# A BRIEF OVERVIEW OF THE REGIME SHIFT DETECTION METHODS

*Sergei Rodionov*

*Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA 98195, U.S.A.*

The methods reviewed here are primarily those that are used in atmospheric and oceanic (physical and biological) studies. These methods are divided into four groups, depending on the type of shifts they are designed to detect, and placed in Tables 1-4. The tables are self-explanatory and provide a brief description of each method, along with the basic references, as well as their strong and weak points.

Shifts in the mean are the most common type of shifts considered in the literature, and as a result, Table 1 is the most populated. The definition of climatic regime shifts, for example, is often based on "differing average climatic levels over a multi-annual duration" (Overland et al., 2005). It is not surprising, therefore, that a major part of the research effort is directed at developing the methods for detecting shifts in the mean. Several approaches can be distinguished:

1. Parametric methods, such as the classical t-test. The methods require an assumption about the probability distribution of the data;
2. Non-parametric methods, such as the Mann–Whitney U-test, Wilcoxon rank sum, or Mann-Kendall test. No assumption about the probability distribution is required;
3. Curve-fitting methods;
4. Bayesian analysis and its variations, such as the Markov chain Monte Carlo method;
5. Regression-based methods;
6. Cumulative sum (CUSUM) methods; and
7. Sequential methods.

There are much fewer methods capable of detecting regime shifts in the second-order statistics, such as the variance and power spectrum (Tables 2 and 3). A separate group of methods includes those multivariate methods that are designed to detect shifts in the entire structure of a complex system (Table 4). This is a fast-growing group of methods, and Table 4 will likely require an update soon.

***Table 1.*** *Shifts in the mean*

| Method | Brief Description | Pros | Cons |
|---|---|---|---|
| Student t-test | The most commonly used techniques for testing a hypothesis on the basis of a difference between sample means. It determines a probability that two populations are the same with respect to the variable tested. Can be applied sequentially for each data point. The position of a change-point corresponds to the location of the greatest *t* value exceeding the given threshold (Ducre-Robitaille et al., 2003). | Strong theoretical basis. Robust to the assumption of normality and equality of variances. | Data-dredging, a problem that arises in testing for change occurring at a specified time (Epstein, 1982). A problem of finding local maxima in the case of a sequential version. |
| Bayesian analysis | Methods based on this approach differ mainly by the prior distributions specified to represent the other unknown parameters, such as the mean before and after the shift, and the variance of the observations. Some recent methods are presented by Perreault et al. (2000) and Chu and Zhao (2004). | Strong theoretical basis. Provides uncertainty estimates of change points and means for predicting. | Requires a mathematical model of the data. A single change-point scenario. |
| Mann–Whitney U-test | The test is based on the rank values of the sequence, but can be used in a combination with the windowing technique (Mauget, 2003). | Strong theoretical basis. Easy to use. | The data need to be detrended. Originally designed for detecting a single change-point. The effect of using a variable window is not clear. |
| Wilcoxon rank sum | A non-parametric test often used for the homogenization of temperature and precipitation series (Karl and Williams, 1987). Ducre-Robitaille et al. (2003) modified the test in order to identify the most probably position of a change-point, since it was originally based on break positions that were inferred from the metadata. | Strong theoretical basis. Easy to use. | The data need to be detrended. Originally designed for detecting a single change-point. Performance of the modified version has not been evaluated. |
| Pettitt test | A non-parametric test based on the Wilcoxon test (Pettitt, 1979). It can also be derived from the Mann–Whitney U-test. | Strong theoretical basis. Easy to use. | A single change-point scenario. The data need to be detrended. |

