

МОДЕЛЬ СТОХАСТИЧЕСКОГО ПРОЦЕССА ЛЕСОПОЖАРНЫХ СИТУАЦИЙ

Парамонов А. В., Марков П. Н.

Проведено исследование статистических свойств совместной выборки ежедневных значений количества очагов лесных пожаров и комплексного метеорологического показателя в Иркутской области за 1969–1990 гг. Построена математическая модель стохастического процесса развития лесопожарных ситуаций $MAR(1)C$, включающая в себя мультипликативную сезонную компоненту и авторегрессионный процесс $AR(1)$. 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.

Введение. Лесные пожары представляют собой масштабное стихийное природное явление, которое регулярно угрожает жизни людей и наносит серьёзный материальный урон российской экономике. Ежегодно на территории Российской Федерации регистрируются десятки тысяч лесных пожаров (Российский... , 2007). 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.

Несмотря на то, что лесные пожары — это природное явление, подавляющее большинство из них в настоящее время происходит по вине граждан и организаций (Российский... , 2007). Любой лесной пожар должен рассматриваться не только как сложнейшее физико-химическое явление, но и как существенно недетерминированный стохастический процесс. В работе (Валендик и др., 1979) предложена следующая классификация уровней моделирования лесопожарных ситуаций:

- **Фундаментальный уровень:** описание аэротермохимических процессов, происходящих в зоне горения.
- **Тактическое моделирование:** форма, направление и скорость распространения пожара в зависимости от силы и направления ветра, горимости леса, рельефа местности.
- **Стратегическое моделирование:** статистические закономерности возникновения и развития лесопожарных ситуаций и их связь с природными, метеорологическими и прочими факторами.

В настоящей работе развитие лесопожарных ситуаций рассмотрено на стратегическом уровне. Проведено исследование статистических свойств совместной выборки ежедневных значений количества очагов возгорания (далее — ОВ) и комплексного метеорологического показателя (далее — КП) в Иркутской области за 1969–1990 гг.:

$$X_{t_k} = \begin{pmatrix} \text{ОВ}_{t_k} \\ \text{КП}_{t_k} \end{pmatrix}, \quad k = 1 \dots P \cdot K,$$

где $P = 22$ — число сезонов (периодов), $K = 214$ — число наблюдений за сезон (наблюдения проводились с 1 апреля по 31 октября). Данные по первичным метеорологическим показателям (температура воздуха, относительная влажность, количество осадков) от 6 метеостанций (Бодайбо, Ербогачен, Иркутск, Киренск, Нижнеудинск, Жигалово) предоставлены Государственным учреждением «Всероссийский научно-исследовательский институт гидрометеорологической информации — Мировой центр данных». Данные по количеству очагов возгорания предоставлены предприятием «Авиалесоохрана».

Комплексный метеорологический показатель Нестерова. Общие требования по мониторингу и прогнозированию лесных пожаров и чрезвычайных лесопожарных ситуаций регулируются ГОСТ «Безопасность в чрезвычайных ситуациях. Мониторинг и прогнозирование лесных пожаров. Общие требования» (ГОСТ Р 22.1.09–99. . . , 1999). Степень пожарной опасности в лесу по условиям погоды должна определяться по принятому в лесном хозяйстве комплексному показателю В. Г. Нестерова (Нестеров, 1949), который вычисляется на основе данных о температуре воздуха,

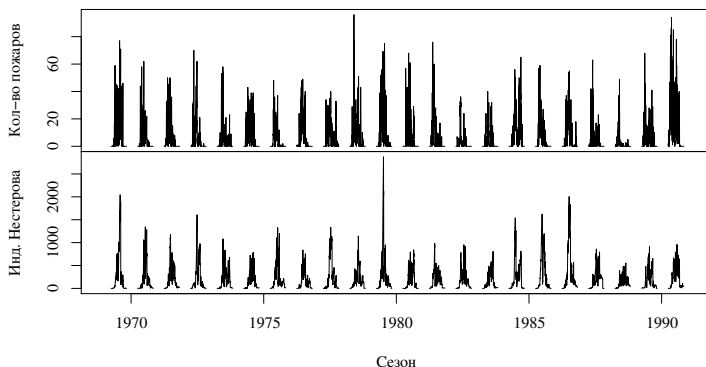


Рис. 1. Исходные данные.

