УДК 504”550.1”

# ВЫЯВЛЕНИЕ ОСНОВНЫХ СЕЗОННЫХ ЗАКОНОМЕРНОСТЕЙ ВРЕМЕННОГО ИЗМЕНЕНИЯ ЭКОЛОГИЧЕСКИХ ПОКАЗАТЕЛЕЙ НА ПРИМЕРЕ ТОМСКОЙ ОБЛАСТИ

**Черкашина Ю.А.**

*Национальный исследовательский Томский политехнический университет, Томск, e-mail: cherr999y@mail.ru*

Статья является результатом научных исследований сотрудников кафедры Прикладной математики Томского политехнического университета и сотрудников Департамента природных ресурсов и охраны окружающей среды. Исследования посвящены применению математических методов для выявления закономерностей временного изменения экологических показателей на основе статистического анализа. В статье приведен анализ исходных данных, а именно: восстановление пропусков в данных; выявление наличия сезонных ритмов; выделение трендов во временных рядах; определения сезонной декомпозиции. Изложены основные математические методы. Приведены результаты исследования. Произведен обзор и сравнение наиболее популярных методов анализа временных рядов – авторегрессионных и нейросетевых.

**Ключевые слова: моделирование, тренд, сезонные колебания, несанкционированный объект размещения отходов, авторегрессионная модель, нейросетевая модель**

# IDENTIFY THE MAIN SEASONAL PATTERNS OF TEMPORAL CHANGES ENVIRONMENTAL INDICATORS AT THE EXAMPLE OF TOMSK AREA

**Cherkashina Y.A.**

*National Research Tomsk Polytechnic University, Tomsk, e-mail: cherr999y@mail.ru*

The article includes results of scientific results achieved at department of Applied Mathematics at Tomsk Polytechnic University and Department of Natural Resources and Environment. Researches are devoted the application of mathematical methods to identify patterns of change in environmental performance based on statistical analysis. In the article analyzes the initial data, namely: recovery of omissions in data; determination seasonal rhythms' existence; determination of trends in time series; seasonal decomposition. Describes the main mathematical methods, the results of the study. Produced review and compare of the most popular methods for the analysis of time series – autoregressive and neural network.

**Keywords: modeling, trend, seasonal rhythms, unsanctioned waste disposal facility, AR model, neural network model**

Экологическая обстановка в мире требует всестороннего и повседневного анализа состояния окружающей среды. Только на основании точных количественных данных можно принимать решения о проведении необходимых мероприятий, предупреждающих ухудшение состояния окружающей среды.

В настоящее время при решении экологических задач отходят от вербального описания и основываются на математических моделях и информационных технологиях. Одним из наиболее сложных и трудоемких процессов является выявление закономерностей из имеющихся массивов данных. Он не всегда заканчивается успешно, поскольку данные содержат разнотипную, противоречивую и неполную информацию.

Цель работы являлось проведение исследования, направленного на выявление тренда и сезонной декомпозиции в экологических данных.

## Содержательная постановка задачи

Результатом хозяйственной или иной деятельности человека становится воздействие на все компоненты окружающей среды, в том числе и на земельные ресурсы. Негативные изменения могут проявляться в их загрязнении, а конечном счете – в деградации. Одна из основных причин – отсутствие единой региональной системы сбора, вывоза, сортировки, использования, обезвреживания и размещения отходов различного происхождения.

Регионы сталкиваются с проблемами наличия несанкционированных свалок, необходимостью их ликвидации и последующей рекультивации загрязненных земель. На несанкционированных свалках встречаются трупы животных, медицинские отходы, содержащие биологические остатки (использованные шприцы с остатками крови и др.). Согласно ст. 1 Федерального закона от 24.06.1998 № 89-ФЗ «Об отходах производства и потребления» отходы производства и потребления – остатки сырья, материалов, полуфабрикатов, иных изделий или продуктов, которые образовались в процессе производства или потребления, а также товары (продукция), утратившие свои потребительские свойства.

Для исследования степени воздействия на окружающую среду отходы принято классифицировать по классам опасности. Критерии отнесения отходов к классам опасности утверждены Приказом Министерства природных ресурсов Российской

Федерации от 15.06.2001 № 511 «Об утверждении критериев отнесения опасных отходов к классу опасности для окружающей природной среды». Класс опасности отходов устанавливается по степени возможного вредного воздействия на окружающую природную среду (ОПС) при непосредственном или опосредованном воздействии опасного отхода на нее.

