

УДК 004.942

ИССЛЕДОВАНИЕ ТЕНДЕЦИИ ВЫДАВАЕМЫХ ПАТЕНТОВ В РОССИИ НА ОСНОВЕ МНОГОФАКТОРНЫХ МОДЕЛЕЙ

Михалев Павел Владимирович

Студент;

Березниковский филиал ФГБОУ ВО «Пермский национальный исследовательский политехнический университет»;

Российская Федерация, Пермский край, г. Березники - ул. Тельмана, 7;

e-mail: pavelmi@list.ru.

Доказана значимость исследования влияния различных факторов на количество выдаваемых патентов на разработки и исследования в Российской Федерации. Выбран вид модели и построена линейная многофакторная модель, которая описывает тенденции развитию выдачи патентов и влияние на нее таких факторов, как количество населения, занятых научными разработками, расходы на научную отрасль и количество организаций, занятых научными разработками. Выполнен прогноз системы в будущем.

Ключевые слова: математическое моделирование, прогнозирование, патент, наука, исследование, разработка.

INVESTIGATION OF THE TRENDS OF ISSUING PATENTS IN RUSSIA BASED ON MULTIFACTOR MODELS

Mihalev Pavel

Student;

Berezniki branch of «Perm national research Polytechnic University»;

Russian Federation, Perm region, Berezniki - St. Thalmann, 7;

e-mail: pavelmi@list.ru.

The proven significance of the study the influence of various factors on the number of patents granted on research and development in the Russian Federation. The selected view of the model and constructed a multivariate linear model that describes the development trend of the granting of patents and the influence of such factors as the number of people employed in research and development expenses scientific sector and number of organizations engaged in scientific developments. The prognosis of the system in the future.

Keywords: mathematical modeling, forecasting, patent, science, research, development.

Введение

Количество выдаваемых патентов на изобретения и исследования в России отражает её общий уровень образования и квалификацию российских ученых и изобретателей, а также показывает уровень подготовки ВУЗами страны квалифицированных кадров, занимающихся исследованиями и разработками. Развитие науки в нашей стране безусловно является важной задачей, поскольку новые открытия могут влиять на все сферы общества. «Следует сосредоточить повышенное внимание на развитии фундаментальной науки, нацелить её на получение принципиально новых знаний, поиск ответов на так называемые большие вызовы завтрашнего дня» – отмечено на заседании Совета по науке и образованию при Президенте РФ от 23/11/2016 [1]. В связи с этим, исследование и прогнозирование дальнейшего развития науки, а именно количества выдаваемых патентов в России, приобретает важное значение.

Целью данной работы является построение прогнозной модели количества выдаваемых патентов (y) на ближайшее время и выявление факторов, влияющих на нее.

1. Выбор факторов

Выберем факторы, которые могут оказывать влияние на выдачу патентов, из числа тех, что есть в открытом доступе на электронном ресурсе «Прогноз – системы бизнес-аналитики» [2]. В качестве факторов выберем количество поданных заявок на патенты (x_1), численность персонала, занятого исследованиями и разработками (x_2), исследователи, занятые исследованиями и разработками (x_3), исследователи с учеными степенями (x_4), внутренние затраты на исследования и разработки (x_5 , млн руб.), количество вновь зарегистрированных предприятий и организаций, занятых исследованиями и разработками (x_6) и количество ликвидированных подобных предприятий и организаций (x_7). Управляемым фактором будем считать внутренне затраты на исследования и разработки (x_5), т.к. государство может это регулировать. Годовые ряды факторов и критерия приведены в таблице (Таблица 1).

Выбор факторов не противоречит здравому смыслу, поскольку, например, количество выдаваемых патентов в первую очередь зависит от количества поданных заявок на них, также это напрямую зависит от количества занятого персонала исследованиями и разработками и его квалификации, зависит от финансирования исследований и разработок, а также от текущего количества организаций и предприятий, занимающихся этим.

