

УДК 519.254.1

Моделирование и прогноз развития лесопожарных ситуаций с использованием векторных авторегрессионных процессов

А. В. Парамонов

*Центр Математических Исследований МГТУ «СТАНКИН»
Вадковский пер. д. 3А, Москва, Россия, 117936*

Проведено исследование статистических свойств совместной выборки ежедневных значений количества очагов лесных пожаров и комплексного метеорологического показателя в Иркутской области за 1969–1988 гг. Построена математическая модель стохастического процесса развития лесопожарных ситуаций с использованием векторного авторегрессионного процесса VAR(1). Разработан вычислительный алгоритм моделирования будущих значений количества очагов возгорания. Показано, как он может быть использован для прогноза развития лесопожарной ситуации с высокой степенью надёжности на 1–2 суток вперёд.

Ключевые слова: лесные пожары, комплексный показатель Нестерова, векторный авторегрессионный процесс, величина риска.

1. Введение

Лесные пожары представляют собой масштабное стихийное природное явление, которое регулярно наносит серьёзный урон российской экономике. Процесс развития лесных пожаров практически не поддаётся контролю со стороны человека.

В то же время существуют и активно используются практические методики профилактики и противодействия лесным пожарам. Эффективность применения этих методик самым непосредственным образом зависит от умения моделировать и предсказывать развитие лесопожарных ситуаций. Точный и своевременный прогноз даёт возможность сосредоточить ограниченные ресурсы на тех мероприятиях, которые позволят минимизировать потенциальные последствия пожара.

2. Показатель Нестерова

Общие требования по мониторингу и прогнозированию лесных пожаров и чрезвычайных лесопожарных ситуаций регулируются ГОСТ Р 22.1.09–99 «Безопасность в чрезвычайных ситуациях. Мониторинг и прогнозирование лесных пожаров. Общие требования» от 25 мая 1999 г. [1]. Степень пожарной опасности в лесу по условиям погоды должна определяться по принятому в лесном хозяйстве комплексному показателю В. Г. Нестерова, который вычисляется на основе данных о температуре воздуха, температуре точки росы и количестве выпавших осадков:

$$\text{КП} = \sum_{i=1}^n T(T - T_d),$$

где T — температура воздуха в градусах Цельсия, T_d — температура точки росы в градусах Цельсия, n — число дней после последнего дождя (осадки до 2,5 мм в расчёт не принимаются) [2]. В настоящей работе расчёт значений точки росы произведён на основе имеющихся значений температуры и влажности воздуха по формулам:

$$T_d = \frac{b\gamma(T, RH)}{a - \gamma(T, RH)}, \quad (1)$$

$$\gamma(T, RH) = \frac{aT}{b + T} + \ln(RH/100), \quad (2)$$

где RH — относительная влажность в процентах, $a = 17,27$, $b = 237,7^\circ\text{C}$. Формулы (1)–(2) обладают погрешностью не более $0,4^\circ\text{C}$ в следующем диапазоне значений [3]:

$$0^\circ\text{C} < T < 60^\circ\text{C}, \quad 1\% < RH < 100\%, \quad 0^\circ\text{C} < T_d < 50^\circ\text{C}.$$

3. Модель МАР(1)С

Проведено исследование статистических свойств совместной выборки ежедневных значений количества очагов возгорания (далее — ОВ) и комплексного показателя В. Г. Нестерова (далее — КП) в Иркутской области за 1969–1988 гг.:

$$X_{t_k} = \begin{pmatrix} \text{ОВ}_{t_k} \\ \text{КП}_{t_k} \end{pmatrix}, \quad k = pk_p,$$

где $p = 20$ — число сезонов (периодов), $k_p = 214$ — число наблюдений за сезон (наблюдения проводились с 1 апреля по 31 октября)¹.

Во-первых, видно (см. рис. 1), что временные ряды обладают выраженной мультипликативной сезонной компонентой. Переход к аддитивной модели осуществлён путём логарифмирования данных. Выделение сезонной компоненты произведено методом медианного усреднения логарифмированных значений для каждой даты за все годы наблюдений (см. рис. 2).

