

ПРОГНОЗИРОВАНИЕ ВРП С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ПРОСТРАНСТВЕННОЙ КОРРЕЛЯЦИИ

Т. Горшкова, н.с., РАНХиГС

При выборе наиболее эффективной экономической политики государств, состоящих из большого количества субъектов с уникальными географическими и экономическими характеристиками, таких как Российская Федерация, необходимо учитывать особенности всех регионов. Однако в связи с трудоемкостью определения потребностей каждого региона в отдельности и объединении их при выработке общей политики представляется эффективным рассматривать экономическое положение не отдельного региона, а нескольких кластеров, в которые входят регионы с похожими характеристиками.

ВВЕДЕНИЕ

Одним из методов исследования региональных данных является построение моделей с учетом пространственных связей между регионами. Предполагается, что регионы, имеющие общую границу или находящиеся на небольшом расстоянии друг от друга, демонстрируют похожую динамику макропоказателей, что связано с одинаковой наделенностью ресурсами и производственным распределением внутри страны. В данной статье исследуются возможности прогнозирования регионального ВРП на российских данных с помощью одномерных пространственных моделей. Несмотря на то, что в общем случае структурные модели позволяют более точно определять параметры, влияющие на ряд данных, в представленном исследовании структурные модели не рассматриваются, так как целью анализа является получение общего вывода о качестве прогнозов ВРП, разработанных с помощью пространственных моделей.

Работа построена следующим образом. В первой части проведен краткий обзор эмпирических исследований, посвященных использованию пространственных моделей. Во второй части одномерные модели, учитывающие пространственную корреляцию между регионами, применены к российскому региональному ряду ВРП, а также к трем подвыборкам ряда для регионов европейской части, Кавказа и Сибири, а также Дальнего Востока. В заключении приводятся основные выводы из исследования и даются рекомендации по использованию моделей прогнозирования регионального ВРП на российских данных.

1. ОБЗОР ЭМПИРИЧЕСКИХ ИССЛЕДОВАНИЙ

Анализ пространственных зависимостей между регионами набирает популярность только в последние годы, в связи с чем в научной литературе имеется довольно небольшой пул работ, посвященных моделям пространственной корреляции. В данной статье кратко рассмотрены только основные модели, а также модели, посвященные анализу регионов, входящих в состав одной страны (США, Германия, Китай). Работы, в которых указанные модели применяются к союзам стран, в текущий обзор не вошли.

Наиболее известным тестом на наличие пространственной автокорреляции является *I* Морана, которое рассчитывается следующим образом:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2}, \quad (1)$$

где x – вектор значений объясняемой переменной, μ – их среднее, N – число наблюдений. *I* Морана принимает значения от -1 до 1 . Значение, равное -1 , свидетельствует о том, что регионы со значениями x выше среднего соседствуют с регионами, в которых значения x ниже среднего. Значение, равное 1 , говорит о положительной корреляции: регионы, в которых x выше среднего, граничат друг с другом. Если значение статистики Морана незначимо, то предполагается, что ошибки распределены между регионами случайным образом. Однако даже в этом случае учет пространственной корреляции может улучшить прогнозные свойства модели^{1,2}.

Витон (2010)³ выделяет следующие базовые типы пространственных моделей, которые сведены в табл. 1.

В работе Семериковой, Демидовой (2016)⁴ модели с пространственной корреляцией используются для прогнозирования безработицы в России и Германии. Авторы исследуют структурные модели, в которых безработица зависит от большого количества экономических показателей, в том числе долей занятых в различных отраслях экономики, ВРП, возрастной структуры рабочей силы. Для учета пространственных эффектов используются модель с пространственной корреляцией в лагах, в ошибках и модель Дарбина. Кроме того, используются два вида пространственных матриц: матрица «соседей первой линии» и обратная матрица расстояний.

На основе проведенного исследования авторы делают вывод, что для Германии выбор взвешивающей матрицы имеет важное значение для получения оценок. Разные матрицы дают разные коэффициенты модели. Для России больших различий в оценках коэффициентов и их значимости между моделями одной спецификации с разными взвешивающими матрицами не наблюдается. Авторы утверждают, что разница в выводах для России и Германии обусловлена тем, что 75 субъектов Российской Федерации занимают гораздо большую площадь, чем 370 регионов Германии. При большом уровне агрегирования пространственное взаимодействие региональных рынков труда, выраженное, например, в виде межрегиональной трудовой мобильности, играет меньшую роль, чем при более мелком региональном делении.

Кроме того, авторы просимулировали российские данные и пришли к выводу, что при моделировании региональной безработицы в России с помощью исследуемых в работе Маддалы и др. (1997)⁵ факторов лучше выбирать матрицу обратных расстояний, отражающую связи между всеми регионами, а не граничную матрицу.

