

СРЕДНЕСРОЧНОЕ ПРОГНОЗИРОВАНИЕ КОЛИЧЕСТВА ПОЖАРОВ С ИСПОЛЬЗОВАНИЕМ АВТОКОРРЕЛЯЦИОННЫХ ФУНКЦИЙ

А.Н. Батуро;

С.А. Техтереков, кандидат педагогических наук;

В.С. Коморовский, кандидат технических наук.

**Сибирская пожарно-спасательная академия –
филиал Санкт-Петербургского университета ГПС МЧС России**

Рассмотрены вопросы оперативного прогнозирования количества техногенных пожаров на региональном уровне. Основным методом прогноза является использование автокорреляционных функций. Также рассмотрена пороговая фильтрация данных для повышения качества прогнозов.

Ключевые слова: прогнозирование, техногенные пожары, автокорреляция, преобразование Фурье, пороговая фильтрация

MEDIUM-TERM FORECASTING OF NUMBER OF FIRES USING AUTOCORRELATION FUNCTIONS

A.N. Baturo; S.A. Tekhterekov; V.S. Komorovskiy. Siberian fire and rescue academy – branch of Saint-Petersburg university of State fire service of EMERCOM of Russia

The article deals with the operational forecasting of number of man-made fires at the regional level. Using the autocorrelation functions is the main method of forecasting. The threshold filtering data is considered to improve the forecasts quality.

Keywords: forecasting, man-made fires, autocorrelation, Fourier transform, threshold filtering

Необходимость составления прогноза пожаров на период один–два месяца обусловлена решением задач планирования обеспечения подразделений ГПС МЧС России, решения кадровых вопросов, определение режимов несения службы. Особенно актуальным это является в труднодоступных районах Крайнего Севера и обширных по площади регионах Сибири, где своевременное обеспечение подразделений всем необходимым может быть затруднено как протяженностью путей сообщения, так и погодными условиями. Кроме того, оперативный прогноз необходим для организации и проведения работ по профилактике и предотвращению пожаров.

Целью данной работы является разработка методики оценочного оперативного прогнозирования количества техногенных пожаров на примере регионов Сибири и Дальнего Востока.

Работа проводилась в рамках проекта по созданию универсальной методики прогнозирования количества техногенных пожаров в различных регионах России. Актуальность работы определяется тем, что существующая в подразделениях МЧС России практика использования показателя «Аналогичный период прошлого года» как основы построения прогнозов и отчетности не в полной мере использует потенциал ретроспективных данных и требует совершенствования.

В настоящей работе прогнозирование основано на использовании автокорреляционных функций для прогнозирования временных рядов. Этот математический аппарат широко используется в экономике, демографии, метеорологии [1–3]. Однако для решения задач обеспечения техносферной безопасности данный метод используется сравнительно редко [4, 5].

Автокорреляционная модель процесса возникновения пожаров

Рассмотрим данные о количестве пожаров в Красноярском крае по месяцам за период с января 2006 г. по июнь 2011 г. [6]. На рис. 1 приведен график изменения количества пожаров за 66 месяцев.



Рис. 1. Динамика пожаров в Красноярском крае

Как можно заметить из рис. 1, зависимость количества пожаров от времени представляет собой периодическую функцию с линейным убывающим трендом (линия тренда добавлена на рисунке). Для составления прогноза воспользуемся методами прогнозирования временных рядов [7].

Предположим, что возникновение техногенных пожаров представляет собой стационарный процесс $y(t)$. Найдем автокорреляционную функцию этого процесса.

Основными статистическими характеристиками временного ряда являются математическое ожидание:

$$\bar{y}(t) = M[y(t)],$$

дисперсия:

$$\sigma^2(t) = D[y(t)],$$

и автокорреляционная функция временного ряда:

$$\rho(t, s) = \frac{M[(y(t) - \bar{y}(t)) * (y(\tau) - \bar{y}(\tau))]}{\sigma(t) * \sigma(s)},$$

где $\sigma(t)$, $\sigma(s)$ – средние квадратические отклонения [8, 9]. В приведенных формулах и далее черта над переменной означает осреднение по времени.

