

---

УДК 621.398  
ББК 32.965.7  
Б 94

**Бучацкая В.В.**

*Кандидат технических наук, доцент кафедры прикладной математики и информационных технологий факультета математики и компьютерных наук Адыгейского государственного университета, Майкоп, тел. (8772) 59-39-04, e-mail: buch\_vic@mail.ru*

**Обработка аномальных значений уровней временного ряда как этап комплексной оценки информации в подсистеме прогнозирования для ситуационного центра**  
(Рецензирована)

**Аннотация**

*Рассматриваются методы предварительной обработки временных рядов, учитывающие наличие случайных составляющих типа аномальных значений. Описывается алгоритм обнаружения и устранения таких значений, структура программного приложения с использованием указанного алгоритма и результаты численных экспериментов.*

**Ключевые слова:** аномальные значения уровней временного ряда, критерий Ирвина, предобработка данных.

**Buchatskaya V.V.**

*Candidate of Technical Sciences, Associate Professor of the Department of Applied Mathematics and Information Technology, Faculty of Mathematics and Computer Science, Adyghe State University, Maikop, ph. (8772) 59-39-04, e-mail: buch\_vic@mail.ru*

**Processing abnormal values of temporary row levels as a stage of information complex assessment in a subsystem of forecasting for the situational center**

**Abstract**

*The paper discusses the methods of preliminary processing the temporary rows, considering existence of casual components of the abnormal value type. The author describes the algorithm of detection and elimination of such values, structure of a program application with the use of the specified algorithm and results of numerical experiments.*

**Keywords:** abnormal values of temporary row levels, Earvin's criterion, data pre-processing.

В современных условиях, характеризующихся неопределенностью и постоянными изменениями, актуально использование инновационных методов управления. Одним из таких методов является ситуационный центр. Он аккумулирует средства сбора и анализа информации, инструменты прогнозирования и построения возможных моделей развития и визуального представления результатов [1].

В рамках ситуационного центра могут решаться задачи, которые можно условно разделить на три группы: информационно-аналитическое обеспечение, прогнозирование развития ситуаций и обеспечение коллективной подготовки решений. Важное место среди них занимает задача прогнозирования. Ее решение позволяет получить сценарий развития на основе анализа текущей ситуации.

Прогнозирующие системы определяются как совокупность методов, приемов и процедур, позволяющих прогнозы при заданной целевой функции развития объекта прогнозирования, при заданном объеме прогнозной информации.

Любая исследуемая совокупность, наряду со значениями признаков, сложившихся под влиянием факторов, непосредственно характерных для анализируемой совокупности, может содержать и значения признаков, полученных под воздействием иных фак-

---

торов, не характерных для основной совокупности. Такие значения резко выделяются и, следовательно, использование методологии статистического моделирования без предварительного анализа и изучения аномальных наблюдений приводит к серьезным ошибкам при анализе. Резко выделяющиеся из общей совокупности (аномальные) наблюдения требуют отдельного изучения и обработки. Таким образом, адекватная подготовка исходных данных является одним из основных этапов проведения статистического исследования на информационной базе, характеризующей реальные социально-экономические явления.

Выбор того или иного метода выявления и анализа аномальных наблюдений определяется объемом совокупности, характером исследуемых процессов и задач (одномерные и многомерные).

Под аномальным уровнем понимается отдельное значение уровня временного ряда, которое не отвечает потенциальным возможностям исследуемой экономической системы и которое, оставаясь в качестве уровня ряда, оказывает существенное влияние на значения основных характеристик временного ряда, в том числе на соответствующую трендовую модель.

Нехарактерные уровни во временном ряду можно подразделить на три группы [2, с. 22-23]:

- значения, отражающие объективное развитие процесса, но сильно отличающиеся от общей тенденции, так как они проявляют свои экстремальные воздействия крайне редко;
- значения, возникающие вследствие изменений методики расчета;
- значения, возникающие вследствие ошибок при измерении показателя, при записи и передаче информации, а также значения, связанные с различными катастрофическими явлениями, не влияющими на дальнейший ход развития явления, агрегировании и дезагрегировании показателей и т.д.

Аномальные значения первой группы не всегда должны исключаться из временного ряда и могут даже оказаться полезными на этапе исследования причинно-следственного механизма развития явления. Наличие нехарактерных пиковых значений для одного и того же момента времени в различных временных рядах свидетельствует, как правило, о причинных связях между соответствующими показателями.