| Method | Brief Description | Pros | Cons |
|---|---|---|---|
| Mann-Kendall test | A non-parametric test that falls under the class of rank tests. It is shown to be useful in analysis of abrupt climate changes (Goossens and Berger, 1987). Gerstengarbe (1999) presented the shifted sequential version of the test. | Strong theoretical basis. The classical version is easy to use. | The data should not be affected by a trend. A single change-point scenario for the classical version. Not automatic. |
| Lepage test | A non-parametric test that investigates significant differences between two samples, even if the distributions of the parent populations are unknown. The Lepage statistic is a sum of the squares of the standardized Wilcoxon's and Ansari–Bradley's statistics. Modified by Yonetani (1993). | The test appears to be more statistically powerful than other similar non-parametric tests. | A single change-point scenario. Not automatic; requires a visual inspection of the Lepage statistic time series if calculated using the windowing technique. |
| Standard normal homoge-neity test | A test to detect discontinuity in a standardized time series often called the standard normal homogeneity test (Alexandersson, 1986). The SNHT also has been discussed in a slightly different version in the statistical literature (Hawkins, 1977). | Performed well in comparison with other similar tests for revealing and dating single and sudden shifts in artificial data (Easterling and Peterson, 1995). | Performance deteriorates when change-points are close together in time or the number of change-points are greater than four (Easterling and Peterson, 1995). |
| Regression-based approach | The two-phase regression technique, in which the tested time series is the predictand and time is the predictor. It was introduced by Solow (1987) and later modified by Easterling and Peterson (1995), Elsner et al. (2000), and Lund and Reeves (2002). | Robust, outperforms the SNHT method in the case of multiple change-points. | Does not work if change-points are separated by less than ten points. Less sensitive to small shifts. Often biased toward an excessive number of unobserved change-points. |
| Cumulative deviation test | The test is based on the adjusted partial sums or cumulative deviations from the mean (Buishand, 1982). A variant of a simple cumulative sum (CUSUM) method (Rebstock, 2002). | Simple, easy to use. | Works with anomalies. Using different base periods may substantially affect the results. |

| Method | Brief Description | Pros | Cons |
|---|---|---|---|
| Oerlemans method | The method is based on a comparison of a typical change, prescribed a priori, with an interval of a given time series (Oerlemans, 1978). A break is defined as the ratio of the amplitude of the change to the corresponding rms difference between the typical and observed change. | Can be applied to any time series. Easy to compare the results. | Requires the “best-fitting” of a curve that represents an idealized break in the data. A statistical significance test cannot be constructed. |
| Signal-to-noise ratio | A regime shift (“climatic jump”) is defined when the signal-to-noise ration is greater than 1, which is equivalent to the 95% significance level (Yamamoto et al., 1986). | Simple, easy to use. | A single change-point. |
| Intervention analysis | The method is an extension of Auto-Regressive Integrated Moving Average (ARIMA) modeling (Box and Tiao, 1975). Wei (1990) provided a detailed review of the method, and Francis and Hare (1994) used it for analysis of Alaska salmon catch records. | Allows for a quantitative estimate of the statistical significance of step interventions, while accounting for autocorrelation in the time series. | Typically, the time and type of intervention should be specified in advance. |
| Markov chain Monte Carlo | A Bayesian approach using Gibbs sampling (Elsner et al., 2004). The MCMC algorithm consists of two steps. Step 1 uses the entire record to determine candidate change points, and step 2 determines the posterior distributions of the data before and after the candidate change point, ignoring the other candidates. | Strong theoretical basis. | Relatively complex. Requires data modeling and a visual inspection of posterior probabilities. |
| Lanzante method | An iterative procedure designed to search for multiple change-points. It involves the application of a non-parametric test, related to the Wilcoxon-Mann-Whitney test, followed by an adjustment step (Lanzante, 1996). | Outperforms the regression-based method of Easterling and Peterson (1995), if the change-points are not too close. Designed to distinguish a shift from a trend. | Performance deteriorates if the change-points are close to each other (separated by less than 25 points). |

| Method | Brief Description | Pros | Cons |
|---|---|---|---|
| STARS | A sequential version of the partial CUSUM method combined with the t-test Rodionov, 2004. | Automatic detection of multiple change-points. Signals a possibility of a regime shift in real time. Outperforms the Lanzante method if the shifts occur on a background of a trend. | Requires some experimentation when choosing the probability level and cutoff length. Does not explicitly take into account the autocorrelation. |

***Table 2.*** *Shifts in the Variance*

| Method | Description | Pros | Cons |
|---|---|---|---|
| Downton-Katz test | The method is somewhat analogous to that developed by Karl and Williams (1987) to test for homogeneities in the mean. It uses a non-parametric bootstrap technique to compute confidence intervals. | No assumptions are required about the frequency distribution. | Requires a reference time series with no potential change-points. The change-points have to be widely separated (at least 10 years apart). |
| Rodionov method | Similar to STARS, but based on the F-test. It is included in the regime shift detection calculator (See the method's description in this volume). | Automatic detection of multiple change-points. Signals a possibility of a regime shift in real time. | The method is not documented yet. It is still in the experimental phase. |

