

АЛГОРИТМЫ ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА: ОБНАРУЖЕНИЕ СТРУКТУРНЫХ СДВИГОВ МЕТЕОПАРАМЕТРОВ И КОНЦЕНТРАЦИЙ ЗАГРЯЗНЯЮЩИХ ВЕЩЕСТВ

Гребенюк Е. А.

В статье рассматриваются проблемы разработки алгоритмического обеспечения для систем экологического мониторинга параметров воздушной среды. Предложена стратегия предупредительного мониторинга, направленного на обнаружение возникающих нарушений до того момента, когда экологическая ситуация приобретет катастрофический характер. Для решения задачи разработаны методы обнаружения изменений параметрических моделей временных рядов, описывающих изменения концентраций загрязняющих веществ в атмосфере.

Введение. Для нормализации экологической обстановки необходимо создание региональных автоматизированных систем экологической защиты. В состав такой системы должны входить подсистемы мониторинга атмосферы, водных ресурсов, почв и других компонентов природной среды и подсистема поддержки принятия управленческих решений. Существующие системы, в основном, обеспечивают лишь наблюдение за параметрами окружающей среды и их регистрацию. Для существенного снижения риска химического загрязнения атмосферы необходимы функции оперативного прогноза возможных экологических нарушений. Среди задач, для решения которых предназначена подсистема мониторинга, важное место занимает задача предупреждающего контроля: оперативного обнаружения экологических нарушений до того момента, когда экологическая ситуация приобретет неблагоприятный или катастрофический характер. Для осуществления предупреждающего контроля и оперативного обнаружения экологических нарушений необходимо использование специального алгоритмического обеспечения, включающего:

— построение параметрических моделей загрязнителей во времени и в пространстве с учетом метеоусловий;

— обнаружение текущих изменений в процессе, которые могут привести процесс к выходу за норму.

Своевременное обнаружение таких изменений состояния атмосферы, которые могут привести к неблагоприятным экологическим ситуациям, позволит принять необходимые решения по их предупреждению и стабилизации состояния среды.

Указанное обеспечение может быть разработано на основе методов построения векторных параметрических моделей изменения концентраций загрязняющих веществ во времени и алгоритмов обнаружения изменения свойств, обнаруживающих отклонения исходного процесса от модели (Basseville, Nikiforov, 1993; Гребенюк, 2008).

Модели, описывающие изменения концентраций загрязняющих веществ. Задача построения моделей, описывающих процессы распространения концентраций загрязняющих веществ в атмосфере, представляет большой интерес для экологии, частности для предупреждения аварийных ситуаций.

В настоящее время существует много методов количественного описания процессов рассеяния газообразных веществ в атмосфере, наиболее широко распространенным являются:

- 1) гауссовские или дисперсионные модели рассеяния;
- 2) модели рассеяния, базирующиеся на интегральных законах сохранения;

- 3) модели, построенные на численном решении системы уравнений сохранения в их оригинальном виде (часто как модели или методы прямого численного моделирования). Однако сложность расчета параметров этих моделей и необходимость сбора и актуализации большого количества эмпирических данных делают их непригодными для оперативного текущего прогноза. Для целей текущего прогноза широко используются модели временных рядов, в которых все соотношения между характеристиками загрязняющих веществ и другими факторами, влияющими на уровень загрязнения, определяются на основе их наблюдаемых значений, а действие ненаблюдаемых факторов описывается моделью случайных возмущений (Roberson, Steyn, 1989). Модели изменения концентраций загрязнителей во времени, построенные только на основе анализа прошлых наблюдений, исследовались многими авторами (Modarres, Khosravi, 2005; Xie-Kang Wanga and Wei-Zhen Lub, 2006). Пусть $\hat{Y}_t = (y_{1t}, \dots, y_{kt})$ — вектор концентраций измеряемых загрязняю-