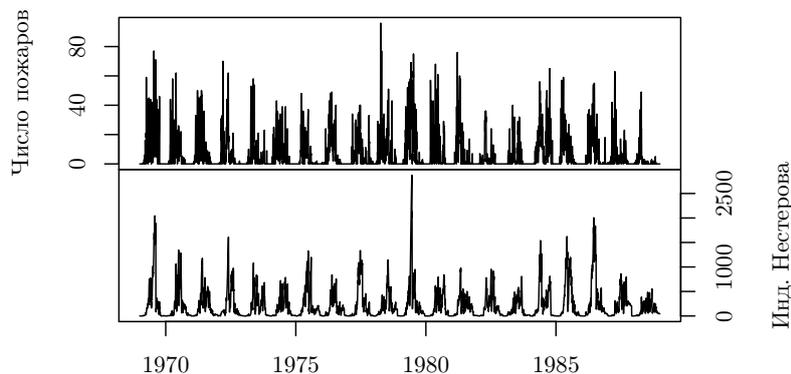


Рис. 1. Исходные данные

К полученным значениям сезонной компоненты применена сплайн-аппроксимация. Это позволило рассчитать значения сезонной компоненты для дат, отсутствующих в исходной выборке (до 1 апреля и после 31 октября).

Стохастическая компонента получена путём вычитания сезонной компоненты из логарифмированных данных:

$$Y_{t_k} = \ln(X_{t_k}) - s(\tau_k), \quad \tau_k = \frac{k \bmod k_p}{k_p}.$$

¹Данные по метеорологическим показателям (температура воздуха, относительная влажность, количество осадков) предоставлены Государственным учреждением «Всероссийский научно-исследовательский институт гидрометеорологической информации — Мировой центр данных». Данные по количеству очагов возгорания предоставлены предприятием «Авиалесоохрана».

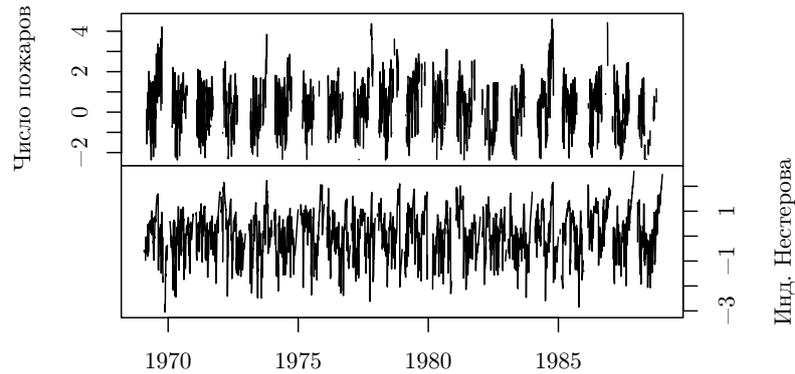


Рис. 2. Стохастическая компонента (логарифмированные данные за вычетом сезонной компоненты)

Полученные значения характеризуют, на сколько порядков наблюдаемая величина отличается от ожидаемой (средней за все годы наблюдений):

$$X_{t_k} = \exp(Y_{t_k})m(\tau_k), \quad m(\tau_k) = \exp(s(\tau_k)).$$

Эмпирические авто- и кросскорреляционные функции совместного распределения стохастической компоненты существенно отклоняются от нуля и имеют вид затухающих экспонент. Это указывает на наличие тренда, который может быть описан векторной авторегрессионной моделью VAR(1) [4]:

$$y_{t_k} = Ay_{t_{k-1}} + e_{t_k}, \quad (3)$$

где $y_{t_k} \in \mathbb{R}^2$ — вектор значений стохастической компоненты в момент времени t_k , $k > 1$, $A \in \mathbb{R}^{2 \times 2}$ — матрица параметров модели, $e_{t_k} \in \mathbb{R}^2$ — вектор остаточных значений в момент времени t_k (случайная компонента). Численные оценки значений параметров модели получены методом Юла-Уолкера [5]:

$$A = \begin{array}{c} \text{ОВ} \quad \text{КП} \\ \text{ОВ} \quad \left(\begin{array}{cc} 0,69 & 0,06 \\ 0,06 & 0,79 \end{array} \right) \\ \text{КП} \end{array}.$$

Авто- и кросскорреляционные функции совместного распределения остаточных членов (случайной компоненты) не отклоняются существенно (на уровне значимости 0,05) от нуля (рис. 3).

