).

Table 3: The upper limits of sea ice area forecasts, million square kilometers

Year	Satellite measurements	Forecasted values
2008	15.22	15.39
2009	15.14	15.59
2010	15.11	15.36
2011	14.58	15.32
2012	15.24	15.51
Mean	15.06	15.44

Figure 2: Actual and forecast sea ice area



The mean difference between the actual and forecast value is 0.378 million square kilometers or 2.5% of the mean value of the maximum sea ice area in the Northern Hemisphere over the period from 2007 to 2012.

4. CONCLUSIONS

As we can see, applying the developed system to forecast climatic factors gives meaningful results. The forecasting error for the sea ice area in the Northern hemisphere is sufficiently small. Therefore, it is possible to use the forecasts to analyze future climatic changes.

The confirmation of the fact that solar radiation impacts sea ice area changes enables us to use long-term calculations of incoming solar radiation (Fedorov, 2014) to predict the trends in sea ice and climatic changes.

REFERENCE LIST

1. Budyiko M. I. (1974) *Izmenenie klimata [Climate change]* L.: Gidrometeoizdat.
2. Bukharov O. E., Bogolyubov D. P., Mizikin A. A. *Razrabotka obolochki sistemyi podderzhki prinyatiya resheniy s ispolzovaniem evolyutsionnyih algoritmov [Development of a decision support system shell based on evolutionary algorithms]* Promyishlennyie ASU i Kontrolleryi [Industrial ACS and controllers] (pp. 37-45) M.: Nauchtehlitizdat, №7.
3. Fedorov V. M. (2014) *Effekt usileniya mezhshirotnogo teploobmena i izmenenie ploschadi morskih ldov v Severnom polusharii [The effect of interwidth heat exchange strengthening and the sea ice area change in the Northern Hemisphere]* Sbornik tezisov vserossiyskoy konferentsii s mezhdunarodnyim uchastiem «Sostoyanie arkticheskih morey i territoriy v usloviyah izmeneniya klimata» [Theses of the Russian conference with the international participation "The state of the Arctic seas and territories in the conditions of climate change"] (pp. 35-36) Arhangelsk, ID SAFU.
4. Fedorov V. M. (2014) Prostranstvennyie i vremennyie variatsii solyarnogo klimata Zemli v sovremennuyu epohu [Current spatial and temporal variations of the Earth's solar climate] *Geofizicheskie protsessyi i biosfera [Geophysical processes and biosphere]13*(4), 67-84.
5. Gorban A. N. (1998) Obobschennaya approksimatsionnaya teorema i vyichislitelnyie vozmozhnosti neyronnyih setey [Generalized approximation theorem and neural network calculation capacity] *Sibirskiy zhurnal vyichislitelnoy matematiki [Siberian Numerical analysis journal] 1*(1),12-24 publishing house: SO RAN.
6. Haykin, S. (2006) *Neyronnye seti: polnyi kurs*, 2-e izd., ispr. [Neural networks: a comprehensive foundation, 2nd ed., Rev.] 388–389. M. Viliams.
7. Ishibuchi, H., Tanaka, H. (1991) A*n extension of the BP-algorithm to interval input vectors* Proc. IEEE Int. Joint Conf. on Neural Networks, Singapore. 1991, 2, 1588–1593.
8. Sea Ice Index. (2013). Retrieved from: http://nsidc.org/data/docs/noaa/g02135_seaice_index/
9. Tsaregorodtsev, V. G. (2010) *Preimuschestva i dostoinstva neyronnykh setey [The benefits and advantages of neural networks]*. Retrieved from: http://www.neuropro.ru/neu3.shtml (accessed 10 March 2014).
10. Voyeykov A. I. (1948) *Klimatyi zemnogo shara, v osobennosti Rossii [Climate of the Earth, Russia in particular]* 1, (pp 163-671) Academy of Sciences of the USSR.