температуре точки росы и количестве выпавших осадков:

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

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). В настоящей работе расчёт значений точки росы произведён на основе имеющихся значений температуры и влажности воздуха по формулам (<http://www.paroscientific.com/dewpoint.htm>):

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

$$\gamma(T, RH) = \frac{a \cdot T}{b + T} + \ln(RH/100),$$

где RH — относительная влажность в процентах, $a = 17.27$, $b = 237.7^\circ\text{C}$.

Значения КП_{t_k} получены путём усреднения 6 значений КП, рассчитанных для каждой из метеостанций:

$$\text{КП}_{t_k} = \frac{1}{6} \left(\text{КП}_{t_k}^{\text{Бодайбо}} + \text{КП}_{t_k}^{\text{Ербогачен}} + \text{КП}_{t_k}^{\text{Иркутск}} + \text{КП}_{t_k}^{\text{Киренск}} + \text{КП}_{t_k}^{\text{Нижнеудинск}} + \text{КП}_{t_k}^{\text{Жигалово}} \right).$$

В (ГОСТ Р 22.1.09–99. . . , 1999) 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.¹

К сожалению, описанный способ прогнозирования не позволяет получать точные и надёжные оценки пожарной опасности. Так, выборочное значение коэффициента ранговой корреляции Спирмена между величинами 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. Таблица 1). 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 (Барановский, 2007; Зденева и Виноградова, 2009). Таким образом, в настоящее время существует реальная необходимость в разработке новых, более точных и надёжных методов прогнозирования пожарной опасности.

Модель MAP(1)C. Для построения математической модели использована подвыборка исходных данных за 1969–1988 гг. Данные за 1989–1990 гг. использованы в дальнейшем для верификации модели.

Сезонная компонента. Исследуемые временные ряды обладают выраженной мультипликативной сезонной компонентой (см. Рис. 1). Переход к аддитивной модели осуществлён путём логарифмирования данных.

В настоящей работе сезонная компонента моделируется с помощью конечного ряда Фурье:

$$s_p(\tau) = \frac{a_0}{2} + \sum_{i=1}^N a_i \cos(2\pi \frac{i\tau}{T}) + b_i \sin(2\pi \frac{i\tau}{T}).$$

Количество базисных функций $2N+1$ выбрано равным 9. Подбор параметров модели a_i, b_i , осуществлён методом наименьших квадратов с допол-

¹Необходимо отметить, что эта шкала не является универсальной и должна быть скорректирована с учётом местных особенностей, причём не вполне ясны принципы корректировки.

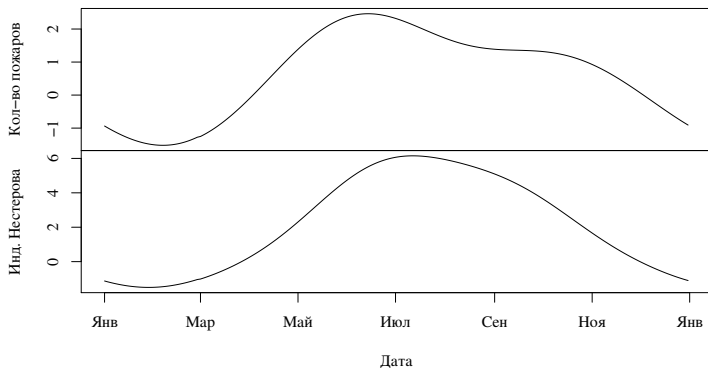


Рис. 2. Сезонная компонента.

