IDENTIFYING “REGIMES” IN THE TIME SEQUENCES OF PROCESSES

*Boris Zalmanovich Belashev[1]

1. Karelian research center of Russian academy of science

A process is represented as a sequence of “regimes” when describing climatic variations and complex systems, detecting the malfunctioning of instruments and illegal interference in computer and hardware networks. For natural, technological and social processes such representation is comfortable to detect precursors of disastrous events.
The shifts of “regimes” in the sequences are commonly detected by checking statistical hypotheses by Student’s test and tests of non-parametric statistics when comparing the characteristics of the sequence at one or another site. Algorithms often use a priori information, work in unstable fashion over short intervals and at sequences boundaries, find it hard to recognize the shifts and trends of “regimes”. Widespread R-algorithm [1] analyze data consecutively in a sliding window, operate with a small number of parameters and filter off “red” noise. The boundary regimes it detect may reflect distortions characteristic of the beginning and end of measurements.
The goal of the present study is to develop B-algorithm which automatically recognizes stable “regimes” in time data sequences.
The main stages of the algorithm are clusterization, comparison and optimization. Clusterization preliminary divides a normalized time sequence into “regimes”. While comparing the characteristics of the “regimes”, the division obtained is re-estimated. The final model of division is formed during optimization, which minimizes the variance of deviations from the trends of the “regimes” by varying the internal boundaries of the sites.
The algorithm, developed as a software in the MATLAB computer mathematics system, fulfils procedures automatically, decreases the amount of a priori data, increases operation speed and brings solutions closer to optima.
The operation of B- and R-algorithms is exemplified by the PDO (Pacific Decade Oscillation) sequence of mean January temperatures [2].
In both algorithms the number of “regimes” and their approximate positions coincide. The delay of part of R-“regimes” emphasize the sensitivity of B-algorithm to new “regimes”. Discrepancies are connected with the last “regime” identified by R- and B-algorithms on 2 and 15 members of sequence, respectively.
B- algorithm can operate with constant, linear and complex trends.
B- and R-algorithms, applied to the time sequences of river runoff, solar and seismic activity and the sequences of the annual rings of trees, have provided new information on the stages of processes.
The simplified schematic “regime” representation of a process reveals trends commonly obscured by the abundance of details. B- algorithm identifies stable “regimes” and recognizes the shifts and trends of “regimes”. The use of MATLAB functions increases the potential of the algorithm, admits the fulfilment of procedures by various methods, brings solutions closer to optimum solutions and ensures high operation speed. The ability of the algorithm to automatically process data makes it suitable for working with extensive data in various earth sciences.

References:

1.Rodionov, S. (2006) Use of prewhitening in climate region shift detection. Geophysical Research Letters 33, L12707.
2.Pacific Decadal Oscillation Index Data.
https://www.rdocumentation.org/packages/rpdo/versions/0.2.5.

Keywords: time sequence, shift regime, algorithm, clusterization, comparision, optimization