

Б.М. Пранов
(РАНХиГС при Президенте РФ; e-mail: boris.pranov@gmail.com)

ОБ ОДНОМ МЕТОДЕ МОДЕЛИРОВАНИЯ ЕВРОПЕЙСКОЙ ПОЖАРНОЙ СТАТИСТИКИ

Предлагается метод построения аппроксимирующих математических зависимостей для пожарной статистики ряда европейских стран с использованием производственной функции Кобба-Дугласа.

Ключевые слова: временной ряд, аппроксимация, коэффициент корреляции.

В.М. Pranov
**ON A METHOD FOR MODELING
THE EUROPEAN FIRE STATISTICS**

A method of constructing approximate mathematical relationships for fire statistics of several European countries using the Cobb-Douglas production function is offered.

Key words: time series, approximation, the correlation coefficient.

Статья поступила в редакцию Интернет-журнала 12 апреля 2015 г.

В работах [1, 2] автором предложен метод построения временных рядов пожарной статистики с использованием производственной функции Кобба-Дугласа. Обычно аппроксимация временных рядов пожарной статистики осуществляется с использованием *линейных* однофакторных или многофакторных *математических зависимостей* вида

$$y_i = a_0 + a_1 x_{i1} + \dots + a_k x_{ik} + \varepsilon_i, \quad (1)$$

где i – номер момента времени;

x_1, \dots, x_k – параметры оценки временного ряда y_i ;

a_0, a_1, \dots, a_k – оцениваемые коэффициенты.

Используются также *нелинейные математические модели*, где в правой части зависимости (1) помещают, помимо первых степеней некоторых параметров x_i , их вторые степени и, возможно, попарные произведения.

Однако, при взгляде на график общего числа пожаров, произошедших в какой-либо конкретной стране (России, США или европейской) на протяжении ряда лет, заметно, что линия графика совершает неперiodическое волнообразное движение и не может быть удовлетворительно описана линейной зависимостью.

В такой ситуации для более точной аппроксимации статистических данных следует использовать нелинейные зависимости. Одной из наиболее популярных моделей, используемых в экономике на протяжении уже более 100 лет, является функция Кобба-Дугласа, моделирующего величину *валового внутреннего продукта (ВВП)* в зависимости от двух факторов – численности работников и оценки основных фондов объекта (страны, региона, предприятия).

Обычно производственную функцию используют в виде зависимости

$$Y = A \cdot X_1^\alpha \cdot X_2^\beta, \quad (2)$$

где Y – аппроксимируемая величина (обычно это ВВП – валовой внутренний продукт);

X_1 – стоимость основных фондов (страны, региона, предприятия);

X_2 – численность работников (также используется эквивалентная величина – фонд заработной платы);

α и β – коэффициенты эластичности;

A – коэффициент размерности.

Таким образом, в зависимости (2) под величиной Y понимается продукция объекта – страны, региона, предприятия. Практика почти столетнего моделирования показывает, что это соотношение удовлетворительно описывает зависимость выпускаемой продукции как для макро-, так и для микроэкономического моделирования. Можно сделать предположение, что общее число пожаров, гибель людей на пожарах, получение травм на пожарах, ущерб от пожаров можно рассматривать как своего рода "продукцию" человеческого сообщества при затратах как материальных, так и человеческих ресурсов. В работах [1, 2] зависимость (2) была применена к аппроксимации общего числа пожаров в России и США. Оказалось, что самым удивительным образом зависимость оказалась неожиданно точной – для России коэффициент корреляции с фактическими данными составляет 0,994, а для США – 0,834.

В связи с этим работа по моделированию временных рядов общего числа пожаров была продолжена для стран европейского сообщества. В число этих стран были включены: Австрия, Великобритания, Венгрия, Германия, Италия, Польша, Румыния, Франция, Чехия и Швеция. При отыскании и первоначальном анализе пожарной статистики этих стран были отмечены следующие особенности, подчас весьма странные:

1. Для некоторых стран пожарная статистика открыта лишь до определенного периода, а для последующего ряда лет эта статистика предлагается уже только за определенную плату; это обстоятельство является удивительным, так как пожарная статистика примыкает к моральной статистике и она открыта для многих, но как теперь оказывается – не для всех стран; так, для Германии – до 2007 года, далее только за плату;