На санкционированные объекты размещения отходов запрещается прием, захоронение, обезвреживание химически и эпидемически опасных отходов; твердых, пастообразных и жидких отходов, которые содержат токсичные вещества (1–2 классов опасности); горючих и взрывоопасных ингредиентов; отходов радиоактивных веществ; трупы павших животных и конфискатов боен мясопереработки; отходов лечебных учреждений; сбор вторичного сырья; размещение приемных пунктов вторичного сырья от учреждений и населения.

Также остро стоит проблема захоронения отходов, образующихся в малых населенных пунктах области, т.к. строительство полноценного полигона требует больших финансовых затрат и не по силам бюджетам сельских поселений.

**Обзор методов и моделей анализа временных рядов**

В основу исследования взяты данные 1998–2013 гг. по обращению с отходами производства и потребления на территории Томской области. Исходные данные содержат такие характеристика, как общее количество: образованных отходов; отходов, поступивших от сторонних предприятий; переданных в переработку или во вторичное использование; отходов, которое осталось на следующий год на свалке; обезвреженных в текущем году; отходов, вывозимое на полигон.

Временной ряд включает в себя два обязательных элемента – отметку времени и значение показателя ряда [2].

В настоящее время существует множество моделей прогнозирования временных рядов: регрессионные и авторегрессионные модели, нейросетевые модели, модели экспоненциального сглаживания, модели на базе цепей Маркова, классификационные модели и др. Наиболее популярными и широко используемыми являются классы авторегрессионных и нейросетевых моделей [4]:

1. Авторегрессионные модели. В основу таких моделей заложено предположение о том, что значение процесса Z(t) линейно зависит от некоторого количества предыдущих значений того же процесса Z(t-1),… ,Z(t-p). В этой модели текущее значение процесса выражается как конечная линейная совокупность предыдущих значений процесса и импульса, который называется «белым шумом»,

$$Z(t) = C + \phi_1 Z(t-1) + \phi_2 Z(t-2) + ... + \phi_p Z(t-p) + \varepsilon_t \quad (1)$$

Формула (1) описывает процесс авторегрессии порядка *p*, здесь *C* – вещественная константа, $\varphi_1,..,\varphi_p$ – коэффициенты, $\varepsilon_t$ – ошибка модели. Для определения $\varphi_i$ и C используют метод наименьших квадратов [2] или метод максимального правдоподобия [3].

Другой тип модели имеет большое значение в описании временных рядов и часто используется совместно с авторегрессией называется моделью скользящего среднего порядка *q* и описывается уравнением

$$Z(t) = \frac{1}{q}(Z(t-1) + Z(t-2) + ... + Z(t-q)) + \varepsilon_t \quad (2)$$

где *q* – порядок скользящего среднего, $\varepsilon_t$ – ошибка прогнозирования.

Важными достоинствами данного класса моделей являются их простота и прозрачность моделирования. На сегодняшний день данный класс моделей является одним из наиболее популярных [2], а потому в открытом доступе легко найти примеры применения авторегрессионных моделей для решения задач прогнозирования временных рядов различных предметных областей.

Недостатками данного класса моделей являются: большое число параметров модели, идентификация которых неоднозначна и ресурсоемка [3]; низкая адаптивность моделей, а также линейность и, как следствие, отсутствие способности моделирования нелинейных процессов, часто встречающихся на практике [8].

2. Другой популярной моделью является модель на основе искусственных нейронных сетей [5]. Нейронные сети состоят из нейронов (рис. 1).