Нехарактерные значения второй группы не должны исключаться из рассмотрения, а приниматься за «повторные» (пороговые), начиная с которых должны быть пересчитаны по новой методике все предыдущие значения временного ряда.

Аномальные значения третьей группы должны быть исключены из рассмотрения в любом случае, так как они искажают представление о характере развития явления и могут оказать существенное влияние на выводы, полученные в результате анализа ряда, содержащего такую искаженную информацию.

Для выявления аномальных уровней временных рядов используются методы, рассчитанные для статистических совокупностей. Одним из них является метод Ирвина [2, с. 18; 3]. Он предполагает использование следующего алгоритма.

Шаг 1. Рассчитывается значение величины

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\sigma_y}, \quad t = 2, 3, \dots, n,$$

где среднеквадратичное отклонение  $\sigma_y$  рассчитывается по формуле:

$$\sigma_y = \sqrt{\frac{\sum_{t=1}^n (y_t - \bar{y})^2}{n-1}}.$$

$$\bar{y} = \frac{\sum_{t=1}^n Y_t}{n}.$$

Шаг 2. Расчетные значения  $\lambda_t$  сравниваются с табличными значениями критерия Ирвина  $\lambda_{\alpha}$  и если оказываются больше табличных, то соответствующее значение  $Y_t$  уровня ряда считается аномальным.

После выявления аномальных уровней ряда обязательно определение причин их возникновения, а затем устранение этих значений [3].

Указанные действия должны быть реализованы в рамках подсистемы прогнозирования для обеспечения качества получаемых прогнозов. В связи с этим был разработан программный модуль, выполняющий указанные действия. Кроме реализации указанного алгоритма обнаружения аномальных значений в нем реализовано прогнозирование временного ряда с использованием трендовых моделей (линейной, полином 2-й степени, полином 3-й степени) и оценка их достоверности [4, 5]. Сглаживание данных реализовано методом взвешенной скользящей средней. Алгоритм работы программы представлен на рисунке 1.