В работе Жирардина и Холодилина (2011)⁶ авторы исследуют темп роста валового регионального продукта (ВРП) в 31 провинции Китая за период 1979–2007 гг. при предположении, что учет пространственной корреляции улучшает прогнозные свойства модели.

Авторы рассматривают несколько моделей:

- AR-модель без ограничений на параметры;
- FE-модель, предполагающую гомогенность коэффициента наклона;
- модель пула, предполагающую гомогенность коэффициента наклона и константы;

¹ Frees E, 2004. Longitudinal and Panel Data: Analysis and Applications for the Social Sciences, New York: Cambridge University Press. ISBN 0-521-82828-7.

² Giacomini, R., Granger, C.W.J., 2004. Aggregation of space-time processes. Journal of Econometrics 118, 7–26.

³ Viton, 2010. Notes on Spatial Econometric Models, City Reg Plann 870(03).

⁴ Семерикова Е.В., Демидова О.А. Взаимодействие региональных рынков труда в России: анализ с помощью пространственных эконометрических моделей // Пространственная экономика. 2016. № 3 С. 57–80.

⁵ Maddala, G.S., Trost, R.P., Li, H., Joutz, F., 1997. Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators. Journal of Business and Economic Statistics 15, 90–100.

⁶ Girardin, E., Kholodilin, K.A., 2011. How helpful are spatial effects in forecasting the growth of Chinese provinces, Journal of Forecasting 30, 622–643.

Таблица 1

БАЗОВЫЕ ТИПЫ ПРОСТРАНСТВЕННЫХ МОДЕЛЕЙ

1	Пространственная модель Дарбина	$y = X\beta + WX\theta + u$ у зависит от значений объясняющих переменных в исследуемом регионе и взвешенных значений этих же объясняющих переменных в соседних регионах
2	SAR – модель пространственной авторегрессии	$y = X\beta + \rho Wy + u$ у зависит от значений объясняющих переменных в исследуемом регионе и от взвешенных значений у в соседних регионах
3	SEM – модель пространственной ошибки	$y = X\beta + (1 - \lambda W)^{-1}v$ у зависит от значений объясняющих переменных в исследуемом регионе и от взвешенной по соседям компоненты ошибки
4	Обобщенная пространственная модель	$y = \rho W_1 v + u$ $u = \lambda W_2 u + v$ у зависит от взвешенных значений у в соседних регионах и от взвешенной по соседям компоненты ошибки
5	Пространственная панельная модель с фиксированными эффектами	$y_{it} = \lambda \sum_j w_{itj} y_{ij} + a_i + x_{it}\beta + u_{it}$ у зависит от фиксированного эффекта, специфичного для исследуемого региона, от значений объясняющих переменных в исследуемом регионе и от взвешенных значений у в соседних регионах
6	Пространственная панельная модель со случайными эффектами	$y_{it} = \lambda \sum_j w_{itj} y_{ij} + x_{it}\beta + v_{it}$ у зависит от специфичного случайного эффекта, входящего в состав ошибки, от значений объясняющих переменных в исследуемом регионе и от взвешенных значений у в соседних регионах

Источник: составлено авторами на основе Витона (2010).

- FE-модель и модель пула с учетом пространственной корреляции;
- модель групповых эффектов, в которой константа специфична для каждой группы провинций (Внутренние и Прибрежные провинции).

Также в работе рассматриваются два типа моделей пространственной корреляции: пространственный лаг (spatial-lag) и пространственная ошибка (spatial-error). Пространственная корреляция учитывается с помощью матрицы пространственных весов W размера $N \times N$. Для обеих моделей пространственной зависимости строятся регрессии пул и FE.

По результатам оценивания можно сделать вывод, что Прибрежные провинции в 1979–2007 гг. росли в среднем быстрее, чем Внутренние. Оцененные коэффициенты пространственной авторегрессии значимы и положительны, что говорит о важности пространственной зависимости среди китайских провинций. Также оценки, полученные в работе, показали, что зависимость между провинциями внутри каждой группы различается. Временная авторегрессия значима, что свидетельствует о важности учета лагов ВРП. Все модели дают лучшие прогнозы, чем наивная модель AR(1). В среднем модели групповых эффектов оказываются лучше, чем пулы, которые в свою очередь лучше фиксированных эффектов.

В статье Longhi, Nijkamp (2007)¹отмечается, что для прогнозирования региональных показателей в пределах одной страны разрабатываются и применяются различные как линейные, так и нелинейные модели, которые зачастую игнорируют пространственную взаимозависимость между регионами. Авторы делают краткосрочные прогнозы уровней региональной безработицы для 326 регионов Западной Германии за период с 1987–2002 гг. Для анализа они используют количество работников, занятых полный рабочий день в 9 секторах экономики, и среднюю дневную зарплату таких рабочих в каждом регионе. Авторы рассматривают непространственные и пространственные модели. В среднем для непространственных моделей модели со случайными эффектами, оцененные мето-

¹ Longhi, S., Nijkamp, P., 2007. Forecasting regional labor market developments under spatial heterogeneity and spatial correlation. *International Regional Science Review* 30, 100–119.