щих веществ, $\hat{X}_t = (x_{t1}, \dots, x_{tk})$ — вектор параметров метеорологических условий, измеренных в момент времени t . В общем случае, изменение концентраций во времени может быть описано в виде авторегрессионной модели с распределенными лагами — ADL (autoregressive distributed lag) модели:

$$\hat{Y}_t = \Lambda + (A_0 + A_1L + \dots + A_pL^p)\hat{Y}_t + \sum_{j=1}^l (B_jL + \dots + B_qL^q)\hat{X}_t + E_t. \quad (1)$$

Здесь $A_i, i = 0, 1, \dots, p, k \times k$ — матрица, $B_j, j = 0, 1, \dots, q, k \times l$ — матрица, L — оператор сдвига: $Ly_t = y_{t-1}$, $Lx_t = x_{t-1}$, $\hat{E}_t = (\varepsilon_{t1}, \dots, \varepsilon_{tk})$ — случайное возмущение, $E(\varepsilon_{ii}) = 0$, $Var(\varepsilon_{ii}) = \sigma^2, i = 1, \dots, k$, Λ — вектор констант. В процессе создания модели изменения концентраций в заданном регионе определяются структура модели (1) (наборы объясняющих переменных, порядки модели), вычисляются ее коэффициенты и проверяется адекватность исходным данным. В зависимости от метеорологических условий и интенсивностей источников загрязнения модель, построенная на участке процесса, может быть либо чисто авторегрессионной, либо чисто регрессионной, либо иметь смешанный характер.

Если модель адекватна процессу, то на момент времени $t + 1$, по наблюдениям, полученным до момента времени t включительно, может быть вычислен прогноз концентраций $\hat{Y}_{t+1,t}^P$:

$$\hat{Y}_{t+1,t}^P = (I - A_0)^{-1}(\Lambda + A_1Y_t + \dots + A_pY_{t-p+1} + B_1X_t + \dots + B_qX_{t-q}). \quad (2)$$

Этот прогноз является прогнозом с минимальной среднеквадратичной ошибкой. Прогноз на 2, 3 и большее число шагов вычисляются подстановкой в формулу (2) значений прогнозов, полученных по наблюдениям до момента времени t , и самих наблюдений.

Значения концентраций загрязнителей поступают на вход модели, по которой вычисляется прогноз. Если свойства процесса изменились, то изменяется и его модель. Если изменения имеют протяженность во времени, то процесс не сразу выходит за норму, и своевременное обнаружение факта наличия изменений позволит предупредить аварийную ситуацию или принять необходимые управленческие решения до того момента, когда ее развитие примет катастрофический характер.

Наличие изменений в процессе может быть определено на основе анализа отклонений процесса от модели. Нас будут интересовать такие

изменения в процессе, которые могут привести к выходу параметров за норму. Так как здесь рассматриваются модели изменений динамики наблюдаемых концентраций, т.е. авторегрессионной части модели (1), то анализ изменений проводится на модели без регрессоров вида:

$$\hat{Y}_t = \Lambda + (A_0 + A_1 L + \dots + A_p L^p) \hat{Y}_t + E_t. \quad (3)$$

1) Увеличение среднего в момент времени $t = t_p + 1$ для отдельных компонент процесса:

$$y_{it} = \lambda_i + (\lambda_{i1} - \lambda_i) \cdot DI_t + \sum_{i=1}^k \sum_{j=1}^p a'_{ij} y_{it-j} + \varepsilon_{it}. \quad (4)$$

Здесь $DI_t = \begin{cases} 1, & \text{если } t \geq t_p + 1 \\ 0, & \text{если } t < t_p + 1 \end{cases}$, $i \in \{i_1, i_2, \dots, i_m\} \subseteq \{1, 2, \dots, k\}$ — компоненты концентраций загрязнителей, для которых увеличилось среднее, $\{a'_{ij}\}$ — компоненты матриц A_j .