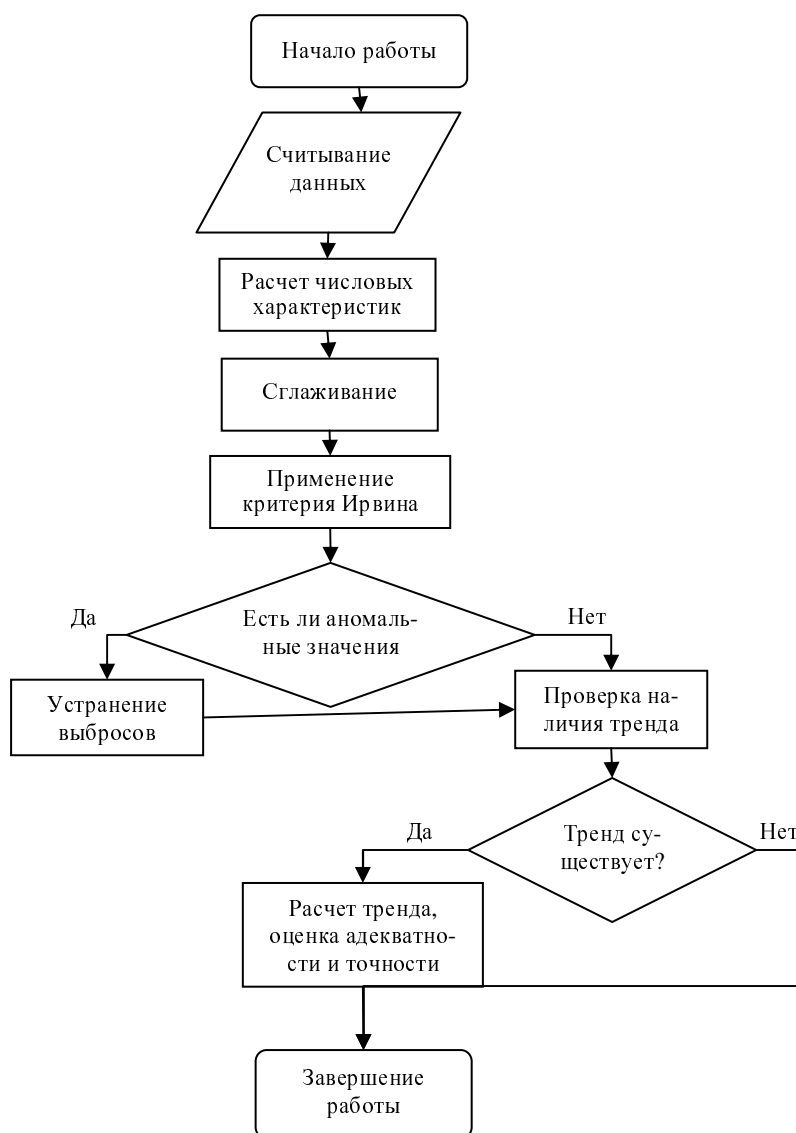


Рис. 1. Блок-схема работы программного модуля

Исходные данные для работы программного модуля содержатся в файле формата .csv, который предусматривает структуру строк следующего вида:

```
DATA,D|NOMINAL,N,38,5|CURS,N,38,5|NUM_CODE,N,11,0|CHAR_CODE,C,3
```

Эта структура выбрана в связи с тем, что она является простой и удобной для интерпретации, кроме того, многие статистические базы предоставляют информацию в виде легко адаптируемом к указанному формату.

В программе предусмотрена возможность графического представления данных и результатов для исходного ряда, сглаженного ряда, проверки критерия Ирвина, прогноза по полиномиальным кривым роста. Кроме того, предусмотрена возможность формирования и вывода отчетов по построенным моделям с указанием их параметров и оценки достоверности.

Работа программного модуля проверена на примерах данных о динамике изменения курсов валют, уровня прожиточного минимума и других социально-экономических показателей. Результат работы программы на примере данных о динамике белорусского рубля [6] представлен на рисунках 2 и 3.