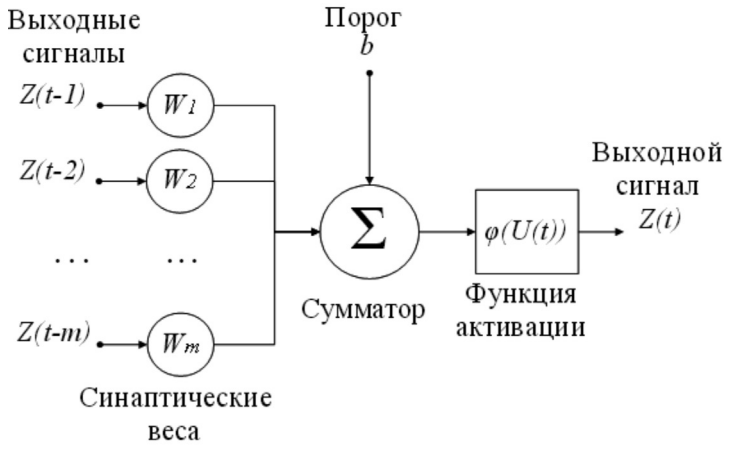


*Рис. 1. Нелинейная модель нейрона*

Модель нейрона можно описать парой уравнений

$$U(t) = \sum_{i=1}^{m} \omega_i Z(t-i) + b,$$

$$Z(t) = \varphi(U(t))$$

где $Z(t-1),..., Z(t-m)$ – входные сигналы; $\omega_1,...,\omega_m$ – синаптические веса нейрона; $p$ – порог; $\varphi(U(t))$ – функция активации.

При помощи нейронных сетей возможно моделирование нелинейной зависимости будущего значения временного ряда от его фактических значений и от значений внешних факторов. Нелинейная зависимость определяется структурой сети и функцией активации.

Основным достоинством нейросетевых моделей является нелинейность, т.е. способность устанавливать нелинейные зависимости между будущими и фактическими значениями процессов. Другими важными достоинствами являются: адаптивность, масштабируемость (параллельная структура ANN ускоряет вычисления) и единообразие их анализа и проектирования [4].

При этом недостатками ANN являются отсутствие прозрачности моделирования; сложность выбора архитектуры, высокие требования к непротиворечивости обучающей выборки; сложность выбора алгоритма обучения и ресурсоемкость процесса их обучения [4].

Одной из популярных современных тенденций является создание комбинированных моделей и методов. Подобный подход дает возможность компенсировать недостатки одних моделей при помощи других, и направлен на повышение точности, как одного из главных критериев эффективности модели.

**Математическая модель**

Для анализа экологических данных использована система, которая позволяет выявить закономерности временного изменения экологических показателей. Она включает такие сервисы, как

- выявление наличия сезонности;
- выделение трендов во временных рядах;
- определение сезонной декомпозиции.

Рассмотрим подробнее каждый.

*Выявление наличия сезонности*

Оценка наличия сезонных ритмов во временных рядах осуществлялась на основе функции автокорреляции и её графического представления – коррелограммы. При помощи анализа коррелограммы можно выявить структуру ряда. Последовательность коэффициентов автокорреляции со смещениями 1, 2, 3 и т.д. называют автокорреляционной функцией, значения которой находятся в диапазоне [– 1; 1]. Автокорреляционную функцию целесообразно использовать для выделения во временном ряде наличия трендовой и сезонной компонент.

*Выделение трендов во временных рядах*

Выявление наличия неслучайной составляющей сводилось к проверке гипотезы о неизменности среднего значения ряда с использованием критерия серий. При его использовании определяется медиана ряда, и образуются «серии» из плюсов и минусов по следующему правилу:

$$y_t = \begin{cases} +, y_t > y_{med} \\ -, y_t < y_{med} \end{cases}$$

Элементы временного ряда, равные $y_{med}$, в полученной последовательности не учитываются. Под «серией» понимается последовательность подряд идущих плюсов или подряд идущих минусов. Наличие

неслучайной составляющей во временном ряде определяется из условия:

$$\begin{cases} v(n) > \left[\frac{1}{2}(n+2-1,96\sqrt{n-1}\right], \\ K_{\max} < [3,3(\lg n+1)] \end{cases}$$

где $v(n)$ – общее число серий, $K_{\max}$ – длина наибольшей серии, [ ] – целая часть от числа [1].

Для построения тренда использовались два метода: скользящих средних и экспоненциального сглаживания [1].

Метод скользящих средних заключается в следующем:

1) определяется количество наблюдений, входящих в интервал сглаживания;

2) вычисляется среднее значение наблюдений в интервале сглаживания по формуле:

$$\overline{y}_t = \frac{1}{m}\sum_{i=t-\frac{m-1}{2}}^{t+\frac{m-1}{2}} y_i$$

где $m$ – количество наблюдений, входящих в интервал сглаживания. И так до тех пор, пока в интервал сглаживания не войдет последнее значение временного ряда.