2. Очертания гистограммы временного ряда для числа всех пожаров для всех стран имеют характерный волнообразный вид; однако для некоторых стран официальная статистика допускает необоснованные выбросы; например, в Австрии статистические данные за 2010, 2011 и 2012 годы составляет 34364, 57994, 42213 соответственно; средняя цифра почти в два раза больше предыдущей; возможно, это объясняется новыми нормативными документами по учёту пожаров;

3. Для ряда европейских стран – Бельгии, Испании – в открытом доступе нет данных пожарной статистики (общего числа пожаров, ущерба от пожаров и т.д.); приводятся только данные о лесных пожарах.

Далее, для большинства европейских стран отсутствуют статистические данные о стоимости основных фондов. Любопытно, что для России и США эти данные за значительный промежуток времени имеются в открытом доступе [2]. Для других стран имеются лишь данные о *валовом внутреннем продукте (ВВП)* за период 1980–2013 гг. [3]. В связи с такой ситуацией автором было принято решение воспользоваться зависимостью Кобба-Дугласа, взяв взамен параметра X_1 (стоимости основных фондов) соответствующее значение ВВП. Разумеется, такая замена не равноценна, тем не менее следует помнить, что эти два параметра (стоимость основных фондов и ВВП) связаны достаточно точной зависимостью Кобба-Дугласа и поэтому с использованием этой зависимости первый из них можно выразить через второй и использовать в нашей зависимости. Заметим, что при такой замене в качестве множителя появляется численность населения, но мы её и так используем.

В соответствии со сказанным, для упомянутых выше стран Европы в качестве основной зависимости автор использует аппроксимирующую

$$Y = A \cdot X_1^\alpha \cdot X_2^\beta, \quad (3)$$

где Y – аппроксимируемая величина (общее число пожаров);

X_1 – ВВП – валовой внутренний продукт;

X_2 – численность населения страны;

α и β – коэффициенты эластичности;

A – коэффициент размерности.

Дальнейшее исследование будет заключаться в следующем:

1. Для упомянутых ранее стран Европы производится расчёт коэффициентов α и β по зависимости (3);

2. Для полученной таким образом таблицы вычисленных коэффициентов A , α и β строится модель кластеризации, то есть вся совокупность исследуемых стран разбивается на группы (кластеры), внутри которых значения этих коэффициентов оказываются близкими (относительно некоторой метрики);

3. К группе коэффициентов для 10 стран добавляются ещё Россия и США, для которых расчёты были проведены в работе [2], и проводится кластеризация для 12 стран;

4. Делаем выводы.

Результаты аппроксимации временных рядов пожарной статистики по зависимости (3) существенно отличаются как по степени приближения (измеряемой коэффициентом корреляции между фактическими данными и вычисленными по зависимости), так и по качеству приближения (которое измеряется F -критерием). Приведём в качестве примеров несколько расчётов.

Для Франции имеются в наличии все данные об общем числе пожаров за исследуемый промежуток наблюдения (2002-2013 гг.). Результаты сравнения фактических и расчётных данных представлены на рис. 1.