Таблица 1. Реакция и факторы развития выдачи патентов

Критерий и факторы / год	2005	2006	2007	2008	2009
Номер года	1	2	3	4	5
y	19 447	19138	18431	22260	26294
x_1	23644	27884	27505	27712	25598
x_2	813207	807066	801135	761252	742433
x_3	391121	388939	392849	375804	369237
x_4	99428	99507	103725	101049	101275
x_5	230785,2	288805,2	371080,3	431073,2	485834,3
x_6	2707	2628	2543	2492	2234
x_7	668	4704	2516	1189	1412

Таблица 1. (Окончание)

Критерий и факторы / год	2010	2011	2012	2013
Номер года	6	7	8	9
y	21627	20339	22481	21378
x_1	28722	26495	28701	28765
x_2	736540	735273	726318	727029
x_3	368915	374746	372620	369015
x_4	105114	109493	109330	108248
x_5	523377,2	610426,7	699869,8	749797,6
x_6	3201	3517	3732	2655
x_7	1474	2240	2645	2940

2. Нормирование и корреляция

Так как практически все социально-экономические показатели формируются благодаря воздействию множества факторов, то и модель, прогнозирующая их, также должна учитывать это – то есть быть многофакторной. Основными в моделировании и прогнозировании являются:

- Линейно-многофакторные модели (ЛММ) [3].
- Модель пространства состояний (МПС) дает возможность использовать в исходной модели большой спектр процедур, содержащих оценивание и прогнозирование [4].
- Авторегрессионные модели (АВРМ). Их основа заключается в том, что прогнозирование показателя не зависит от факторов. Оно зависит лишь от самого себя за прошлый период [5].

Прогнозирование, прежде всего, предполагает исследование тенденции факторов, а после уже прогнозирование этих факторов на основе обнаруженной тенденции.

Для начала нормируем все факторы и критерии для исключения влияния размерности по формуле:

$$\tilde{y} = \frac{y(t) - \min_t y(t)}{\max_t y(t) - \min_t y(t)},$$

где $\min_t y(t)$ – минимум из критериев y ; $\max_t y(t)$ – максимум из критерия y . Аналогично нормируются и факторы. В конечном итоге получим, что $\langle \tilde{x}, \tilde{y} \rangle \in [0,1]$.

Перед началом моделирования, обычно, проводят анализ выбранных факторов на их взаимную корреляцию, поскольку факторы с высокой взаимной корреляцией не позволят определить их изолированное влияние на результат [6], потому они подлежат исключению из модели. Корреляция рассчитывается по формуле:

$$r_{x,y} = \frac{\sum ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}},$$

где \bar{x} – среднее значение фактора, \bar{y} – среднее значение критерия.

Таблица 2. Парная корреляция факторов и критерия

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1	-0,4772	-0,4047	0,5036	0,5793	0,3087	0,5253
x_2		1	0,9526	-0,7627	-0,9267	-0,4672	0,1108
x_3			1	-0,5423	-0,8009	-0,2728	0,1925
x_4				1	0,8959	0,7438	0,1444
x_5					1	0,5116	0,1078
x_6						1	0,0886
y	-0,0060	-0,6317	-0,7600	0,0551	0,3998	-0,1349	-0,3546

Корреляционный анализ показал, что среди факторов нет взаимно коррелирующих показателей, потому ни один из них не подлежит исключению из модели.

3. Построение моделей

Перейдем к построению ЛММ вида:

$$y_{calc}(t_i) = a_0 + \sum_j a_j \tilde{x}_j(t_i),$$

где a_0 – независимый коэффициент, a_i – коэффициенты влияния i -х факторов $x_i(t)$ в момент времени (номер года) t на значение критерия.

Чтобы определить данные коэффициенты, необходимо минимизировать квадратичное отклонение исходных данных от рассчитанных по формуле:

$$s = \sum (\tilde{y}(t_i) - y_{calc}(t_i))^2 \rightarrow \min.$$

Для минимизации используем «поиск решения» *MS Excel*. Результатом получили коэффициенты ЛММ $a_0 = 1,771$, $a_1 = -0,269$, $a_2 = -3,339$, $a_3 = 1,832$, $a_4 = -1,636$, $a_5 = 0,286$, $a_6 = -0,025$, $a_7 = 0,114$. Квадратичная погрешность в данном случае составляет $s = 0,03051$. Полученные коэффициенты говорят о том, что наиболее значащими факторами являются x_2 , x_3 и x_4 . Значит, наиболее высокое значение в данной модели имеет численность персонала и исследователей, занятых разработками и количество исследователей с учеными степенями соответственно, причем x_2 и x_4 факторы влияют на систему отрицательно. А вот количество заявок на патенты (x_1) и внутренние затраты на разработки (x_5) влияют на критерий крайне слабо. Судя по полученным коэффициентам, можно сказать, что модель не совсем логична.