нительным условием на гладкость результирующей кривой (Ramsay and Silverman, 2002).

За счёт требования непрерывности, метод позволяет получить более точные оценки значений сезонной компоненты, и даже восстановить значения сезонной компоненты для дат, отсутствующих в исходной выборке (в нашем случае — до 1 апреля и после 31 октября).

Стохастическая компонента. Стохастическая компонента получена путём вычитания сезонной компоненты из логарифмированных данных:

$$Y_{t_k} = \ln(X_{t_k}) - s(\tau_k).$$

Полученные временные ряды не могут быть описаны с помощью простейшей модели независимых одинаково распределённых (н.о.р.) случайных величин. Профиль эмпирической автокорреляционной функции стохастической компоненты количества очагов возгорания приведён на Рис. 4. Эмпирическая автокорреляционная функция существенно отклоняется от нуля и имеет вид затухающей экспоненты. Это указывает на наличие тренда, который может быть описан с помощью авторегрессионной модели первого порядка (Бокс и Дженкинс, 1974). Значения стохастической компоненты КП также не являются н.о.р. величинами, по построению.

Кроме того, временные ряды обладают выраженной статистической связью: коэффициент корреляции совместного распределения значений Y_{t_k} равен 0.41. Рассматриваемый стохастический процесс является существенно многомерным.

Существует несколько моделей авторегрессионного типа, способных описать обнаруженные эмпирические эффекты:

- VARMA(p, q) — векторная модель авторегрессии—скользящего среднего (Sims, 1980):

$$y_{t_k} = (A_1 y_{t_{k-1}} + \dots + A_p y_{t_{k-p}}) + e_{t_k} + (B_1 e_{t_{k-1}} + \dots + B_q e_{t_{k-q}}), \quad (2)$$

где $A_i \in \mathbb{R}^{n \times n}$, $B_i \in \mathbb{R}^{n \times n}$ — матрицы параметров модели, $e_{t_k} \in \mathbb{R}^n$ — вектор остаточных значений в момент времени t_k (случайная компонента), n — размерность вектора стохастической компоненты y_{t_k} .

- ARMAX(p, q) — модель авторегрессии—скользящего среднего с внешними (экзогенными) факторами (Льюнг, 1991):

$$y_{t_k} = \begin{pmatrix} y_{t_k}^* \\ y_{t_k}^{**} \end{pmatrix},$$

$$y_{t_k}^* = (A_1 y_{t_{k-1}}^* + \dots + A_p y_{t_{k-p}}^*) + e_{t_k} + (B_1 e_{t_{k-1}} + \dots + B_q e_{t_{k-q}}) + G y_{t_k}^{**}. \quad (3)$$

где $y_{t_k}^*$ — зависимая переменная, $y_{t_k}^{**}$ — вектор внешних факторов, A_i, B_i, G — параметры модели, e_{t_k} — случайная компонента.

- ARMA(p, q) с внешними регрессорами (линейная регрессия с остатками)

ARMA(p, q), (Бокс и Дженкинс, 1974)):

$$y_{t_k} = \begin{pmatrix} y_{t_k}^* \\ y_{t_k}^{**} \end{pmatrix},$$

$$y_{t_k}^* = z_{t_k} + G y_{t_k}^{**}, \quad (4)$$

$$z_{t_k} = (A_1 z_{t_{k-1}} + \dots + A_p z_{t_{k-p}}) + e_{t_k} + (B_1 e_{t_{k-1}} + \dots + B_q e_{t_{k-q}}),$$

где $y_{t_k}^*$ — зависимая переменная, $y_{t_k}^{**}$ — вектор внешних факторов, A_i, B_i, G — параметры модели, e_{t_k} — случайная компонента.

Во всех случаях предполагается, что значения случайной компоненты представляют собой последовательность н.о.р. величин с нулевым математическим ожиданием.