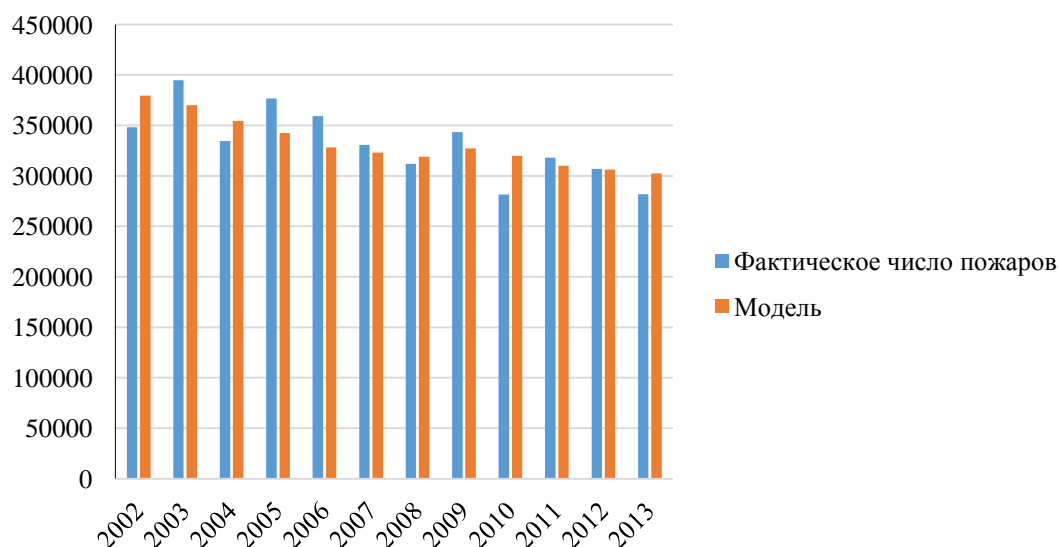


Рис. 1. Результаты сравнения фактических и расчётных данных для Франции

Вычисления показали, что коэффициент корреляции между фактическими данными и вычисленными по зависимости равен 0,718.

Для Германии имеются в наличии данные об общем числе пожаров только за период 2001-2007 гг. Сравнение фактических и расчётных данных представлено на рис. 2.

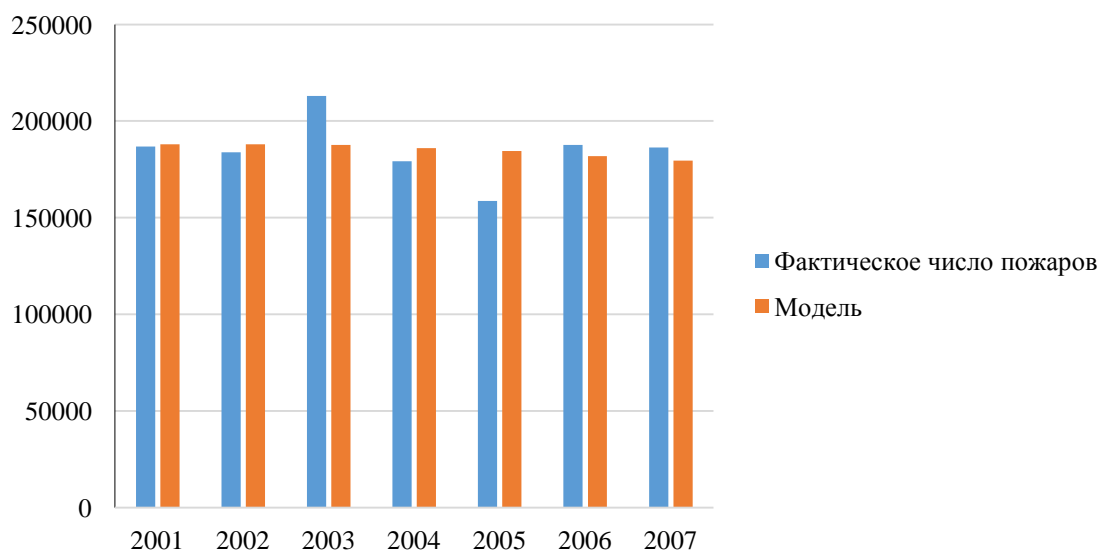


Рис. 2. Результаты сравнения фактических расчётных данных для Германии

Коэффициент корреляции между фактическими и расчётными данными является весьма неудовлетворительным: он равен 0,208.

Для Великобритании имеются в наличии все данные об общем числе пожаров за исследуемый промежуток наблюдения (2002-2013 гг.). Сравнение фактических и расчётных данных и представлено на рис. 3.