Частное распределение значений случайной компоненты ОВ хорошо описывается нормальным законом, особенно для положительных значений. Статистические тесты Андерсона—Дарлинга, Крамера—фон—Мизеса и Шапиро—Франсиа требуют отвергнуть гипотезу об отклонении эмпирического распределения от нормального даже на уровне значимости 0,1. Наблюдаемое расхождение для отрицательных значений обусловлено, по-видимому, дискретностью исходных данных.

Частное распределение значений случайной компоненты КП является асимметричным и, в отличие от распределения значений случайной компоненты ОВ, не может быть описано нормальным законом. Как видно, распределения случайной компоненты ОВ и КП существенно различаются по своим статистическим свойствам. Это подтверждает мнения некоторых исследователей о том, что индекс Нестерова в том виде, как он описан в ГОСТ, не может считаться вполне адекватной характеристикой пожароопасности [6, 7].

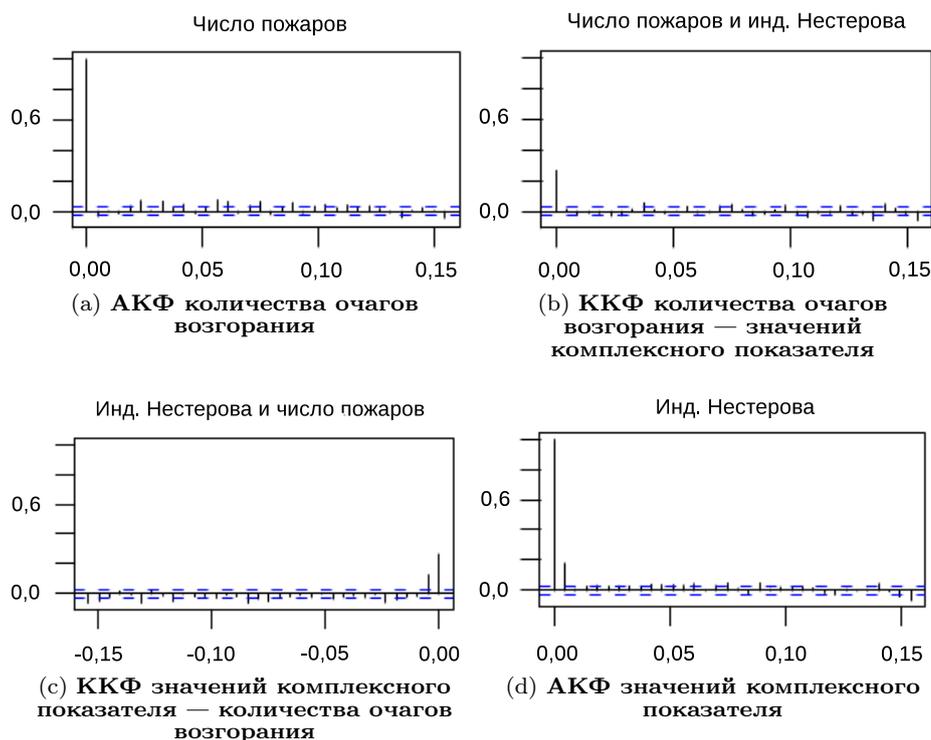


Рис. 3. Эмпирические авто- и кросскорреляционные функции совместного распределения остаточных членов модели ВАР(1). Пунктиром обозначен 95% доверительный интервал в предположении о стохастической независимости

Как бы то ни было, распределения обладают выраженной статистической связью: коэффициент корреляции совместного распределения остаточных членов равен 0,28. В качестве параметрической модели совместного распределения случайной компоненты использовано скошенное нормальное распределение [8]:

$$e_{t_k} \sim SN_2(\sigma\Omega\sigma^T, \alpha), \quad \text{т.е.} \quad p(e_{t_k}) = 2\phi_2(e_{t_k}, \sigma\Omega\sigma^T)\Phi(\alpha^T e_{t_k}),$$