***Table 3***. *Shifts in the Frequency Structure*

| Method | Description | Pros | Cons |
|---|---|---|---|
| Nikiforov method | The method is based on ARIMA modeling of time series before and after the shift combined with a likelihood ratio test (Nikiforov, 1983; Basseville and Nikiforov, 1993). An application of the method can be found in Rodionov (1994). | Strong theoretical basis. | A single change-point scenario. |

***Table 4.*** *Shifts in the System*

| Method | Description | Pros | Cons |
|---|---|---|---|
| Principal component analysis | The method is widely used to identify coherent patterns of variability among large sets of time series Von Storch and Zwiers, 1999; Mantua, 2004). Although not a regime shift detection method per se, it has been applied to 100 biotic and abiotic time series in the North Pacific to analyze the scale of regime shifts in 1977 and 1989 Hare and Mantua, 2000. | Reduces the dimensionality of the data matrix. Requires no a priori assumption about candidate regime shift years. | Additional time series analysis methods must be used to assess the statistical significance and character of temporal changes in the PCs. |
| Average standard deviates | An ad hoc composing method that creates a single “regime index” consisting of average standard deviates Ebbesmeyer et al., 1991; Hare and Mantua, 2000; Mantua, 2004. | Easy to use. | Requires an a priori specification of a regime shift date and a sign reversal of some time series, which leads to a spurious amplification of the shift Rudnick and Davis, 2003. |
| Fisher information | An information theory approach using Fisher information as an indicator for tracking steady and transient ecosystem states Fath et al., 2003. The information is characterized by a ratio of a system’s acceleration to its speed along the state space trajectory. | Simple and easily applied to collections of time series with many variables. | Requires a careful choice of input variables and their weighting. Interpretation of the results is not straightforward. No means for assessing the statistical significance. |
| Vector auto-regressive method | A formal statistical approach to detecting regime shifts in a multivariate system Solow and Beet, 2005. The regime shift is identified objectively as the point at which the system changes from one steady state to another. The states are described by a vector autoregressive model of the first order. | Relies on standard statistical theory to provide a significance test. Allows for serial dependence in the vector of time series. | Requires a large number of observations to fit the model. The results are sensitive to the selection of variables. Computationally demanding. |

**References**

Alexandersson, H. (1986). A homogeneity test applied to precipitation data, *J. Climatol.,* 6, 661-675.

Basseville, M. and I. V. Nikiforov. (1993). *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, Englewood Cliffs, NJ, 528 pp.

Box, G. E. P. and G. C. Tiao. (1975). Intervention analysis with applications to economic and environmental problems, *J. Am. Statist. Assoc.,* 70, 70-79.

Buishand, T. A. (1982). Some methods for testing the homogenety of rainfall records, *J. Hydrol.,* 58, 11-27.

Chu, P. S. and X. Zhao. (2004). Bayesian change-point analysis of tropical cyclone activity: The central North Pacific case, *J. Climate*, 17, 4893-4901.

Ducre-Robitaille, J. F., L. A. Vincent, and G. Boulet. (2003). Comparison of techniques for detection of discontinuities in temperature series, *Int. J. Climatol.,* 23, 1087-1101.

Easterling, D. R. and T. C. Peterson. (1995). A new method for detecting undocumented discontinuities in climatological time series, *Int. J. Climatol.,* 15, 369-377.

Ebbesmeyer, C. C., Cayan, D. R., Mclain, D. R., Nichols, F. N., Peterson, D. H., and Redmond, K. T. (1991). 1976 step in Pacific climate: Forty environmental changes between 1968-1975 and 1977-1984. pp. 115-126. In: J.L. Betancourt and V.L. Tharp (Eds.) Proceedings of the 7th Annual Pacific Climate (PACLIM) Workshop, April 1990. California Department of Water Resources. Interagency Ecological Study Program Technical Report 26,126 pp.

Elsner, J. B., T. Jagger, and X. F. Niu. (2000): Changes in the rates of North Atlantic major hurricane activity during the 20th century, *Geophys. Res. Lett.,* 27, 1743-1746.

Elsner, J. B., X. Niu, and T. Jagger. (2004). Detecting Shifts in Hurricane Rates Using a Markov Chain Monte Carlo Approach, *J. Climate,* 4, 2652-2666.