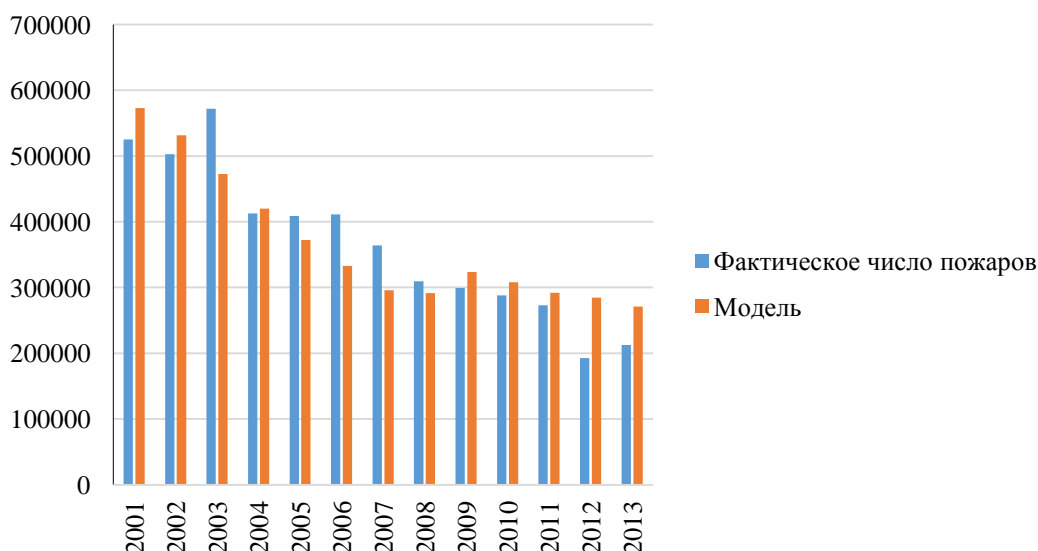


Рис. 3. Результаты сравнения фактических и расчётных данных для Великобритании

Коэффициент корреляции между фактическими и расчётными данными равен 0,878.

Перейдем теперь к разделению стран Европы на кластеры, содержащие страны, наиболее близкие в смысле некоторой метрики. В данном случае в качестве такой метрики взят квадрат евклидовой метрики. Как показали дальнейшие расчёты (их методика описана в справочной документации SPSS), разумно считать, что всю совокупность следует разбить на три кластера. Полученное распределение по кластерам представлено на рис. 4.

Принадлежность к кластерам

Наблюдение	Кластеры 3
1: Австрия	1
2: Великобритания	2
3: Венгрия	1
4: Германия	1
5: Италия	1
6: Польша	1
7: Румыния	3
8: Франция	1
9: Чехия	1
10: Швеция	1

Рис. 4. Распределение стран Европы по кластерам

Соответствующая дендрограмма представлена на рис 5.

Следует отметить, что дендрограмма создается в процессе формирования кластеров, так что видно, что на уровне 5 (по горизонтали, отмечено сверху), сформированы три кластера, к одному из них относится Великобритания, ко второму – Румыния, а к третьему – все оставшиеся 8 стран.