Попробуем применить АвРМ, которые имеют вид:

$$y_{calc}(t_i) = a_0 + \sum_j^N a_j y_{calc}(t_{i-j}),$$

где a_0 – независимый коэффициент, a_j – коэффициенты влияния i - j расчетного значения критерия системы, N – порядок.

Аналогично найдем коэффициенты, используя «поиск решения» *MS Excel*. Для модели 1-го порядка получили $a_0 = 0,160$, $a_1 = 0,679$ и квадратичную погрешность $s = 0,542$ (рис. 1); для модели 2-го порядка $a_0 = 0,437$, $a_1 = -1,598$, $a_2 = 1,853$ и квадратичную погрешность $s = 0,524$; для модели 3-го порядка $a_0 = 0,370$, $a_1 = -0,237$, $a_2 = -0,490$, $a_3 = 1,053$ и квадратичную погрешность $s = 0,405$. В результате видим (рис. 1), что данный вид модели крайне плохо аппроксимирует наши данные.

Следующая МПС имеет вид:

$$\begin{cases} \vec{x}(t_i) = \vec{a} + B\vec{x}(t_{i-1}) \\ y(t_i) = c + \vec{d}\vec{x}(t_i) \end{cases},$$

где x' – вектор состояния; \vec{d} – матрица выхода, B – матрица перехода.

Также находим коэффициенты, используя «поиск решения» *MS Excel*. Квадратичная погрешность аппроксимации получилась $s = 0,009469$.

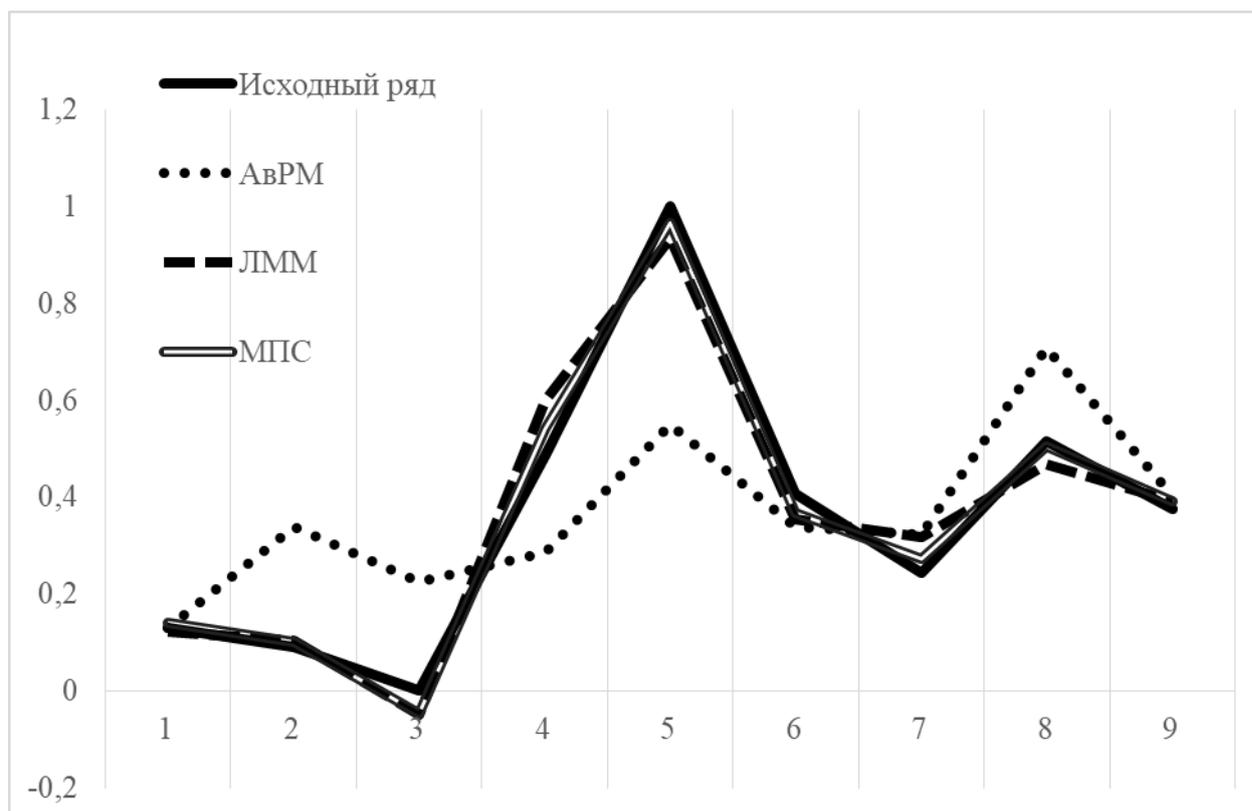


Рис. 1. Анализ приближения моделей к $y(t)$

Таблица 3. Погрешности аппроксимации всех моделей

	ЛММ	АВРМ 1-го порядка	АВРМ 2-го порядка	АВРМ 3-го порядка	МПС
s	0,03051	0,542	0,524	0,405	0,009469