Оценками автокорреляционной функции являются выборочные коэффициенты автокорреляции:

$$r(t, t-L) = \frac{\overline{y(t) * y(t-L)} - \overline{y(t)} * \overline{y(t-L)}}{s(t) * s(t-L)},$$

где:

$$s(t) = \sqrt{\overline{y(t) * y(t)} - \overline{y(t)} * \overline{y(t)}};$$

$$s(t-L) = \sqrt{\overline{y(t-L) * y(t-L)} - \overline{y(t-L)} * \overline{y(t-L)}};$$

$$\overline{y(t) * y(t-L)} = \left(\frac{\sum_{i=1+L}^N y(i)y(i-L)}{N-L} \right), \quad \overline{y(t) * y(t)} = \left(\frac{\sum_{i=1+L}^N y(i)y(i)}{N} \right), \quad \overline{y(t)} = \left(\frac{\sum_{i=1+L}^N y(i)}{N-L} \right);$$

$$\overline{y(t-L)} = \left(\frac{\sum_{i=1+L}^N y(i-L)}{N-L} \right).$$

здесь N – длина ряда, величина лага L .

Коррелограмма количества возникших пожаров в Красноярском крае представлена на рис. 2. Пунктирными линиями показан интервал доверительной трубки при доверительной вероятности 0,95. Величина лага была выбрана от $L=1$ до $L=15$, что было несколько больше, чем визуально определяемый период колебания количества пожаров.



Рис. 2. Коррелограмма количества возникших пожаров в Красноярском крае

Статистический критерий Дарбина–Уотсона для приведенных данных составляет $DW=0,059$ при $DW_{up}=1,634$, $DW_{low}=1,574$. Следовательно, можно предположить наличие положительной автокорреляции, по крайней мере, первого порядка. Проверка значимости коэффициентов автокорреляции по критерию Стьюдента на уровне значимости 0,05 показывает, что все коэффициенты значимо отличаются от нуля. Как видно из рис. 2, максимальные коэффициенты автокорреляции приходятся на $L=1$ и $L=12$, что говорит о наличии выраженной тенденции (тренда) и выраженной периодичности колебаний количества пожаров, с периодом в 12 месяцев. Это хорошо согласуется с априорной информацией.

Модель процесса будем искать в виде:

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + \dots + a_n y(t-n) + \varepsilon(t), \quad (1)$$

где n – максимальный учитываемый лаг; $\varepsilon(t)$ – ошибка модели в момент времени t . Параметры модели a_i могут быть выражены через выборочные коэффициенты автокорреляции $r(i)$ в системе уравнений Юла–Уокера:

$$\begin{cases} r(1) = a_1 + a_2 r(1) + \dots + a_n r(n-1), \\ r(2) = a_1 r(1) + a_2 + \dots + a_n r(n-2), \\ \vdots \\ r(n) = a_1 r(n-1) + a_2 r(n-2) + \dots + a_n, \end{cases}$$

где n – максимальный лаг. Представим решение данной системы в матричной форме:

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & r(1) & \dots & r(n-1) \\ r(1) & 1 & \dots & r(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(n-1) & r(n-2) & \dots & 1 \end{pmatrix}^{-1} \times \begin{pmatrix} r(1) \\ r(2) \\ \vdots \\ r(n) \end{pmatrix}.$$

Для удобства анализа временного ряда сначала проведем его центрирование, вычтя из значений $y(t)$ трендовую составляющую:

$$z(t) = y(t) - y_{tr}(t), \quad (2)$$

где $y_{tr}(t)$ – линейный тренд процесса:

$$y_{tr}(t) = a \cdot t + b.$$

Коэффициенты линейной модели тренда, найденные по методу наименьших квадратов, равны $a=-0,9528$, $b=435,93$. Величина коэффициента a говорит о снижении количества пожаров в среднем на один пожар/мес. Далее, для центрированного процесса пересчитаем коэффициенты автокорреляции и решим систему уравнений Юла–Уокера. Подставим найденные значения в модель (1), при $n=12$. Таким образом, получим прогнозирующую модель в виде:

$$z(t)=0,533 \cdot z(t-1)-0,015 \cdot z(t-2)-0,098 \cdot z(t-3)+0,042 \cdot z(t-4)-0,213 \cdot z(t-5)-0,363 \cdot z(t-6)-0,187 \cdot z(t-7)+0,032 \cdot z(t-8)-0,127 \cdot z(t-9)-0,084 \cdot z(t-10)+0,298 \cdot z(t-11)+0,579 \cdot z(t-12).$$

Подставив значения центрированной функции, получаем вектор модельных данных, без учета трендовой компоненты. Проведя преобразование, обратное (2), получим модель количества пожаров. Оценим ее погрешность:

$$\varepsilon = t_{\alpha, n-1} \cdot \sqrt{\frac{\sum_{i=1}^n (z(t) - y(t))^2}{n \cdot (n-1)}},$$

где $t_{\alpha, n-1}$ – критерий Стьюдента при уровне значимости $\alpha = 0,05$, n – объем выборки.