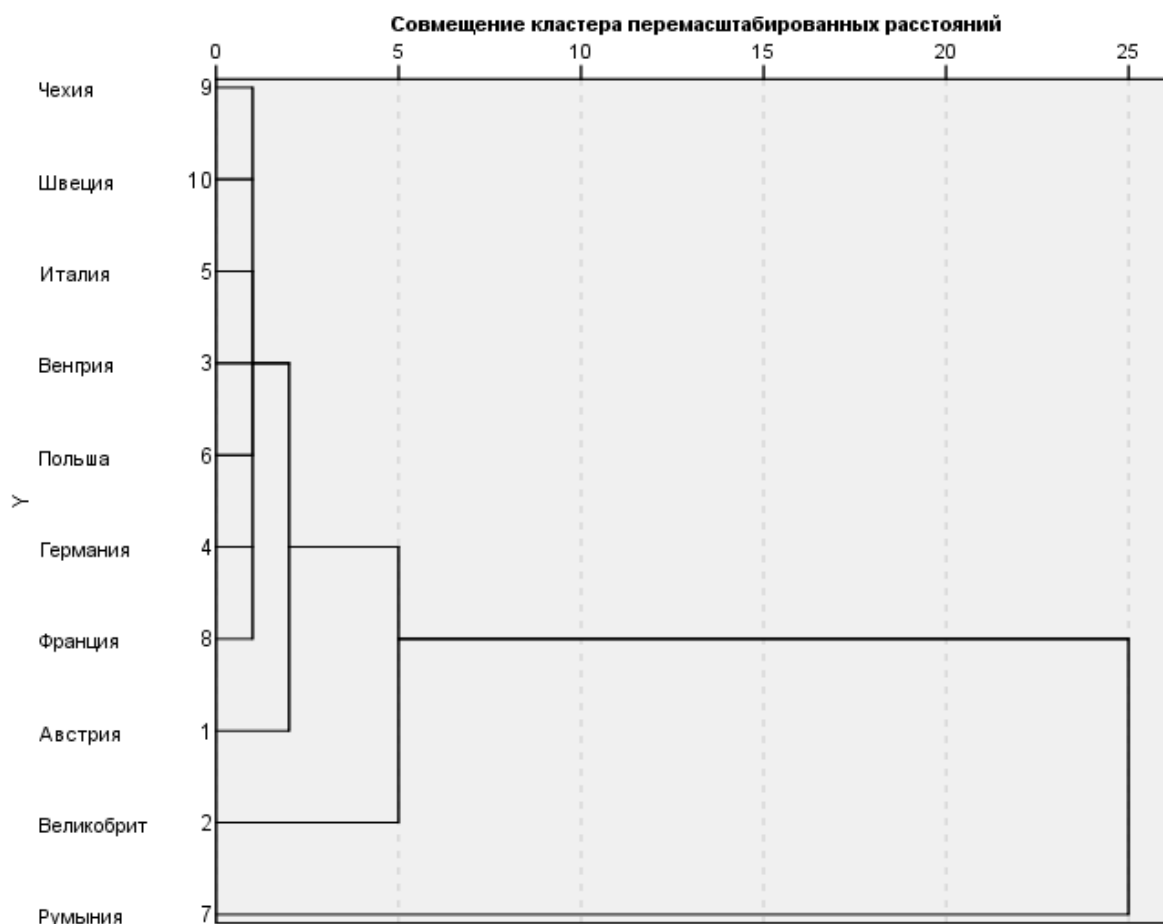


Рис. 5. Дендрограмма с использованием метода межгрупповых связей

При добавлении к 10 рассмотренным странам Европы ещё России и США получается набор из 12 стран, расчётные данные по зависимости (3) представлены в табл. 1.

Таблица 1

Расчётные коэффициенты зависимостей

№	Страна	A0	A1	A2
1	Россия	292,195	-0,251	0,098
2	США	94731,820	1,889	-3,896
3	Австрия	0,61	0,912	0,888
4	Великобритания	602165,259	1,401	-2,096
5	Венгрия	1,245	1,275	-0,379
6	Германия	34,551	0,896	-0,199
7	Италия	7221,955	0,108	0,299
8	Польша	0,084	1,386	-0,021
9	Румыния	0,229	-1,403	4,586
10	Франция	124680,103	0,698	-0,887
11	Чехия	10747,667	0,151	-0,137
12	Швеция	6278,208	0,316	-0,244

Результаты кластерного анализа группы из 12 стран представлены на рис. 6 и 7.

Принадлежность к кластерам

Наблюдение	Кластеры 3
1: Россия	1
2: США	2
3: Австрия	1
4: Великобрит	2
5: Венгрия	1
6: Германия	1
7: Италия	1
8: Польша	1
9: Румыния	3
10: Франция	1
11: Чехия	1
12: Швеция	1