Поскольку целью текущей работы является исследование возможности увеличения количества выдаваемых патентов в будущем, путем поддержки принятия решений, нас также интересуют прогнозные свойства нашей модели, а не только погрешность аппроксимации. Чтобы понять, насколько точно модель может давать прогноз, применим распространённый метод постпрогноза [Ошибка! Источник ссылки не найден.], основанный на расчете реакции системы при исходных факторах на протяжении последних лет. В одной из работ [Ошибка! Источник ссылки не найден.] увеличение интервала дает возможность определить горизонт прогнозирования.

Для всех вышеперечисленных моделей выполнен расчет постпрогноза на один, два и три года.

Таблица 4. Постпрогноз по ЛММ и МПС на три года

Номер года	7	8	9
$u_{исх}$	0,242	0,515	0,374
$u_{расч\ ЛММ}$	0,429	0,341	0,858
$u_{расч\ МПС}$	-0,306	-0,837	-1,449

Получена следующая погрешность постпрогноза разных моделей в зависимости от выбранного интервала.

Таблица 5. Погрешность постпрогноза разных моделей

Модель	Постпрогноз на 1 год	Постпрогноз на 2 год	Постпрогноз на 3 год
ЛММ	77%	-33%	129%
МПС	-226%	-262%	-487%

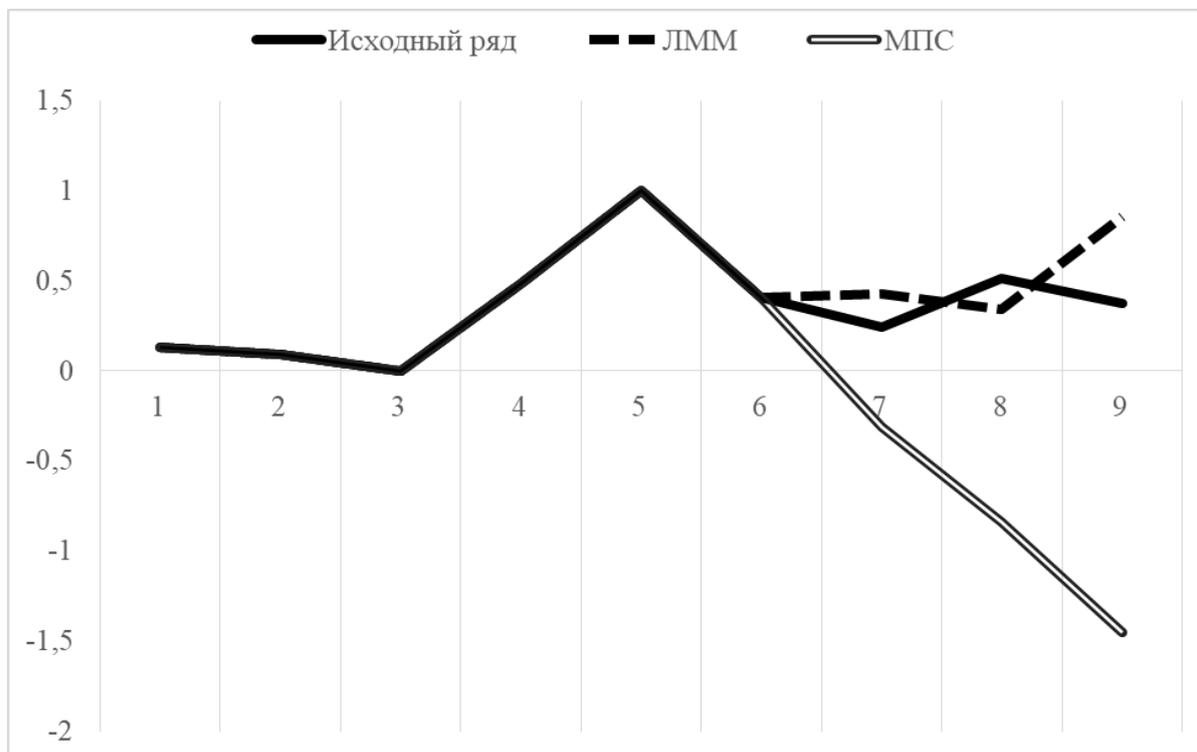


Рис. 2. График постпрогнозов на три года по всем моделям

По графику видно, что постпрогноз ЛММ соответствует исходным данным лучше, чем МПС. Следовательно, для дальнейшего прогноза будет выбрана именно ЛММ.