дом максимального правдоподобия, показали лучшие результаты, чем модели с фиксированными эффектами. Модели, учитывающие пространственную гетерогенность, дают худшие прогнозы по сравнению с гомогенными моделями. Таким образом, можно сделать вывод об отсутствии значительной разницы в уровнях занятости в урбанизированных и в сельскохозяйственных регионах. Объединение гетерогенных коэффициентов в пул может дать более надежные прогнозы. Большинство моделей дают лучшие прогнозы в 2000 и 2001 гг., чем «наивная» модель, однако в 2002 г. «наивная» модель оказывается лучше всех остальных. В среднем самые хорошие результаты дала обобщенная модель со случайными эффектами, не учитывающая специфичные эффекты регионов.

Кроме того, авторы рассматривают возможность моделирования пространственных эффектов и пространственной корреляции с помощью модели пространственной структуры ошибок. Для регрессии, учитывающей пространственную корреляцию путем включения средних зарплат в регионе, наилучшие результаты показала модель случайных эффектов, оцененная с помощью метода максимального правдоподобия. В среднем пространственные модели дают более хорошие прогнозы, чем соответствующие непространственные модели. Единственное исключение – модель с фиксированными эффектами.

Аналогично непространственным моделям пространственные оказываются лучше «наивной» в 2000 и 2001 гг., однако хуже в 2002 г. Также модели, не учитывающие гетерогенность пространственных коэффициентов, лучше, чем гетерогенные.

В работе Холодилиной, Силиверстовс, Кутц (2008)¹ прогнозируется годовой темп роста ВВП для 16 немецких регионов. Авторы делят все данные на несколько групп. По пространственному признаку выделяются регионы Восточной и Западной Германии по отдельности и все регионы вместе.

В качестве базовой модели используется линейная AR(1), построенная для каждого региона отдельно. Кроме того, в работе проанализированы следующие модели:

- модель пула, оцененная с помощью двух методов – МНК и ОММ, в которой предлагается идентичность как константы, так и коэффициентов наклона для каждого региона;
- модель с фиксированными эффектами, в которой константа различается для регионов;
- модели, учитывающие пространственную корреляцию путем включения пространственных лагов: пространственная модель Дарбина и пространственная модель Дарбина с фиксированными эффектами;
- пул с пространственной ошибкой;
- модель пространственной ошибки с фиксированными эффектами.

Модели, учитывающие пространственную корреляцию, оцениваются с помощью MLE.

Оценки, полученные с помощью моделей, учитывающих пространственную корреляцию, практически не отличаются от оценок более простых моделей. Оценки пространственных авторегрессионных коэффициентов оказываются незначимыми. Однако Элхорст (2003)² показал, что следует использовать модели пространственной автокорреляции, даже если ее коэффициенты оказываются незначимыми. Оценки модели пространственной ошибки, полученные с помощью MLE, близки к МНК-оценкам, но в отличие от моделей пространственной авторегрессии коэффициент пространственной корреляции λ оказывается значимым и положительным. Прогнозы по моделям, учитывающим пространственную корреляцию, оказались более точными. Среди них наибольшую точность показали оценки, полученные для пула.

Результаты всех рассмотренных моделей приведены в *табл. 2*.

¹ Kholodilin, K.A., Siliverstovs, B., Kooths, S., 2008. A dynamic panel data approach to the forecasting of the GDP of German Länder. *Spatial Economic Analysis* 3, 195–207.

² Elhorst, J. P., 2003. Unconditional maximum likelihood estimation of dynamic models for spatial panels. Research Report 03C27, University of Groningen, Research Institute SOM (Systems, Organisations and Management).