Для данного примера получим $\varepsilon = 114$. Проверим метод с помощью ретро-прогнозирования на один месяц вперед, то есть на июль (67-й месяц в нашей постановке задачи):

$$z(67) = 291 \mp 114 \text{ пожаров.}$$

Фактически, на июль 2011 г. в Красноярском крае было зарегистрировано 269 пожаров, ошибка прогноза составила 22 пожара, в сторону завышения, или 8,2 % от фактического числа пожаров.

Средняя ошибка прогноза составляет 51 пожар по абсолютной величине или около 13 % от среднего количества пожаров за рассматриваемый период.

На рис. 3 приведен график исходных данных и модели с прогнозом на один месяц.

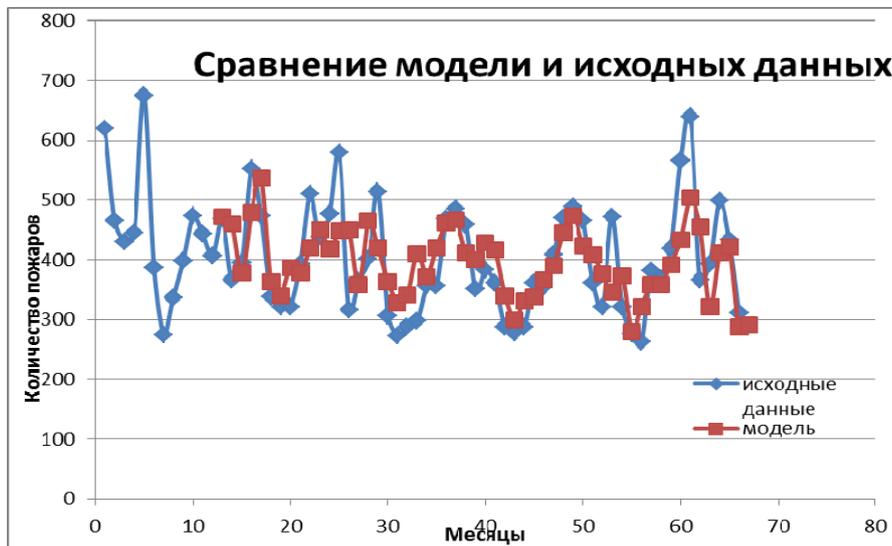


Рис. 3. Сравнение исходных данных и модели

Таким образом, получили удовлетворительное совпадение фактических и модельных данных. Построим аналогичные модели для других регионов. Основные результаты моделирования приведены в табл. 1.

Наихудшие результаты имеем для Красноярского края, что, по-видимому, связано с наибольшей географической неоднородностью территории данного региона.

Таблица 1. Характеристики автокорреляционных моделей для разных регионов

Регион	Математическое ожидание	Полуширина доверительного интервала	Средняя ошибка	Доля ошибки от среднего, %
Красноярский край	406	114	51	13
Забайкальский край	177	49	19	10
Новосибирская область	337	77	30	9
Иркутская область	360	84	32	9

Исследование спектра процесса возникновения пожаров

Поскольку процесс возникновения пожаров можно рассматривать как некоторый сигнал дискретного времени, для построения прогнозирующих моделей имеет смысл изучить спектр этого сигнала. Для получения спектра будем использовать преобразование Фурье:

$$Y(i\omega) = \int_0^{\infty} y(t)e^{-i\omega t} dt ,$$

где ω – частота сигнала.

Вновь обратимся к данным по Красноярскому краю, сократив временной ряд до $N=64$ значений (это связано с реализацией алгоритма быстрого преобразования Фурье в системе Mathcad). На рис. 4 приведен спектр возникновения пожаров в Красноярском крае.

Как видно из рис. 4, существенные амплитуды имеются во всех частотных областях спектра, хотя больший вклад вносят области средних и низких частот. Применим пороговый фильтр:

$$G_j = Y_j \Phi(|Y_j| - p) ,$$

где p – величина порога, выбираемая из условия:

$$\varepsilon = \sum_{j=1}^N (y_j(t) - g_j(t))^2 \rightarrow \min ,$$

где $g_j(t)$ – сигнал после фильтрации, полученный обратным преобразованием Фурье над G . На рис. 4 приведена величина выбранного порога $p=85$.