где $\phi_2(e_{t_k}, \sigma\Omega\sigma^T)$ — плотность 2-мерного нормального распределения с ковариационной матрицей $\sigma\Omega\sigma^T \in \mathbb{R}^{2 \times 2}$, $\alpha \in \mathbb{R}^2$ — параметр формы. Численные оценки значений параметров модели получены методом максимального правдоподобия:

$$\alpha = \begin{matrix} \text{ОВ} \\ \text{КП} \end{matrix} \begin{pmatrix} 0,0 \\ -2,9 \end{pmatrix}, \quad \Omega = \begin{matrix} \text{ОВ} & \text{КП} \\ \text{ОВ} & \text{КП} \end{matrix} \begin{pmatrix} 0,79 & 0,17 \\ 0,17 & 0,26 \end{pmatrix}.$$

Таким образом, построена математическая модель $MAP(1)C$ (Мультипликативная $AP(1)$ с Сезонностью) стохастического процесса лесопожарных ситуаций, включающая в себя следующие компоненты:

- мультипликативная сезонная компонента $s(\tau)$;
- тренд, описываемый векторной авторегрессионной моделью ВАР(1);
- случайная компонента e_{t_k} , описываемая двумерным скошенным нормальным распределением:

$$\begin{aligned} x_{t_k} &= \exp(y_{t_k} + s(\tau_{t_k})), \\ y_{t_k} &= Ay_{t-1} + e_{t_k}, \\ e_{t_k} &\sim SN_2(\sigma\Omega\sigma^T, \alpha). \end{aligned} \tag{4}$$

4. Предсказание и прогноз

В соответствии с пунктом 5.2.6 ГОСТ Р 22.1.09–99, количество лесных пожаров прогнозируют, исходя из:

- степени пожарной опасности в лесу по условиям погоды;
- класса пожарной опасности лесных участков на рассматриваемой территории;
- количества потенциальных источников огня;
- количества пожаров в ретроспективе в аналогичных условиях;
- теоретических законов распределения случайных событий, которыми удовлетворительно описываются (аппроксимируются) некоторые ряды распределения лесных пожаров во времени и по территории.

В стандарте приведена простая методика прогнозирования пожарной опасности с использованием комплексного показателя Нестерова. В соответствии с этой методикой, значение КП определяет так называемый *класс пожарной опасности*, который характеризует ожидаемый масштаб развития лесопожарной ситуации. Существует общероссийская шкала соответствия значений КП уровню пожарной опасности.²

Описанный способ прогнозирования с использованием значений КП хотя и является очень простым, но не позволяет получать точные и надёжные оценки пожарной опасности. Так, корреляция между значениями КП и ОВ в наиболее пожароопасный период (с 1 мая по 30 сентября в сезон 1990 г.) не превышает 0,5, а в сезон 1989 г. не превышает 0,25 (см. табл. 1).

В настоящей работе для прогнозирования будущих значений ОВ использована предложенная модель MAP(1)С. Значения метеорологических показателей и рассчитываемого на их основе комплексного показателя Нестерова могут быть получены на несколько суток вперёд. Математические модели и методы прогнозирования метеорологических показателей выходят за рамки настоящей работы. Следует отметить, что обычно значения комплексного показателя доступны с приемлемой степенью надёжности не более чем на 3 суток [9]. Поэтому прогнозирование развития лесопожарных ситуаций на больший срок является затруднительным.

Разработан функциональный вычислительный алгоритм предсказания значений ОВ при заданных значениях КП на основе предложенной модели MAP(1)С. Прогнозирование уровня пожарной опасности осуществлено методом стохастического моделирования.

Алгоритм позволяет получить не только ожидаемое значение ОВ, но и, что особенно важно, оценить потенциальный риск развития лесопожарной ситуации, т.е. рассчитать доверительный интервал для заданного уровня достоверности. В настоящей работе для обозначения верхней границы доверительного интервала используется термин *Величина Риска* (*Value at Risk*):

$$\text{VaR}_p = \inf \{x | F(x) \geq p\},$$

где $F(x)$ — функция распределения случайной величины x . Концепция VaR широко используется в финансовой математике для расчёта потенциальных убытков. В случае моделирования лесопожарных ситуаций использование этой концепции также выглядит обоснованным, поскольку экономический урон, наносимый лесными пожарами, также как и затраты на его профилактику и ликвидацию, являются монотонными функциями количества очагов возгорания.