Epstein, E. S. (1982). Detecting climate change, *J. Appl. Meteorol.,* 21, 1172.

Fath, B. D., H. Cabezas, and C. W. Pawlowski. (2003). Regime changes in ecological systems: an information theory approach, *J Theor. Biol.,* 222, 517-530.

Francis, R. C. and S. R. Hare. (1994). Decadal-scale regime shifts in the large marine ecosystems of the Northeast Pacific: a case for historical science, *Fish. Oceanogr.,* 3, 279-291.

Gerstengarbe, F. W. W. (1999). Estimation of the beginning and end of recurrent events within a climate regime, *Climate Research,* 11, 97-107.

Goossens, C. and A. Berger. (1987). How to recognize an abrupt climatic change? *Abrupt Climatic Change: Evidence and Implications*, W. H. Berger and L. D. Labeyrie, Eds., Kluwer Academic Publishers, Dordrecht, 31-46.

Hare, S. R. and N. J. Mantua. (2000). Empirical evidence for North Pacific regime shifts in 1977 and 1989, *Progr. Oceanog.,* 47, 103-146.

Hawkins, P. M. (1977). Testing a sequence of observations for a shift in random location, *J. Am. Statist. Assoc.,* 73, 180-185.

Karl, T. R. and C. W. Jr. Williams. (1987). An approach to adjusting climatological time series for discontinuous inhomogeneities, *J. Clim. Appl. Meteorol.,* 26, 1744-1763.

Lanzante, J. R. (1996). Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data, *Int. J. Climatol.,* 16, 1197-1226.

Lund, R. and J. Reeves. (2002). Detection of Undocumented Changepoints: A Revision of the Two-Phase Regression Model, *J. Climate,* 15, 2547-2554.

Mantua, N. J. (2004). Methods for detecting regime shifts in large marine ecosystems: a review with approaches applied to North Pacific data, *Progr. Oceanog.,* 60, 165-182.

Mauget, S. A. (2003). Multidecadal regime shifts in US streamflow, precipitation, and temperature at the end of the twentieth century, *J. Climate,* 16, 3905-3916.

Nikiforov, I. V. (1983). *Sequential Detection of Abrupt Changes in Time Series Properties*, Nauka, Moscow (in Russian),

Oerlemans, J. (1978). An objective approach to breaks in the weather, *Mon. Wea. Rev.,* 106, 1672-1679.

Overland, J. E., D. B. Percival and H. O. Mofjeld. (2005). Regime shifts and red noise in the North Pacific (Submitted).

Perreault, L., J. Bernier, and E. Parent. (2000). Bayesian changepoint analysis in hydrometeorological time series. Part 1. The normal model revisited, *J. Hydrol.,* 235, 221-241.

Pettitt, A. N. (1979). A nonparametric approach to the change-point problem, *Appl. Stat.,* 28, 126-135.

Rebstock, G. A. (2002). Climatic regime shifts and decadal-scale variability in calanoid copepod populations off southern California, *Global Change Biology,* 8, 71-89.

Rodionov, S. (2004). A sequential algorithm for testing climate regime shifts, *Geophys. Res. Lett.,* 31, L09204, doi:10.1029/2004GL019448.

Rodionov, S. N. (1994). *Global and Regional Climate Interactions: The Caspian Sea Experience*, Kluwer Academic Pub., Dordrecht, The Netherlands,

Rudnick, D. I. and R. E. Davis. (2003). Red noise and regime shifts, *Deep-Sea Research,* 50, 691-699.

Solow, A. R. (1987). Testing for climate change: an application of the two-phase regression model, *J. Clim. Appl. Meteorol.,* 26, 1401.

Solow, A. R. and A. R. Beet. (2005). A test for a regime shift, *Fish. Oceanogr.,* 14, 236-240.

Von Storch, H. and F. W. Zwiers. (1999). *Statistical analysis in climate research*, Cambridge University Press, 494 pp.

Wei, W. W. S. (1990). *Time Series Analysis: Univariate and Multivariates Methods*, Addison-Wesley, Redwood City, CA, 478 pp.

Yamamoto, R., T. Iwashima, and N. K. Sange. (1986). An analysis of climate jump, *J. Meteorol. Soc. Jap.,* 64, 273-281.

Yonetani, T. (1993). Detection of long term trend, cyclic variation and step like change by the Lepage test, *J. Meteorol. Soc. Jap.,* 71, 415-418.