2) Модель изменения дисперсии в момент времени $t = t_p + 1$ для отдельных компонент процесса:

$$y_{it} = \begin{cases} \lambda_i + \sum_{i=1}^k \sum_{j=1}^p a'_{ij} y_{it-j} + \varepsilon_{it}, & \text{если } t < t_p + 1, \\ \lambda_i + \sum_{i=1}^k \sum_{j=1}^p a'_{ij} y_{it-j} + \eta_{it}, & \text{если } t \geq t_p + 1. \end{cases} \quad (5)$$

Здесь $\text{var}(\varepsilon_{it}) = \sigma_0^2$, $\text{var}(\eta_{it}) = \sigma^2 > \sigma_0^2$, $i \in \{i_1, i_2, \dots, i_m\} \subseteq \{1, 2, \dots, k\}$ — компоненты концентраций загрязнителей, для которых увеличилась дисперсия, $\{a'_{ij}\}$ как в (4).

3) Модель изменения типа процесса — от стационарного к нестационарному:

$$y_{it} = \begin{cases} \lambda_i + \sum_{i=1}^k \sum_{j=1}^p \alpha'_{ij} y_{it-j} + \varepsilon_{it}, & \text{если } t < t_p + 1, \\ \lambda_i + \sum_{i=1}^k \sum_{j=1}^p \alpha_{ij} y_{it-j} + \varepsilon_{it}, & \text{если } t \geq t_p + 1. \end{cases} \quad (6)$$

Здесь $\sum_{j=1}^p \alpha'_{ij} < 1$, $\sum_{j=1}^p \alpha_{ij} = 1$, $i \in \{i_1, i_2, \dots, i_m\} \subseteq \{1, 2, \dots, k\}$ — компоненты концентраций загрязнителей, которые вследствие изменений стали описываться нестационарными процессами.

Алгоритмы обнаружения изменений (4)–(6), приводящих к выходу процесса за норму. Для обнаружения изменений свойств процесса, подверженного воздействию случайных возмущений, по текущим наблюдениям широко применяются алгоритмы последовательного анализа (Siegmund, 1985). После получения очередного наблюдения алгоритм анализирует его и принимает решение о наличии либо отсутствии изменения свойств. Такой алгоритм представляет собой последовательную процедуру проверки гипотез; проверяется гипотеза H_0 «плотность распределения процесса равна P_1 », против альтернативной гипотезы H_1 «плотность распределения процесса равна P_2 », где P_1 и P_2 различаются своими параметрами. Для оценки качества работы алгоритма используются критерии:

— $E_1(\tau | \tau < t_p)$ — среднее время между «ложными тревогами», представляющее собой среднее число наблюдений через которое алгоритм подает сигнал о наличии изменений (отвержении гипотезы H_0), хотя процесс не изменился;

— $E_2(\tau - t_p | \tau \geq t_p)$ среднее время задержки в обнаружении — среднее число наблюдений после момента изменения, через которое алгоритм подает сигнал о наличии изменений.

В условиях точно известных параметров θ_1 и θ_2 до и после изменения свойств и неизвестного момента изменения t_p разработанный в (Page, 1954) алгоритм кумулятивных сумм, (CUSUM) обладает оптимальными свойствами в смысле критерия минимизации средней задержки при ограничении среднего времени до ложной тревоги. Условия, в которых доказана оптимальность CUSUM являются слишком жесткими. Например, предположения о точно известных параметрах до или после момента обнаружения очень редко выполняются на практике. Поэтому были предприняты значительные усилия для разработки алгоритмов текущего обнаружения, которые могли бы быть полезными для практических целей. Обнаружение изменений (4)–(6) при мониторинге

процесса с использованием алгоритмов текущего обнаружения включает следующие шаги:

1) Оценивается модель процесса (1) по наблюдениям $\hat{Y}_t, \hat{Y}_{t-1}, \dots, \hat{Y}_{t-k}, \dots, \hat{X}_t, \dots, \hat{X}_{t-k}, \dots$. Затем для каждой компоненты процесса последовательным добавлением и вычитанием членов $\alpha_{ij}y_{it-j}$ авторегрессионной части каждая компонента модели (1) приводится к виду:

$$\Delta y_{it} = \lambda_i + \beta_0 y_{it-1} + \sum_{i=1}^k \sum_{j=1}^{p-1} \beta'_{ij} \Delta y_{it-j} + \sum_{i=1}^l \sum_{j=1}^q \gamma_{ij} x_{it-j} + \varepsilon_{it}. \quad (7)$$