В методе экспоненциального сглаживания каждое сглаженное значение рассчитывается путем сочетания предыдущего сглаженного значения и текущего значения временного ряда. В этом случае текущее значение временного ряда взвешивается с учётом сглаживающей константы:

$$S_t = \alpha y_t - (1-\alpha)S_{t-1}$$

где $S_t$ – текущее сглаженное значение; $y_t$ – текущее значение временного ряда; $S_{t-1}$ – предыдущее сглаженное значение; $\alpha$ – сглаживающая константа, значение которой варьируется в диапазоне от 0 до 1.

## Результаты исследования

В результате анализа экологических данных были получены следующие результаты, представленные на рис. 2–3. В результате исследования временных рядов на наличие периодичности и сезонности получено, что данные не являются периодичными и не содержат сезонности. Полученные результаты можно объяснить особенностью структуры данных.

Отсутствие в результатах проведенного исследования сезонности и периодичности свидетельствует о том, что основную массу отходов на исследуемых территориях составляют бытовые отходы (отходы потребления). Потребление населением товаров и услуг, в результате которых образуются отходы (упаковка, пластиковые бутылки, полиэтилен и др.) постоянны в течение года.

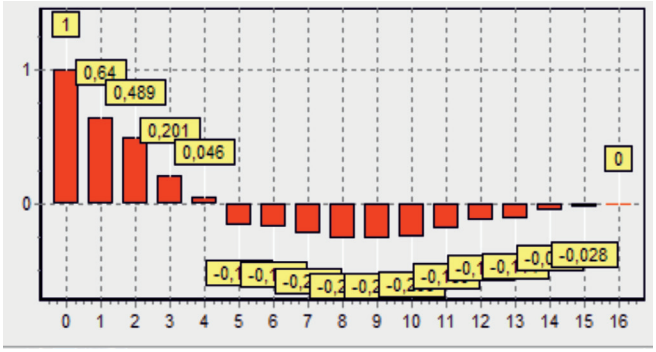


*Рис. 2. Коррелограмма (нет сезонной составляющей)*

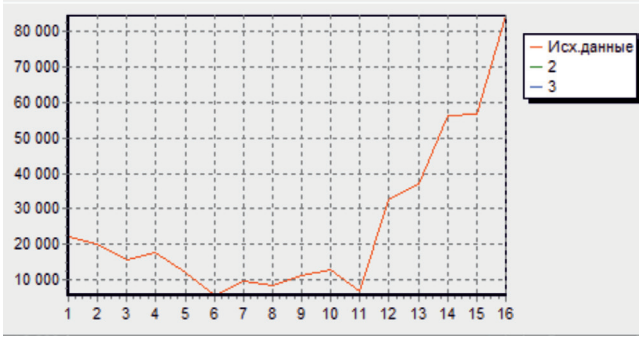


*Рис. 3. Исходный ряд*

Что касается отходов производства, то отсутствие сезонности и периодичности может свидетельствовать о более-менее постоянном процессе производства на территории Томской области, т.е. сфера производства, развитая на территории области, в целом работает в круглогодичном режиме, без особых сезонных изменений.

Хотя здесь необходимо отметить тот факт, что отчетность по отходам собирается ежегодно, 1 раз в год, а не поквартально или помесячно, что так же может отражаться на результатах исследования.

## Заключение

Отсутствие единой региональной системы сбора, вывоза, сортировки, использования, обезвреживания и размещения отходов различного происхождения, имеет значительные долгосрочные последствия для здоровья населения, экономики и окружающей среды. Поэтому важно содействовать экологически рациональному обращению с отходами и развивать программы их удаления, для чего необходимо проводить точные математические расчеты, на основании которых можно принимать решения о проведении защитных мероприятий по предупреждению ухудшения окружающей среды.

### Список литературы

1. Гергет О.М., Милешин А.А., Информационные технологии выявления основных закономерностей временного изменения показателей биосистемы // Современные проблемы науки и образования, 2013. – № 6.

2. Бокс Дж., Дженкинс Г.М. Анализ временных рядов, прогноз и управление. – М.: Мир, 1974. – 406 с.

3. Draper N., Smith H. Applied regression analysis. New York: Wiley, 1981. – 693.

4. Morariu N., Iancu E., Vlad S. A neural network model for time series forecasting // Romanian Journal of Economic Forecasting. – 2009. – № 4. – P. 213–223.

5. Хайкин С. Нейронные сети: полный курс. – М.: ООО «И.Д. Вильямс», 2006. – 1104 с.