Таблица 2
 РЕЗУЛЬТАТЫ ПРОСТРАНСТВЕННЫХ МОДЕЛЕЙ

Авторы	Данные	Сравниваемые модели	Выводы
Семерикова, Демидова (2016)	Уровень безработицы в России и Германии	Модель с пространственной корреляцией в лагах и в ошибках, модель Дарбина	Для Германии выбор матрицы дают разные коэффициенты модели. Для России различия не существенны. При большом уровне агрегирования пространственное взаимодействие региональных рынков труда играет меньшую роль, чем при более мелком делении.
Girardin, Kholodilin (2011)	Темп роста ВРП в Китае	AR-модель, модель с пространственной корреляцией в лагах и в ошибках, модель групповых эффектов, в которой константа специфична для каждой групп провинций	Коэффициенты пространственной авторегрессии свидетельствуют о важности пространственной зависимости среди китайских провинций. Модели групповых эффектов оказываются лучше, чем пулы, которые в свою очередь лучше фиксированных эффектов.
Longhi, Nijkamp (2007)	Уровень безработицы в Германии	Пространственные и непространственные модели	Пространственные модели дают более хорошие прогнозы, чем непространственные, однако, в целом не улучшают прогнозы, полученные с помощью непространственных моделей.
Kholodilin, Siliverstovs, Kooths (2008)	Темп роста ВРП в Германии	AR(1), пул, модель с фиксированными эффектами, модели, пространственной корреляции (пул, с фиксированными эффектами в лагах и в ошибках)	Оценки пространственных моделей, практически не отличаются от оценок более простых моделей, но прогнозы по пространственным моделям оказались более точными.

Источник: составлено авторами.

2. ПРОГНОЗИРОВАНИЕ ВРП РОССИИ

При прогнозировании ВРП используется информация о годовых значениях валового регионального продукта для 83 регионов России за период 2000–2015 гг. Прогнозирование показателей для Республики Крым и города федерального значения Севастополь не проводится в связи с отсутствием данных до 2014 г.

В качестве базовых моделей, с которыми будут сравниваться остальные прогнозы, используются линейные индивидуальные модели авторегрессии AR(1), оцененные для каждого из субъектов отдельно с помощью метода наименьших квадратов. Для каждого региона в отдельности оцениваются уравнения следующего вида:

$$y_{it} = \alpha_i + \beta_i y_{it-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim NID(0, \sigma_i^2), \quad (2)$$

где y_{it} – уровень ВВП для субъекта i в период времени t .

Из 83 регионов только в 15 коэффициент при лаге ВРП оказывается значимым на 10%-ном уровне значимости. Все значимые коэффициенты оказались отрицательными.

Константа практически во всех регрессиях оказывается незначимой. Таким образом, согласно результатам AR(1) моделей, прошлое значение ВРП отрицательно влияет на текущее значение. Отрицательное значение коэффициента может служить косвенным подтверждением наличия безусловной β -конвергенции в российских регионах. В рамках модели безусловной конвергенции темпы экономического роста положительно коррелированы с разрывом между стартовым душевым доходом данного региона и уровнем душевого дохода в устойчивом состоянии равновесия, одинаковым для всех регионов, в начальный момент времени. В соответствии с моделью бедные регионы должны расти более быстрыми темпами, чем богатые, так что в долгосрочной перспективе должно происходить выравнивание региональных уровней экономического развития.

Основной недостаток индивидуальных оценок наименьших квадратов заключается в том, что в связи с малым количеством временных данных оценки коэффициентов, оцененные с помощью МНК, оказываются смещенными. Средняя ошибка прогноза, полученная по наивной AR(1) модели, равна 23%.

Следующая модель – модель пула, в которой константа и коэффициент наклона предполагаются гомогенными:

$$y_{it} = \alpha + \beta y_{it-1} + \varepsilon_{it}, \varepsilon_{it} \sim NID(0, \sigma_i^2), \quad (4)$$

Данная модель оценивается 2-шаговым обобщенным методом моментов на основании методики Ареллано-Бонда. При использовании данного метода лаги зависимой переменной могут использоваться одновременно как в качестве объясняющих переменных, так и в качестве инструментов. При анализе ВРП применялись различные комбинации лаговых значений, критерием качества служил тест Саргана, проверяющий сверхидентифицируемость инструментов.

Также, как и в «наивных» индивидуальных AR(1) моделях, первый лаг выпуска в панельной регрессии оказывается отрицательным и значимым. Тест Вальда на совместную незначимость всех коэффициентов отвергается. J-тест Саргана не отвергает гипотезу об идентифицируемости инструментов.

На основе данной модели были построены прогнозы для ВРП. Средняя ошибка модели пула без учета пространственной корреляции равна 25%.

Следующая модель – панельные данные с фиксированными эффектами, в которой константа предполагается регионально специфической переменной:

$$y_{it} = \alpha_i + \beta y_{it-i} + \varepsilon_{it}, \varepsilon_{it} \sim NID(0, \sigma_i^2), \quad (5)$$

Гипотеза о том, что индивидуальные фиксированные эффекты данной модели равны 0, отвергается, т.е. индивидуальные эффекты различаются в зависимости от региона, и использование модели пула или модели со случайными эффектами даст смещенные оценки. Об этом же свидетельствует и очень высокое значение корреляции между фиксированными эффектами и значением ВРП, которая равна 0,94. Ошибка прогноза, построенного по данной модели, оказывается ниже ошибки «наивной» AR(1) модели и составляет 17%.