2) По наблюдениям $\Delta \hat{Y}_t, \dots, \Delta \hat{Y}_{t-p-1}, \hat{Y}_t, \hat{X}_t, \dots, \hat{X}_{t-q}$ и вычисленным значениям коэффициентов модели (7) для каждой наблюдаемой компоненты процесса вычисляются статистики: (Bascompte, Solé, 1995).

$$T_{it}^1 = \Delta y_{it} - \lambda_i - \beta_0 y_{it-1} - \sum_{i=1}^k \sum_{j=1}^{p-1} \beta'_{ij} \Delta y_{it-j} - \sum_{i=1}^l \sum_{j=1}^q \gamma_{ij} x_{it-j},$$

$$T_{it}^2 = \Delta y_{it} - \sum_{i=1}^k \sum_{j=1}^{p-1} \beta'_{ij} \Delta y_{it-j} - \sum_{i=1}^l \sum_{j=1}^q \gamma_{ij} x_{it-j}.$$

В отсутствие изменений математическое ожидание этих статистик равно нулю. Если в процессе происходят изменения (4), то увеличиваются средние значения T_1 и T_2 , а дисперсия их не меняется. При изменениях вида (5) увеличивается дисперсия обеих статистик, причем выполняется соотношение:

$$Var(T_1) < Var(T_2). \quad (8)$$

В случае изменений (6) увеличиваются математическое ожидание T_1 и дисперсии T_1 и T_2 , при этом выполняется соотношение:

$$Var(T_1) > Var(T_2). \quad (9)$$

3) При получении очередного наблюдения вычисляются статистики T_1 и T_2 , для которых алгоритмом последовательного наблюдения проверяется:

- увеличение математического ожидания этих статистик относительно нулевого значения,
- увеличение их дисперсий,
- сравнение дисперсий T_1 и T_2 .

При обнаружении алгоритмом увеличения математического ожидания и постоянной дисперсии диагностируется изменение (4), при уве-

личении дисперсии, постоянном математическом ожидании и выполнении условия (8) — изменение типа (5). Если, одновременно, увеличивается математическое обеих статистик и их дисперсия, то диагностируется изменение (6) переход процесса в нестационарное состояние. Алгоритмы последовательного обнаружения изменений среднего и дисперсии независимой случайной последовательности рассматривались во многих работах, например (Han, Tsung, 2007; Basseville, Nikiforov, 1993). Поскольку точная оценка увеличения среднего и дисперсии после изменений неизвестна, то для мониторинга изменений целесообразно использование упрощенных алгоритмов, в которых вместо известных значений параметров после изменений задается граничное значение параметра, выход за которое нужно обнаруживать (Гребенюк, 2003). Показано, что средняя задержка в обнаружении для алгоритма (3) стремится к своему оптимальному значению при приближении заданных граничных значений к фактическим.

Пример. Рассмотрим процесс мониторинга содержания в атмосфере диоксида азота. График процесса представлен на рис. 1, горизонтальная линия - пороговое значение загрязнителя, вертикальная линия отмечает момент превышения порогового значения загрязнителя.

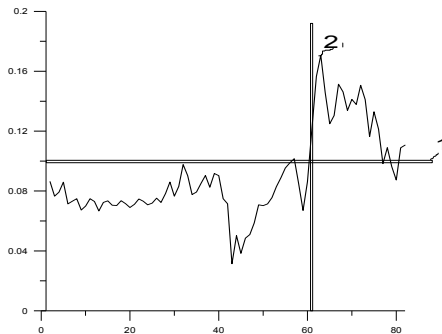


Рис. 1. 1 — предельно допустимая разовая концентрация, 2 — концентрация диоксида азота.

Модель процесса, построенная с учетом влияния метеоусловий, представлена на рис. 2. Внизу — график отклонений процесса от модели.