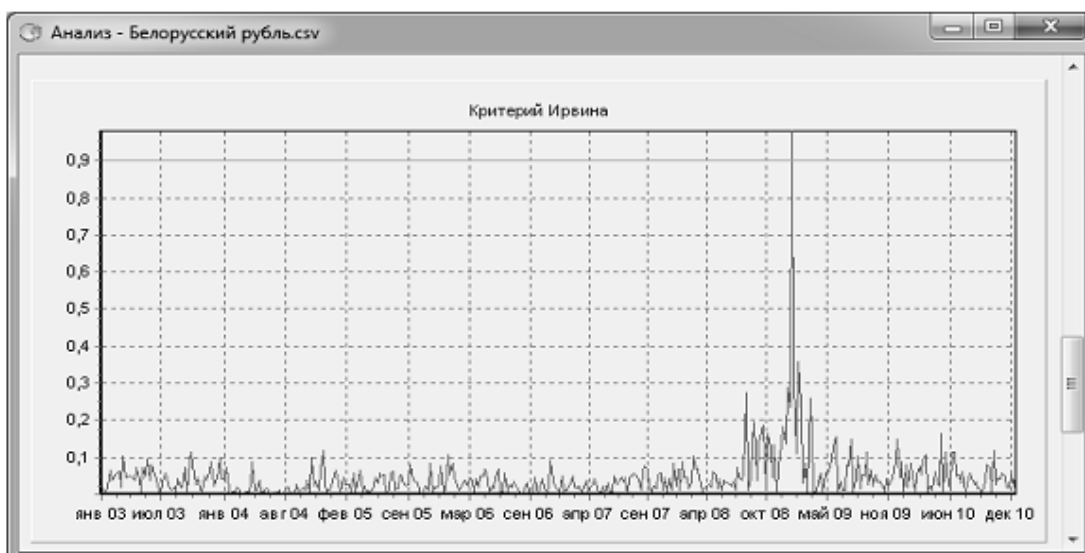


Рис. 2. Результат проверки критерия Ирвина

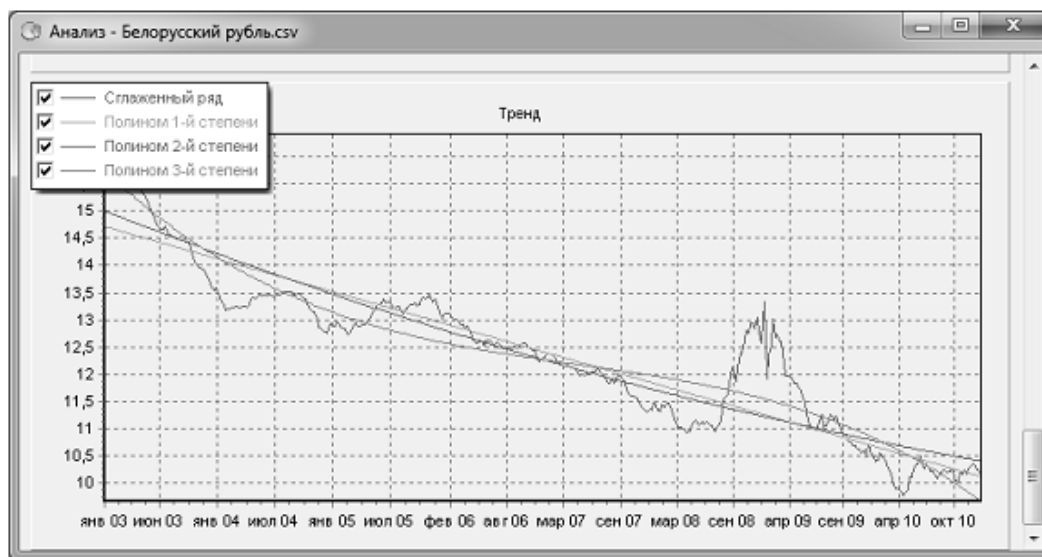


Рис. 3. Результат построения прогнозных моделей

---

Таким образом, созданное программное приложение реализует хороший функционал для базового статистического анализа временных рядов, в то же время обладая простым, интуитивно понятным интерфейсом. Отметим, что программный модуль универсален, не привязан к определенной предметной области исследования. Это позволит включать его в состав подсистемы прогнозирования для ситуационного центра любого уровня.

#### Примечания:

1. Симанков В.С., Колесников Д.А. Режимы работы ситуационного центра // Вестник южного научного центра РАН. Ростов н/Д, 2010. Т. 6, № 1. С. 96-99.
2. Садовникова Н.А., Шмойлова Р.А. Анализ временных рядов и прогнозирование: учеб. пособие. Вып. 2 / Моск. гос. ун-т экономики, статистики и информатики. М., 2004. 184 с.
3. Христиановский В.В., Щербина В.П. Анализ временных рядов в экономике: практика применения: учеб. пособие. Донецк: ДонНУ, 2011. С. 11-13.
4. Бучацкая В.В. Методика определения интервальных оценок при прогнозировании методами экстраполяции // Вестник Адыгейского государственного университета. Сер. Естественно-математические и технические науки. 2012. Вып. 3. С. 136-140. URL: <http://vestnik.adygnet.ru>
5. Бучацкая В.В. Сравнительный анализ некоторых моделей прогнозирования временных рядов // Актуальные проблемы прикладной математики, информатики и механики: материалы междунар. конф. Воронеж, 2012. С. 8-10.
6. База данных по курсам валют. Банк России. URL: [http://www.cbr.ru/currency\\_base/dynamics.aspx](http://www.cbr.ru/currency_base/dynamics.aspx) (дата обращения 25.09.2013 г.).

#### References:

1. Simankov V.S., Kolesnikov D.A. Working conditions of a command center // Bulletin of the RAS southern scientific center. Rostov-on-Don, 2010. Vol. 6, No. 1. P. 96-99.
2. Sadovnikova N.A., Shmoylova R.A. Analysis of time series and prognostication: a manual. Iss. 2 / Mosc. state un-ty of economics, statistics and informatics. M., 2004. 184 pp.
3. Khristianovskiy V.V., Shcherbina V.P. The analysis of time series in economics: application practice: a manual. Donetsk: DonNu, 2011. P. 11-13.
4. Buchatskaya V.V. Technique to define the interval estimates when forecasting by extrapolation methods // The Bulletin of the Adyghe State University. Series Natural-Mathematical and Technical Sciences. 2012. Iss. 3. P. 136-140. URL: <http://vestnik.adygnet.ru>
5. Buchatskaya V.V. Comparative analysis of some models of time series prediction // Actual problems of applied mathematics, informatics and mechanics: materials of international conf. Voronezh, 2012. P. 8-10.
6. The database of currency. The bank of Russia. URL: [http://www.cbr.ru/currency\\_base/dynamics.aspx](http://www.cbr.ru/currency_base/dynamics.aspx) (date of address of 25.09.2013).