APPLICATION OF THE CELLULAR AUTOMATA METHOD IN TIME SERIES FORECASTING

Antonova I.V., Chikina N.O.
National Technical University «Kharkiv Polytechnic Institute», Kharkiv

This study considers time series (TS) characterizing the incidence of some skin infections in Ukraine. Classical Statistical Analysis based on trends and autocorrelation confirmed the conclusion that results obtained in this case indicate a weak predictability of the considered TS. As an alternative to the classical methods, the method of Cellular Automata (CA) was used in this study. It is known that in order to apply the CA method, it is necessary to confirm the presence of long-term memory in the studied TS and to estimate its depth. In addition, the behavior of the series should be chaotic.

For this purpose, the authors carried out the Fractal Analysis where the value of the Hurst exponent H was determined, confirming the chaotic nature of the studied TS. Besides, during the analysis of H - and R/S - trajectories of the studied series the intervals with long-term memory were found. In this way the series has a long-term memory, which is not constant throughout the whole observation period T, but there is a possibility of numerical estimation of memory depth at the intervals $T_i \left(i = \overline{1,s} \right)$ of the observation period. Such estimation was represented in the form of fuzzy set $\left\{ \left(g_i, \mu\left(g_i \right) \right) \right\}, \left(i = \overline{1,s} \right)$, where g_i is the numerical estimation of memory depth, $\mu\left(g_i \right)$ is the corresponding to g_i value of the membership function.

The received information made it possible to build a CA model to obtain a predictive value. The CA works with memory and is a system of cells, where the each one stores information. The character of information change at each time moment is programmed with the help of some genetic algorithm, which in its turn determines the interaction of cells with all nearest neighbors. Application of the genetic algorithm allows to obtain the best configuration of a cellular automaton, and, consequently, the best predictive model.

To apply the CA method, a memory basis was created, while the original TS was transformed into a linguistic TS using the method of trend corridors. A linguistic variable (term) was assigned to each value of the TS level. The set of terms forms a term-set of 3 elements $T = \left\{ L, A, H \right\}$, where L, A, H mean low, acceptable and high morbidity, respectively. Since the value of the term is already known, finding the predicted value was reduced to calculating the values of the membership function $\mu\left(L \right), \mu\left(A \right), \mu\left(H \right)$. The forecast is also represented as a fuzzy linguistic set.