Получены ожидаемые значения уровня пожарной опасности, а также значения VaR для уровней достоверности 90% и 95% в наиболее пожароопасный период с 1 мая по 30 сентября, в сезоны 1989 и 1990 гг. Прогноз производился на 1, 2 и 3 суток вперёд. Значения КП в этот период считались известными заранее. Результаты для $n = 2$ суток приведены на рис. 4.

²Необходимо отметить, что эта шкала не является универсальной и должна быть скорректирована с учётом местных особенностей, причём не вполне ясны принципы корректировки.

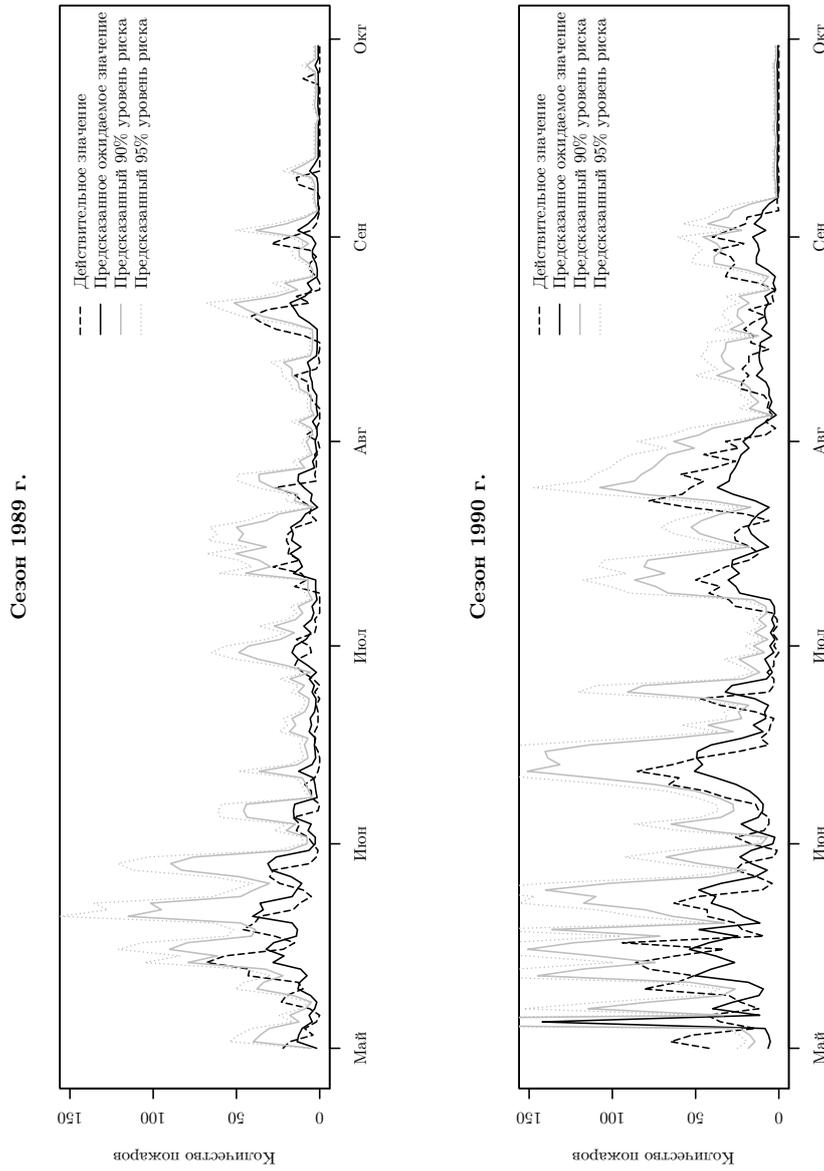


Рис. 4. Прогноз количества пожаров на 2 суток вперёд

Рассчитаны значения ранговой корреляции Спирмена между предсказанными и действительными значениями ОВ. Результаты расчёта приведены в табл. 1. Как видно, прогноз, полученный с использованием модели MAP(1)C, является существенно более близким к действительным значениям, чем прогноз с использованием только показателя Нестерова.