Далее были построены регрессии, учитывающие пространственную корреляцию. Для оценки этих моделей также использовались данные, собранные в пул, и модели с фиксированными эффектами. Все модели пространственной автокорреляции оценивались методом максимального правдоподобия.

Матрица весов представляет собой матрицу смежности, в которой регионам, имеющим общую морскую или сухопутную границу, присвоено значение 1, остальным – 0. Граф смежности регионов представлен на рис. 1, где каждому региону присвоен свой порядковый номер от 1 до 83.

Диагональные элементы весовой матрицы равны 0. Элементы каждой строки матрицы нормированы таким образом, что сумма весовых коэффициентов по каждой строке равна 1. На основе матрицы

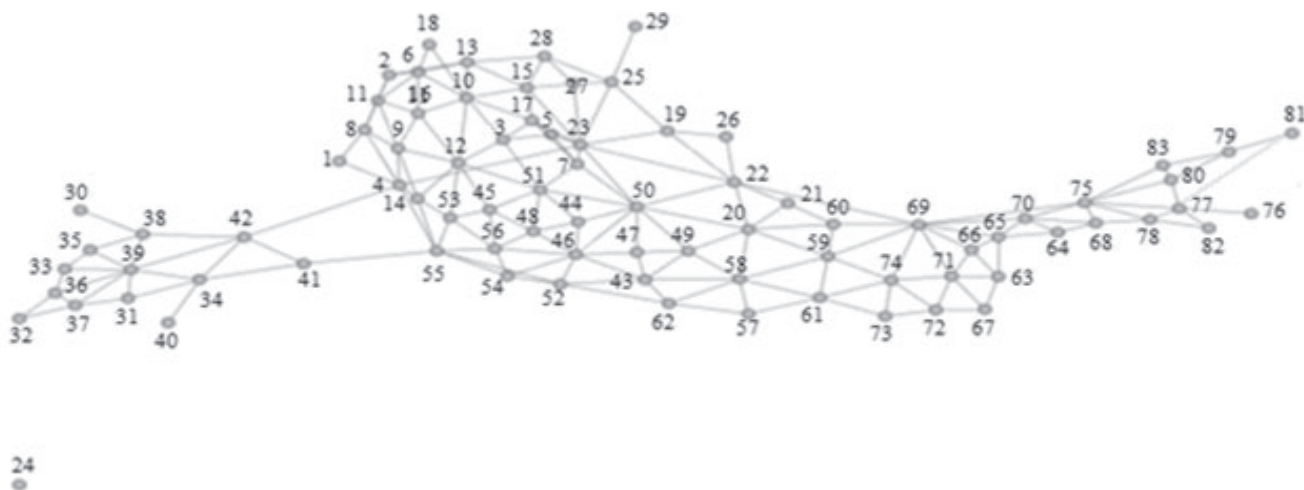


Рис. 1. Граф смежности регионов

Источник: расчеты авторов.

весов и данных по ВРП, очищенных от тренда, были рассчитаны значения I Морана по следующей формуле

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2}, \quad (6)$$

где N – количество регионов, w_{ij} – элемент весовой матрицы, x_i – значение ВРП в регионе i , μ – среднее значение ВРП.

Все значения I Морана значимо положительные, что говорит о наличии положительной пространственной корреляции между регионами.

Сначала были проанализированы данные, объединенные в пул:

$$y_{it} = \alpha + \beta y_{it-1} + \delta \sum_{j=1}^{83} w_{ij} y_{jt} + \varepsilon_{it}. \quad (7)$$

В данной спецификации прошлый ВРП отрицательно влияет на текущее значение, что соответствует результатам индивидуальных AR(1) моделей и может быть связано с наличием безусловной конвергенции. Тест Вальда на совместную незначимость всех коэффициентов отвергается. J-тест Саргана свидетельствует о хорошем подборе инструментальных переменных. Средняя абсолютная ошибка прогноза по данной модели равна 19%, что выше, чем ошибка моделей с фиксированными эффектами, не учитывающих пространственную корреляцию между регионами, но ниже ошибок других непространственных моделей.

В модели с фиксированными эффектами первый лаг ВРП оказывается положительным, что соответствует результатам, полученным для панели с фиксированными эффектами без учета пространственной корреляции, и противоречит выводам, полученным на основе пула и индивидуальных моделей. В целом регрессия значима. Гипотеза о том, что индивидуальные фиксированные эффекты равны 0, отвергается, корреляция между фиксированными эффектами и значением ВРП равна по модулю 0,62. Средняя абсолютная ошибка прогноза по данной модели равна 10%, что ниже ошибок других моделей.