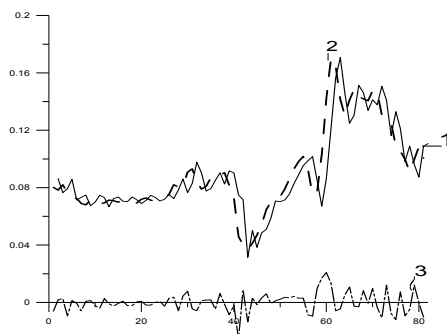


Рис. 2. 1 — наблюдаемые значения концентраций, 2 — значения, вычисленные по модели, 3 — разность между фактическими значениями и моделью.

На рис 3 представлены результаты мониторинга отклонений от модели. Вертикальными линиями обозначены моменты фактического превышения порогового значения загрязнителя ($N = 229$) и момент обнаружения изменения модели ($N = 220$).

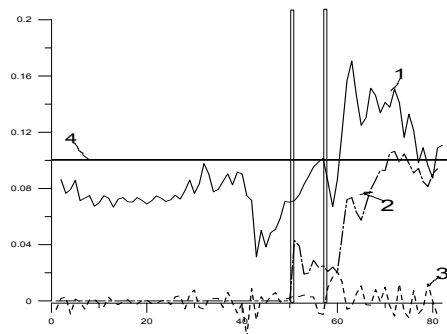


Рис. 3. 1 — наблюдаемые значения концентраций, 2 — решающая функция алгоритма обнаружения; 3 — разность между фактическими значениями и моделью, 4 — порог алгоритма обнаружения.

Таким образом, сигнал о несоответствии модели процессу, свидетельствующий о возникших изменениях, может быть подан существенно раньше, чем произойдет фактический выход параметров за норму.

Заключение. Предложено решение задачи предупреждения возникновения неблагоприятных экологических ситуаций, основанное на

обнаружении изменений в процессах распространения загрязнителей, приводящих к выходу этих концентраций за норму. Предложенные алгоритмические процедуры включают описание процессов с помощью параметрических моделей и обнаружение изменений по отклонениям процесса от модели с использованием CUSUM алгоритмов.

СПИСОК ЛИТЕРАТУРЫ

- Basseville M., Nikiforov I.V.* Detection of abrupt changes — theory and application. — New Jersey: Prentice Hall in Information and System Sciences, 1993.
- Гребенюк Е.А.* Построение моделей, описывающих динамику изменения во времени концентраций загрязняющих веществ в атмосфере // *Проблемы управления.* — 2008. — № 6. — С. 55–64.
- Гребенюк Е.А.* Анализ и оперативная диагностика систем, описываемых нестационарными случайными процессами // *Проблемы управления.* — 2003. — № 4. — С. 23–29.
- Han, D. and Tsung, F.* Detection and Diagnosis of Unknown Abrupt Changes using CUSUM Multi-Chart Schemes // *Sequential Analysis.* — 2007. — Vol. 26. — P. 225–249.
- Modarres R and Khosravi Dehkordi A.* Daily air pollution time series analysis of Isfahan City // *International Journal of Environmental Science and Technology.* — 2005. — Vol. 2, no 3. — P. 259–267.
- Page E.S.* Continuous inspection schemes // *Biometrika.* — 1954. — Vol. 41, no 1. — P. 110–115.
- Siegmund, D.* Sequential Analysis. Tests and Confidence Intervals. — New York, Springer-Verlag, 1985.
- Xie-Kang Wang and Wei-Zhen Lub* Seasonal variation of air pollution index: Hong Kong case study // *Chemosphere.* — 2006. — Vol. 63, no 8. — P. 1261–1272.

**SOFTWARE OF ECOLOGICAL MONITORING: DETECTION OF
STRUCTURAL SHIFTS IN METEOROLOGICAL
PARAMETERS AND POLLUTION CONCENTRATIONS**

Grebenuk E. A.

The paper is devoted to the problem of monitoring of pollutant concentration and preventing excessive contamination of the environment. Algorithms for sequential change detection in stationary and non-stationary processes were developed. The designed algorithms of monitoring were probed on real data of ozone concentration in troposphere.