Таблица 1

Численные оценки коэффициента ранговой корреляции Спирмена между действительными и предсказанными значениями количества очагов возгорания

Метод прогноза	1989 г.	1990 г.
Инд. Нестерова	0,21	0,39
MAP(1)C, 1 сутки	0,76	0,81
MAP(1)C, 2 суток	0,56	0,66
MAP(1)C, 3 суток	0,41	0,56

С целью установления степени адекватности рассчитанных значений VaR проведена проверка статистической гипотезы H_0 о равенстве действительного и теоретического количества превышений значения $VaR_{90\%}$. Для прогноза на 1 и 2 суток гипотеза H_0 не может быть отвергнута на уровне значимости 0,01. Для прогноза на 3 суток и более гипотеза H_0 должна быть отвергнута на любом разумном уровне значимости. Таким образом, рекомендуемый максимальный горизонт прогноза количества очагов возгорания с помощью модели MAP(1)C составляет 2 суток.

5. Заключение

В настоящей работе проведен анализ статистических свойств совместной выборки значений КП и ОВ за 1969–1990 гг. Построена трёхуровневая математическая модель MAP(1)C, связывающая исследуемые показатели. Разработан вычислительный алгоритм моделирования будущих значений ОВ, который может быть использован для прогноза развития лесопожарной ситуации на 1–2 суток вперёд.

Предложенная модель MAP(1)C использует значения комплексного показателя Нестерова, но также учитывает динамические свойства процесса лесопожарных ситуаций. Благодаря этому она позволяет получать существенно более точные и надёжные оценки пожарной опасности.

Необходимо отметить, что существует определённый потенциал для совершенствования методики расчёта комплексного метеорологического показателя. Это тем более важно и актуально, поскольку точность оценок и максимальный горизонт прогноза модели MAP(1)C во многом определяются свойствами используемого комплексного показателя.

Литература

- ГОСТ Р 22.1.09–99 «Безопасность в чрезвычайных ситуациях. Мониторинг и прогнозирование лесных пожаров. Общие требования».
- Нестеров В. Г. Горимость леса и методы её определения. — М.: Гослесбумиздат, 1949.
- MET4 and MET4A Calculation of Dew Point. — <http://www.paroscientific.com/dewpoint.htm>.
- Sims C. A. Macroeconomics and Reality // *Econometrica*. — 1980. — № 48. — С. 1–48.
- Brookwell P. J., Davis R. A. Time Series and Forecasting Methods. Second edition. — New York: Springer, 1991.

6. Барановский Н. В. Методика прогнозирования лесной пожарной опасности как основа нового государственного стандарта // Пожарная безопасность. — 2007. — № 4. — С. 80–84.
7. Зденева М. Я., Виноградова М. В. Метод среднесрочного прогноза степени пожарной опасности в лесах по метеорологическим условиям / Методический кабинет Гидрометцентра России. — 2008. — <http://method.hydromet.ru/methods/fire/fire.html>.
8. Azzalini A., Capitanio A. Statistical Applications of the Multivariate Skew-Normal Distribution // Journal of the Royal Statistical Society, B. — 1999. — № 61. — С. 579–602.
9. Толстых М. А. Прогноз осадков по станциям регионов внетропической зоны Северного полушария на срок до 72 ч на основе полулагранжевой модели с постоянным разрешением ПЛАВ-2005 / Методический кабинет Гидрометцентра России. — 2008. — <http://method.hydromet.ru/model/tolstih2/tolstih2.html>.

UDC 519.254.1

Modelling and Forecasting the Evolution of Forest Fire Using Vector Autoregression Processes

A. V. Paramonov

*Center of Mathematics Research, MSUT «STANKIN»
Vadkovskiy lane, 3A, Moscow, Russia, 117936*

The joint sample of daily values of the number of forest fire seats and the complex meteorological index in Irkutsk region, seasons 1969–1988, is investigated. It appears that the evolution of forest fire is well described by a vector autoregression process based model. Computer algorithm of simulation of the future numbers of fire seats is presented. It is shown how the algorithm can be used for reliable forecasting of the evolution of forest fire up to 2 days ahead.