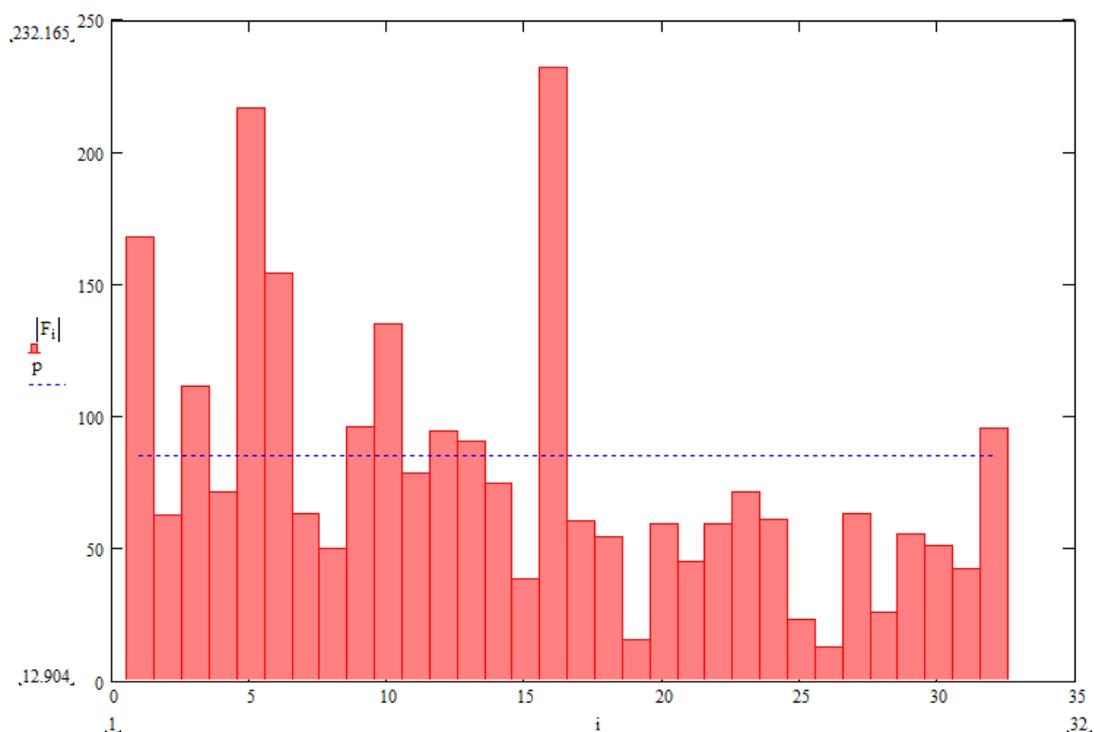


Рис. 4. Спектр временного ряда количества пожаров в Красноярском крае

На рис. 5 приведены исходная модель и модель после фильтрации. Построим автокорреляционную модель полученного после фильтрации сигнала по методике, описанной выше. Результаты прогнозирования для различных регионов приведены в табл. 2.

Как следует из табл. 2, имеется незначительное улучшение прогнозных характеристик модели по двум регионам, по Новосибирской области характеристики практически не изменились, по Забайкальскому краю наблюдается незначительное ухудшение характеристик. Таким образом, на данном этапе исследования нельзя дать однозначный ответ на вопрос о необходимости применения спектров Фурье и пороговой фильтрации в задачах прогнозирования пожаров. Однако полученные результаты свидетельствуют о перспективности дальнейшего изучения этого вопроса, в том числе привлечение оконного преобразования Фурье и вейвлет-преобразования.

Таблица 2. Характеристики автокорреляционных моделей для разных регионов после фильтрации

Регион	Математическое ожидание	Полуширина доверительного интервала	Средняя ошибка	Доля ошибки от среднего, %
Красноярский край	418	100	44	11
Забайкальский край	179	57	24	13
Новосибирская область	338	63	31	9
Иркутская область	364	73	30	8