Модель с фиксированными эффектами для пространственной корреляции в ошибках оценивалась методом максимального правдоподобия по аналогии с работой Элхорста (2003). В данной работе предполагается оценивать уравнения, очищенные от индивидуальных эффектов путем усреднения рядов объясняемой и объясняющих переменных. Далее оценивается функция правдоподобия следующего вида:

$$-\frac{NT}{2} \ln(2\pi\sigma^2) + T \sum_{i=1}^N \ln(1 - \delta w_i) - \frac{1}{2\sigma^2} \sum_{t=1}^T e_t' e_t, \quad e_t = (I - \delta W)[Y_t - \bar{Y} - (X_t - \bar{X})\beta], \quad (8)$$

где w_i – характеристические корни весовой матрицы W , σ^2 – дисперсия ошибок оцениваемого уравнения.

Полученные результаты несравнимы с ОММ-оценками, поэтому для сравнимости также были построены оценки максимального правдоподобия для пространственной SAR модели. Функция правдоподобия в таком случае принимает следующий вид:

$$-\frac{NT}{2} \ln(2\pi\sigma^2) + T \sum_{i=1}^N \ln(1 - \delta w_i) - \frac{1}{2\sigma^2} \sum_{t=1}^T e_t' e_t, \quad e_t = (I - \delta W)(Y_t - \bar{Y}) - (X_t - \bar{X})\beta. \quad (9)$$

Средняя абсолютная ошибка прогноза по данной модели равна 15%, что оказывается ниже ошибок всех остальных моделей.

Таким образом, наиболее точным методом прогнозирования ВРП на основе данных по всем российским регионам оказывается модель пространственной корреляции в ошибках, рассчитанная ме-

тодом максимального правдоподобия. При этом точность моделей пространственной корреляции не сильно отличается от точности модели с фиксированными эффектами без учета пространственных связей. В частности, точность модели с учетом пространственной корреляции в ошибках превышает точность всего на 2 п.п. Точность остальных моделей оказывается ниже моделей с фиксированными эффектами.

Средние ошибки прогнозов всех рассмотренных моделей сведены в *табл. 3*.

Затем все регионы были разбиты на три части в соответствии с административно-территориальным делением РФ. Такое деление обуславливается социальными, политическими и климатическими причинами. Например, для северных регионов Сибири и Дальнего Востока Трудовым Кодексом РФ и законом РФ от 19.02.1993 № 4520-1 предусмотрены надбавки к зарплатам за стаж. Регионы Кавказа являются крупнейшими дотационными регионами в России. Эти два фактора (надбавки и дотации) влияют на ВРП регионов, что приводит к необходимости рассматривать их отдельно.

Таблица 3
МАРЕ МОДЕЛЕЙ

	Без пространственных эффектов			С пространственными эффектами		
	AR(1)	P_{OLS}	F_{OLS}	P_{SAR}	F_{SAR}^*	F_{SEM}
MAPE	23%	25%	17%	19%	20%	15%

* Приведена минимальная ошибка из моделей, рассчитанных OMM и MLE

Источник: расчеты авторов.

Деление на более мелкие округа (по географическим или экономическим зонам) невозможно в связи с малым количеством пространственных наблюдений, которые будут включены в такие зоны. В европейскую часть России входят 49 регионов, в кавказскую – 10 регионов, Сибирь и Дальний Восток состоят из 24 регионов.

Для каждого района была рассчитана своя весовая матрица смежности субъектов, входящих в тот же район и построены те же оценки, что и для всех субъектов одновременно.

В целом результаты AR(1) модели не сильно отличаются от результатов моделей с учетом пространственной корреляции. Ошибки всех моделей, приведенные в *табл. 4*, практически идентичны.

Таблица 4
МАРЕ МОДЕЛЕЙ РЕГИОНОВ ЕВРОПЕЙСКОЙ ЧАСТИ РОССИИ

	Без пространственных эффектов			С пространственными эффектами		
	AR(1)	P_{OLS}	F_{OLS}	P_{SAR}	F_{SAR}^*	F_{SEM}
MAPE	15%	10%	13%	11%	10%	12%

* Оценивание OMM не проводилось

Источник: расчеты авторов.

В отличие от предыдущего случая наиболее точной оказалась модель пространственной корреляции в лагах, рассчитанная методом максимального правдоподобия. При этом для европейских регионов России модель, учитывающая пространственную корреляцию в ошибках, дает менее точные результаты не только по сравнению с SAR моделью, но и относительно модели пула, не учитывающей пространственные связи между регионами.

В целом все модели с учетом пространственной корреляции дают результаты, сравнимые по точности с SAR моделью. В частности, точность пула уступает точности F_{SAR} модели всего на 1 п.п. Полученные результаты соответствуют результатам эмпирических исследований на зарубежных данных, в том числе и относительно схожести результатов пространственных моделей с результатами «наивной» AR(1) модели.