Рис. 6. Распределение по кластерам

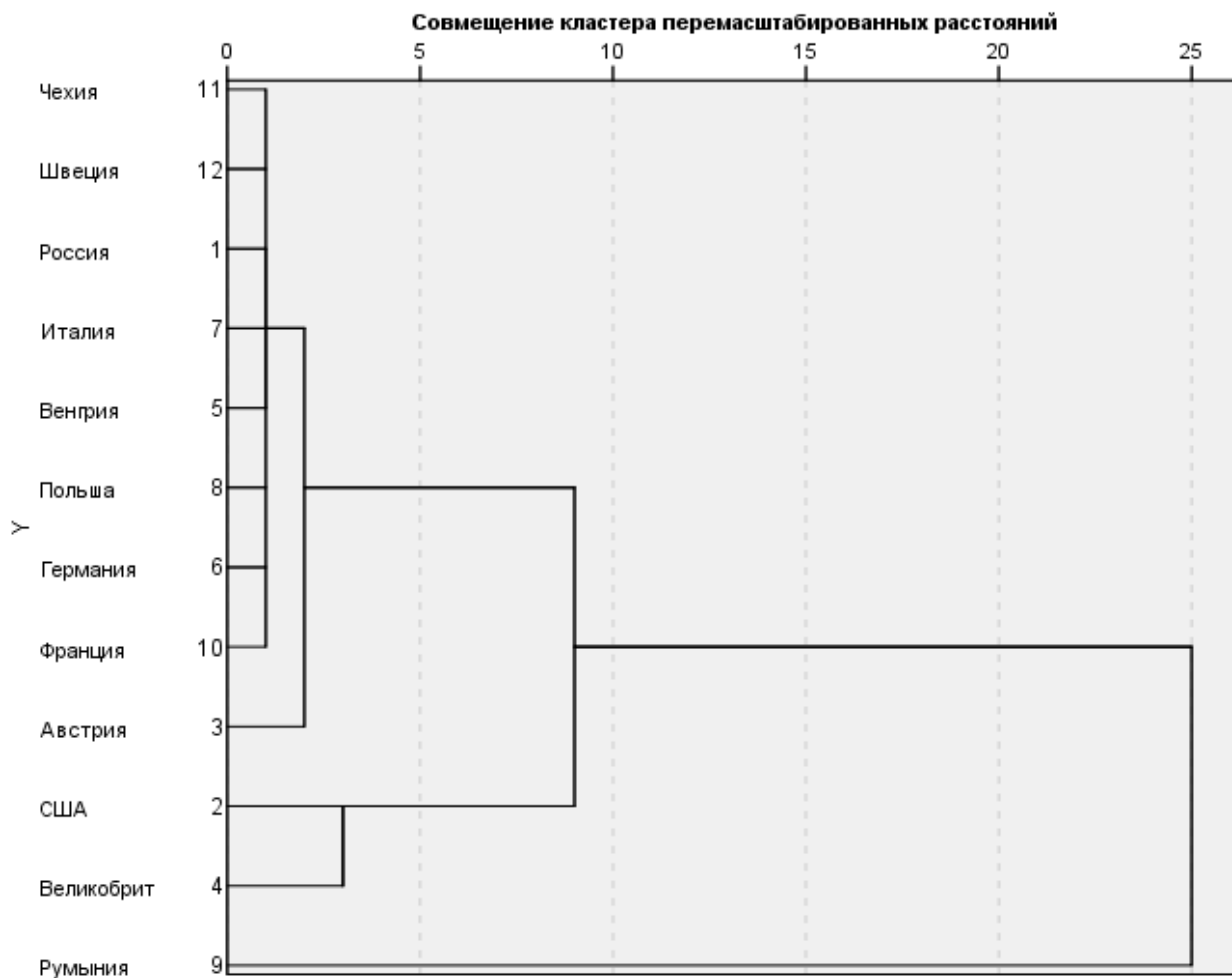


Рис. 7. Дендрограмма с использованием метода межгрупповых связей

Из рис. 6 и 7 можно сделать следующее заключение: Россия отнесена в самый большой кластер Европейских стран, Румыния с её оригинальной статистикой – в особый кластер, а США и Великобритания образуют отдельный от других стран кластер. С точки зрения проведенной кластеризации, можно сделать вывод, что взаимодействие экономики и пожарной охраны имеют: 1) сходный характер для Великобритании и США и отличающийся для других стран; 2) сходный характер для Австрии, Венгрии, Германии, Италии, Польши, России, Франции, Чехии и Швеции; 3) отличный от всех остальных стран – для Румынии.

Для более подробного анализа необходимы дальнейшие исследования.

Литература

1. **Пранов Б.М.** О некоторых аспектах моделирования и прогнозирования временных рядов пожаров // Матер. 23-й междунар. науч.-техн. конф. "Системы безопасности – 2014". М.: Академия ГПС МЧС России, 2014. С. 22-25.
2. **Пранов Б.М.** О некоторых подходах к моделированию и прогнозированию временных рядов пожарной статистики // Технологии техносферной безопасности. Вып. № 5 (57). 2014. С. 209-213. <http://ipb.mos.ru/ttb>.
3. <http://svspb.net/danmark/vvp-stran.php>.