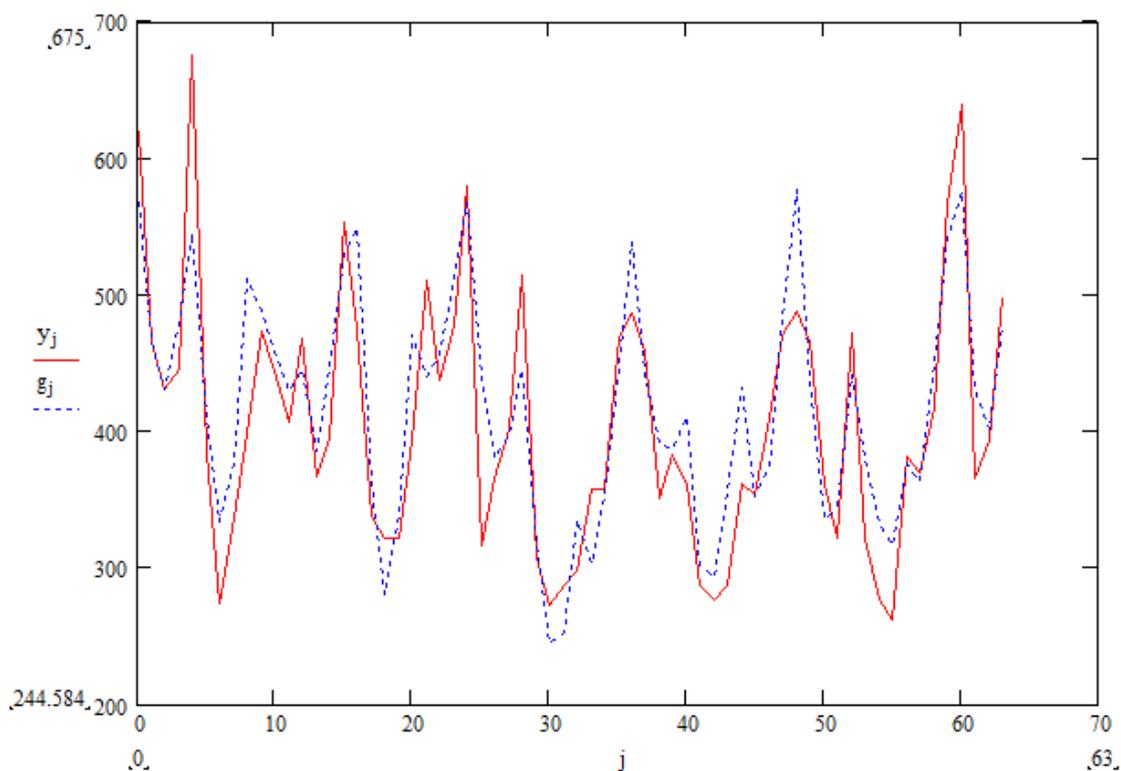


Рис. 5. Исходные данные и модель после пороговой фильтрации

Основные результаты и выводы

В настоящее время существует практика прогнозирования количества пожаров, основанная на использовании данных за аналогичный период прошлого года. Подобный подход нуждается, по крайней мере, в уточнении. Для уточнения прогнозов использовался хорошо известный в экономике и эконометрике метод автокорреляции.

Полученные автокорреляционные модели позволяют прогнозировать количество пожаров на месяц вперед с точностью порядка 87–92 % от среднего количества пожаров за весь исследуемый период. Исследования проводились по четырем регионам: Красноярский край, Забайкальский край, Новосибирская область, Иркутская область.

Дополнительно, для повышения качества прогнозирующих моделей, был исследован Фурье-спектр процесса возникновения пожаров и применена пороговая фильтрация. Однако на данном этапе существенных улучшений добиться не удалось.

Предложенный подход к прогнозированию количества техногенных пожаров в масштабе региона позволяет добиться удовлетворительного качества прогноза на период в один месяц. Перспективой данной работы является изучение спектров количества пожаров, полученных путем оконного преобразования Фурье и вейвлет-преобразования.

Литература

1. Ширнаева С.Ю. Краткосрочное прогнозирование факторов макроэкономической стабилизации экономики России // Вестник Оренбургского государственного университета. 2011. № 120. С. 133–137.
2. Прогноз урожайности зерновых культур как основа прогнозирования объемов сельскохозяйственного производства / К.П. Личко [и др.] // Проблемы прогнозирования. 2007. № 3. С. 60–67.

3. Разложение, анализ и прогноз временных рядов метеопараметров / Л.Ж. Шугунов [и др.] // Известия Таганрогского государственного радиотехнического университета. 2005. Т. 55. № 11. С. 15–21.
4. Shumway R., Stoffer D.S. Time Series Analysis and its Applications. Berlin: Springer, 2000. P. 549.
5. Моделирование взаимодействия природных пожаров и объектов защиты / М.В. Берестенькова [и др.] // Хвойные бореальной зоны. 2012. Т. XXX. № 5–6. С. 103–106.
6. Статистика пожаров // Статистика пожаров, 2011. URL: <http://statp.ru/index.php> (дата обращения: 11.08.2014).
7. Батуро А.Н. Прогнозирование количества пожаров в регионе на основе теории временных рядов // Технологии гражданской безопасности. 2013. № 3. С. 84–88.
8. Арутюнян С.А. Системы цифровой обработки сигналов. Красноярск: Сиб. гос. технол. ун-т, 2010. 53 с.
9. Доррер Г.А. Теория информационных процессов и систем. Красноярск : Сиб. гос. технол. ун-т, 2009. 202 с.