Выбор модели определяется целью исследования и априорными сведениями о динамической структуре переменных стохастической компоненты. При использовании модели (2) подразумевается, что все входящие в стохастическую компоненту переменные являются взаимосвязанными и однородными по своей динамической структуре. При использовании моделей (3), (4) стохастическая компонента должна быть разделена на переменные двух различных типов: «внутренние» y^* и «внешние» y^{**} . При этом модель фактически описывает только динамическую структуру внутренних переменных. Характер динамической структуры внешних факторов y^{**} не описывается моделью, и может быть произвольным.

Целью моделирования в настоящей работе является краткосрочный прогноз будущих значений ОВ; получение будущих значений КП выходит за рамки исследования. Кроме того, значения $Y_{t_k}^{\text{КП}}$ по определению (1) обладают асимметричной динамической структурой, не свойственной авторегрессионным моделям. Таким образом, в рассматриваемом случае стохастическая компонента естественным образом разбивается на зависимую переменную $y^* = y^{\text{ОВ}}$ и внешний фактор $y^{**} = y^{\text{КП}}$, в соответствии с (3), (4).

В настоящей работе для описания стохастической компоненты использована модель типа (4):

$$y_{t_k}^{\text{ОВ}} = z_{t_k} + G y_{t_k}^{\text{КП}},$$

$$z_{t_k} = A z_{t_{k-1}} + e_{t_k}.$$

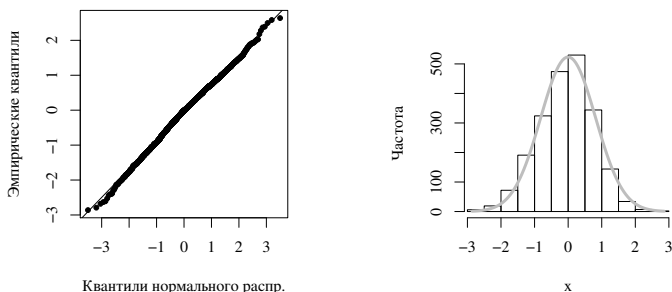


Рис. 3. Эмпирическое распределение значений случайной компоненты ОВ. Справа: квантильный график. Слева: гистограмма (линией обозначена теоретическая частота аппроксимирующего нормального распределения).

Оценки параметров модели получены методом максимального правдоподобия (указан 95% доверительный интервал):

$$A = (0.64 \pm 0.04),$$

$$G = (0.56 \pm 0.08).$$

Случайная компонента. Распределение полученных значений случайной компоненты e_{t_k} является симметричным, легкохвостым и хорошо описывается нормальным законом, особенно для положительных значений (см. Рис. 3). Статистические тесты Андерсона—Дарлингга, Крамера—фон—Мизеса и Шапиро—Франсия (Закс, 1976) требуют принять гипотезу о нормальном распределении положительных значений стохастической компоненты на уровне значимости 0.05. Наблюдаемое расхождение для отрицательных значений обусловлено, по видимому, дискретностью исходных данных.

В качестве оценки параметра масштаба модели использовано выборочное значение среднеквадратичного отклонения (указан 95% доверительный интервал):

$$\sigma = (0.82 \pm 0.03).$$

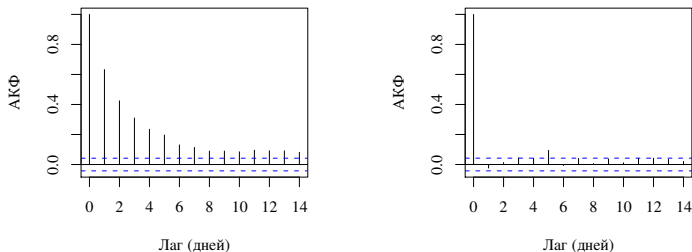


Рис. 4. Автокорреляционная функция стохастической (слева) и случайной (справа) компонент ОВ. Пунктиром обозначен 95% доверительный интервал в предположении о стохастической независимости.