По моделям пространственной корреляции были рассчитаны средние ошибки прогнозов, которые приведены в *табл. 5*. Прогнозы по моделям, не учитывающим пространственную корреляцию, не строились в связи с незначимостью первого лага ВРП.

Таблица 5

МАРЕ МОДЕЛЕЙ ПРОСТРАНСТВЕННОЙ КОРРЕЛЯЦИИ РЕГИОНОВ КAVKAZA

	P_{SAR}	F_{SAR}	F_{SEM}
МАРЕ	6.4%	2.7%	5%

Источник: расчеты авторов.

Ошибки всех моделей практически идентичны, однако наиболее точной, так же, как и в моделях для всех регионов и для регионов европейской части России, оказывается модель пространственной корреляции в лагах, рассчитанная методом максимального правдоподобия. Оценки коэффициентов в моделях, не учитывающих пространственную корреляцию, оказались незначимыми.

По всем моделям были рассчитаны средние ошибки прогнозов, которые приведены в *табл. 6*.

Таблица 6

МАРЕ МОДЕЛЕЙ РЕГИОНОВ СИБИРИ И ДАЛЬНЕГО ВОСТОКА

	Без пространственных эффектов			С пространственными эффектами		
	AR(1)	P_{OLS}	F_{OLS}	P_{SAR}	F_{SAR}	F_{SEM}
МАРЕ	2.6%	2.3%	2.5%	3.1%	4.2%	4.5%

Источник: расчеты авторов.

В отличие от предыдущих моделей для регионов Сибири и Дальнего Востока прогнозы, полученные методом максимального правдоподобия на основе моделей, учитывающих пространственную корреляцию, оказались наименее точными. Этот результат не соответствует выводам, сделанным на основе статистики Морана, которая свидетельствует о наличии значимой положительной корреляции между регионами Сибири и ДВ на всем исследуемом промежутке. Однако ошибка прогноза пула с учетом пространственной корреляции оказывается меньше ошибок соответствующих моделей с фиксированными эффектами, хотя и выше ошибок непространственных моделей. Полученный результат может свидетельствовать о незначимости индивидуальных фиксированных эффектов среди сибирских и дальневосточных регионов. Сложившаяся ситуация может объясняться отсутствием значительных различий в экономической специализации данных регионов, практически одинаковой удаленностью от центра и похожей скоростью экономического развития (удаленность регионов Дальнего Востока и суровые природные условия сибирских регионов практически одинаково замедляют процессы экономического развития и усложняют государственное управление).

В целях оценивания точности полученных результатов также были построены модели для других весовых матриц. Были использованы еще две матрицы взаимосвязей между регионами, которые описывались в предыдущем разделе. Вторая матрица представляет собой матрицу регионов «второй линии» – регионы считаются соседними, если их разделяет только один регион. Третья матрица – матрица расстояний, элементы которой представляют собой не 1 и 0, а обратная матрица расстояний между центрами регионов.

Для этих двух матриц также были построены прогнозы по моделям с учетом пространственной корреляции и рассчитаны соответствующие ошибки прогнозов. В *табл. 7* приведено сравнение точностей всех моделей для трех матриц: (1) – матрица смежностей, (2) – матрица «второй линии», (3) – матрица расстояний.

Таблица 7

МАРЕ МОДЕЛЕЙ С РАЗЛИЧНЫМИ ВЕСОВЫМИ МАТРИЦАМИ

	Все регионы			Европейская часть			Кавказ			Сибирь и ДВ		
	P_{SAR}	F_{SAR}	F_{SEM}	P_{SAR}	F_{SAR}	F_{SEM}	P_{SAR}	F_{SAR}	F_{SEM}	P_{SAR}	F_{SAR}	F_{SEM}
(1)	19%	20%	15%	11%	10%	12%	6.4%	2.7%	5%	3.1%	4.2%	4.5%
(2)	18%	20%	15%	13%	10%	11%	5%	3%	4%	15%	16%	19%
(3)	18%	21%	14%	12%	10%	12%	4.5%	3%	4%	26%	24%	14%

Источник: расчеты авторов.

Таким образом, по результатам прогнозирования ВРП для трех различных весовых матриц можно сделать вывод, что все три матрицы дают в целом идентичные результаты. Исключением являются регионы Сибири и Дальнего Востока, для которых прогнозы по модели, использующей матрицу смежностей, существенно лучше остальных. Для регионов Сибири и Дальнего Востока наименьшей точностью обладает модель с матрицей «второй линии». Это может быть связано с большими расстояниями даже между соседними регионами, что приводит к практически нулевым весам при учете влияния «соседей через регион».

Наиболее точные результаты по всем моделям были получены для регионов Кавказа, что может быть связано с тесными историческими и геополитическими связями между данными регионами. Для регионов Сибири и Дальнего Востока матрица «второй линии» и матрица расстояний дают прогнозы с очень высокой ошибкой. Таким образом, прогнозирование ВРП с помощью стандартной матрицы соседей дает наиболее точные результаты.