Проведем исследование прогноза развития системы, изменяя неуправляемые факторы x_3 (Исследователи, занятые исследованиями и разработками) и x_4 (Исследователи с учеными степенями).

Изменяя тенденцию развития факторов на $\pm 5\%$, получили результаты прогноза системы на 3 года.

Таблица 6. Отклонения неуправляемых факторов

	$x_4 - 5\%$	$x_4 - 0\%$	$x_4 + 5\%$
$x_3 - 5\%$	0,9896	1,2929	1,6282
$x_3 - 0\%$	0,5190	0,8223	1,1576
$x_3 + 5\%$	-0,0010	0,3022	0,6374

В итоге, наихудшей комбинацией неуправляемых факторов является $x_4 - 5\%$ и $x_3 + 5\%$, при которой оценка системы снижается до минимального значения за весь период.

Рассмотрим возможность лица принимающего решения (ЛПР) компенсировать негативное влияние неуправляемых факторов путем регулирования управляемого фактора x_5 (Внутренние затраты на исследования и разработки). Например, можно увеличить или сократить расходы на научные исследования в пределах 5%.

Изменяя на $\pm 5\%$ тенденцию развития управляемого фактора, получили еще одни прогнозы развития системы на 3 года.

Таблица 7. Изменение управляемого фактора

	$x_5 - 5\%$	$x_5 - 0\%$	$x_5 + 5\%$
y	-0,0609	-0,0010	0,0650

Изменение тенденции развития управляемого фактора на $\pm 5\%$ не дало желаемого результата, поскольку наилучший результат изменения управляемого фактора не приводит к существенным изменениям. Следовательно, ЛПР не в состоянии как-либо компенсировать негативное влияние неуправляемых факторов путем изменения управляемых факторов.

Заключение

Как показало исследование, количество заявок на патенты в построенной модели не оказывает в принципе никакого влияния на критерий системы. Это наглядно видно по значению коэффициента парной корреляции, равного $-0,006$. Куда более значащим фактором можно указать численность исследователей, занятых какими-либо научными разработками, который сильно влияет на количество выдаваемых патентов. Но даже небольшое изменение данной численности приводит к крайне негативным последствиям, которые изменением управляемого фактора, то есть расходом на научные исследования, не удастся исправить. Из данной модели следует, что наиболее важное влияние на количество выдаваемых патентов оказывает именно численность исследователей, причем не так важно с учеными степенями они или нет. Еще менее важное влияние на систему оказывают расходы на научные исследования и разработки, а количество заявок на патенты практически никак не отражает количество выдаваемых патентов.

Список литературы

1. Заседание совета по науке и образованию. — [Электронный ресурс].
URL: <http://kremlin.ru/events/president/news/53313>.
PROGNOZ Data Portal. — [Электронный ресурс]. URL: <http://dataportal.prognoz.ru>.
2. Линейные многофакторные модели. — [Электронный ресурс]. URL: <http://economy-ru.info/info/189276/>.
3. Математические модели в пространстве состояний. — [Электронный ресурс].
URL: <http://drive.ispu.ru/elib/lebedev/14.html>.
4. Авторегрессионные модели. — [Электронный ресурс].
URL: https://ru.wikipedia.org/wiki/Авторегрессионная_модель.
5. Затонский А.В., Сиротина Н.А. Прогнозирование экономических систем по модели на основе регрессионного дифференциального уравнения // Экономика и математические методы. — 2014. Т. 50. — № 1. — С. 91-99.
6. Экс-пост-значениям. — [Электронный ресурс]. URL: http://help.sap.com/saphelp_scm70/helpdata/ru/ac/216b77337b11d398290000e8a49608/frameset.htm.
7. Сиротина Н. А., Янченко Т. В., Затонский А.В. Об аппроксимации факторов дифференциальной модели социально-экономической системы. — 2012. — № 11 (19). — С. 6. — [Электронный ресурс]. URL: <http://sisp.nkras.ru/e-ru/issues/2012/11/sirotina.pdf>.