Значения автокорреляционной функции случайной компоненты не отклоняются существенно от нуля (см. Рис. 4). Статистический тест серий (Закс, 1976) показывает, что гипотеза о стохастической независимости значений e_{t_k} не может быть отвергнута на уровне значимости 0.01.

Таким образом, построена параметрическая математическая модель *МАР(1)С* (Мультипликативная *АР(1)* с *Сезонностью*) стохастического процесса лесопожарных ситуаций, включающая в себя следующие компоненты:

- Мультипликативная сезонная компонента $s(\tau)$, описываемая гладкими периодическими функциями;
- Тренд, описываемый авторегрессионной моделью *АР(1)* с внешними регрессорами;
- Случайная компонента e_{t_k} , описываемая нормальным распределением.

$$x_{t_k} = \exp(y_{t_k} + s(\tau_k)), \quad (5)$$

$$y_{t_k}^{\text{ОВ}} = z_{t_k} + G y_{t_k}^{\text{КП}}, \quad (6)$$

$$z_{t_k} = A z_{t_{k-1}} + e_{t_k},$$

$$e_{t_k} \sim N(0, \sigma). \quad (7)$$

Предсказание и прогноз. В настоящей работе термины «предсказание» и «прогноз» используются в следующих смыслах:

Предсказание — получение точечной оценки наиболее вероятного развития лесопожарной ситуации.

Прогноз — построение доверительного интервала для возможных сценариев развития лесопожарной ситуации, т.е. оценивание потенциального риска, на заданном уровне достоверности.

Умение строить прогноз особенно важно в случае существенно недетерминированного, стохастического характера исследуемых величин, каковым является поведение лесных пожаров.

Для обозначения верхней границы доверительного интервала в настоящей работе используется термин *Величина Риска* (*Value at Risk*, (Jorion, 2006)):

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

где $F(x)$ — функция распределения случайной величины x . Концепция VaR широко используется в финансовой математике для расчёта потенциальных убытков (International. . . , 2006; Ширяев, 2007). 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, резонно считать монотонными функциями количества очагов возгорания.

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 (Толстых, 2008). 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)C. В настоящей работе прогноз уровня пожарной опасности осуществлен методом стохастического моделирования.

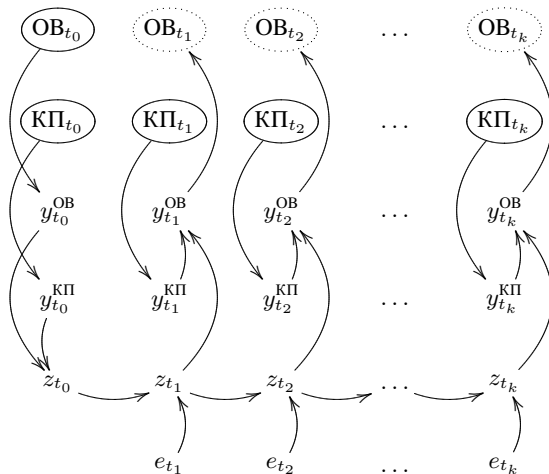


Рис. 5. Моделирование количества очагов возгорания при заданных значениях комплексного показателя. Схема расчёта. Сплошной рамкой обозначены входные данные, пунктирной рамкой — результат расчёта.

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.

Рассчитаны значения ранговой корреляции Спирмена между предсказанными и действительными значениями ОБ. Результаты расчёта приведены в Таблица 1. Как видно, прогноз, полученный с использованием модели MAP(1)С, является существенно более близким к действительным значениям, чем прогноз с использованием «чистого» показателя Нестерова.

С целью установления степени адекватности рассчитанных значений VaR проведена проверка статистической гипотезы H_0 о равенстве действительного и теоретического количества превышений значения $VaR_{90\%}$ («безусловное покрытие»). Для прогноза на 1 и 2 суток гипотеза H_0 не может быть отвергнута на уровне значимости 0.01. Для прогноза на 3 и более суток гипотеза H_0 должна быть отвергнута на этом уровне значи-

Таблица 1. Численные оценки коэффициента ранговой корреляции Спирмена между действительным и предсказанными значениями количества очагов возгорания.