По результатам прогнозирования для четырех подвыборок можно сделать выводы, что, несмотря на различия в ошибках прогноза для частей РФ, модели пространственной корреляции дают наилучшие результаты. При этом для большинства подвыборок AR(1) модели также демонстрируют высокую точность прогнозов. Однако, согласно исследованиям Балтаджи, Ли (2004)¹, даже при отсутствии значимой пространственной связи между регионами для получения эффективных оценок лучше использовать пространственные модели.

Также необходимо заметить, что в данной работе большинство непространственных моделей оценивались методом наименьших квадратов. При оценивании динамических панельных данных МНК оценки оказываются смещенными. В связи с этим для оценивания обычно используется метод обобщенных моментов. Однако на малом временном промежутке ОММ оценки дают нестабильные результаты, поэтому, согласно Жирардину, Холодилину (2011), в целях прогнозирования лучше использовать смещенные, но стабильные МНК оценки.

Таким образом, наиболее точные результаты в среднем для ВРП дают модели с фиксированными эффектами с учетом и модели пула без учета пространственных связей, построенные не для полной выборки регионов, а для регионов, объединенных в кластеры.

ЗАКЛЮЧЕНИЕ

В этой работе осуществлен анализ методов прогнозирования макроэкономических панелей, а также их приложения к российским региональным данным.

В первой части проведен обзор работ, посвященных различным свойствам рядов данных и их оценкам. В частности, рассмотрены модели, учитывающие пространственную зависимость между различными способами и проведен обзор работ, посвященных вопросу влияния пространственных эффектов на прогнозы региональных данных. На основе большинства рассмотренных работ можно сделать вывод, что в среднем включение пространственных эффектов либо не улучшает результаты «наивных» AR моделей, либо улучшает незначительно.

¹ Baltagi, B.H., Li, D., 2004. Prediction in the panel data model with spatial correlation. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (Eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Springer, Berlin, 283–295 (Chapter 13).

Во второй части исследованы эмпирические свойства методов прогнозирования в панелях применительно к российским макроэкономическим региональным панелям для ряда ВРП. Данный показатель прогнозировался с учетом и без учета пространственной взаимосвязи между регионами. В среднем при прогнозировании ВРП наиболее точные результаты дала модель пространственной корреляции с фиксированными эффектами в лагах, оцененная методом максимального правдоподобия. Были использованы три весовые матрицы, которые в среднем дали практически идентичные результаты. При этом для большинства подвыборок AR(1) модели демонстрируют более низкую точность прогнозов.

Таким образом, можно сделать вывод, что российские регионы демонстрируют сильную взаимосвязь и при прогнозировании, в частности ВРП, необходимо учитывать пространственную корреляцию между регионами. При этом спецификация весовой матрицы не имеет большого значения, однако определение регионов в три кластера на основе административно-территориального деления значительно снижает ошибки прогнозов, полученных всеми методами.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Frees E, 2004. *Longitudinal and Panel Data: Analysis and Applications for the Social Sciences*, New York: Cambridge University Press. ISBN 0-521-82828-7.
2. Giacomini, R., Granger, C.W.J., 2004. Aggregation of space-time processes. *Journal of Econometrics* 118, 7–26.
3. Viton, 2010. *Notes on Spatial Econometric Models*, City Reg Plann 870(03).
4. Семерикова Е.В., Демидова О.А. Взаимодействие региональных рынков труда в России: анализ с помощью пространственных эконометрических моделей // *Пространственная экономика*. 2016. № 3 С. 57–80.
5. Maddala, G.S., Trost, R.P., Li, H., Joutz, F., 1997. Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators. *Journal of Business and Economic Statistics* 15, 90–100.
6. Girardin, E., Kholodilin, K.A., 2011. How helpful are spatial effects in forecasting the growth of Chinese provinces? *Journal of Forecasting* 30, 622–643.
7. Longhi, S., Nijkamp, P., 2007. Forecasting regional labor market developments under spatial heterogeneity and spatial correlation. *International Regional Science Review* 30, 100–119.
8. Kholodilin, K.A., Siliverstovs, B., Kooths, S., 2008. A dynamic panel data approach to the forecasting of the GDP of German Länder. *Spatial Economic Analysis* 3, 195–207.
9. Elhorst, J. P., 2003. Unconditional maximum likelihood estimation of dynamic models for spatial panels. Research Report 03C27, University of Groningen, Research Institute SOM (Systems, Organisations and Management).
10. Baltagi, B.H., Li, D., 2004. Prediction in the panel data model with spatial correlation. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (Eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Springer, Berlin, 283–295 (Chapter 13).