Метод прогноза	Сезон 1989 г.	Сезон 1990 г.
Инд. Нестерова	0.21	0.39
МАР(1)С, 1 сутки	0.75	0.80
МАР(1)С, 2 суток	0.56	0.67
МАР(1)С, 3 суток	0.45	0.59

мости. Таким образом, рекомендуемый максимальный горизонт прогноза количества очагов возгорания с помощью модели МАР(1)С составляет 2 суток.

Заключение. В настоящей работе проведен анализ статистических свойств совместной выборки значений КП и ОВ за 1969–1990 гг. Построена трёхуровневая параметрическая математическая модель МАР(1)С, 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 МАР(1)С 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.

Следует отметить, что полученные результаты не отменяют необходимость совершенствования методики расчёта комплексного метеорологического показателя. Это тем более важно и актуально, поскольку точность оценок и максимальный горизонт прогноза модели МАР(1)С во многом определяются свойствами используемого комплексного показателя.

СПИСОК ЛИТЕРАТУРЫ

- Барановский Н. В.* Методика прогнозирования лесной пожарной опасности как основа нового государственного стандарта // *Пожарная безопасность.* — 2007. — № 4. — С. 80–84.
- Бокс Дж., Дженкинс Г.* Анализ временных рядов. Прогноз и управление. — Москва: Мир, 1974.

Валендик Э. Н., Матвеев П. М., Софронов М. А. Крупные лесные пожары и борьба с ними. — Москва: Наука, 1979.

ГОСТ Р 22.1.09–99 «Безопасность в чрезвычайных ситуациях. Мониторинг и прогнозирование лесных пожаров. Общие требования». — Москва: Изд-во стандартов, 1999.

Закс Л. Статистическое оценивание. — Москва: Статистика, 1976.

Зденева М. Я., Виноградова М. В. Метод среднесрочного прогноза степени пожарной опасности в лесах по метеорологическим условиям // *Метеорология и гидрология*. — 2009. — № 1. — С. 16–26.

Льюнг Л. Идентификация систем. Теория для пользователя: Пер. с англ., Цыпкин Я. З. (ред.). — Москва: Наука, 1991.

Нестеров В. Г. Горимость леса и методы её определения. — Москва: Гослесбумиздат, 1949.

Российский статистический ежегодник. 2007, Гос. ком. Рос. Федерации по статистике. — Москва: Госкомстат России, 2007.

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.

Ширяев В. И. Модели финансовых рынков. Оптимальные портфели, управление финансами и рисками. — Москва: УРСС, 2007.

International Convergence of Capital Measurement and Capital Standards: A Revised Framework — Comprehensive Version, Basel Committee on Banking Supervision. — 2006. — <http://www.bis.org/publ/bcbs128.pdf>.

Jorion P. Value at Risk: The New Benchmark for Managing Financial Risk, 3rd edition. — New York: McGraw-Hill, 2006.

MET4 and MET4A Calculation of Dew Point.
<http://www.paroscientific.com/dewpoint.htm>.

Ramsay J. O., Silverman B. W. Applied Functional Data Analysis. — New York: Springer, 2002.

Sims C. A. Macroeconomics and Reality // *Econometrica*. — 1980. — no. 48. — Pp. 1–48.

A MODEL OF THE FOREST FIRE STOCHASTIC PROCESS

Paramonov A. V., Markov P. N.

*The joint sample of daily values of the number of forest fire seats and the complex meteorological index in Irkutsk region, seasons 1969–1990, is investigated. The stochastic process of forest fire evolution is well described by proposed **MAR(1)S** model, which includes multiplicative seasonal component and AR(1) autoregressive process. Computer algorithm for the 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.*