

УДК 33

DOI: 10.34670/AR.2023.11.60.051

Формирование информационной базы стратегического планирования муниципального образования в условиях некомплектности данных

Павленков Михаил Николаевич

Доктор экономических наук, профессор,
Нижегородский государственный университет им. Н.И. Лобачевского (Дзержинский филиал),
606037, Российская Федерация, Дзержинск, ул. Буденного, 17;
e-mail: kaf-fin-df@yandex.ru

Парамонов Александр Васильевич

Кандидат экономических наук, доцент,
директор,
Нижегородский институт управления,
Российская академия народного хозяйства и государственной службы при Президенте РФ,
603950, Российская Федерация, Нижний Новгород, пр. Гагарина, 46;
e-mail: paramonov-alv@ranepa.ru

Аннотация

В статье рассмотрены вопросы формирования информационной базы стратегического планирования муниципального образования в условиях некомплектности ретроспективных данных. В управлении деятельностью муниципального образования требуются качественные данные, в том числе при прогнозировании показателей развития территории. Основными причинами некомплектности данных являются: пропуски; сокрытие данных; недостоверность данных. Для обеспечения информационной поддержки прогнозирования показателей развития территории разработан метод оценки достоверности данных. Метод, используя синтез оценок по периоду и источнику возникновения, обеспечивает интегрированную оценку достоверности исходных данных, что позволяет адекватно оценить исходные данные. Для реализации метода разработаны модель «Восстановление данных» и модель «Формирование временного ряда». Разработана модель и алгоритм восстановления данных, используя полиграммную оценку. Модель позволяет качественно обработать некомплектность данных и сформировать матрицу восстановленных данных. Для формирования временного ряда, применяются экспертные методы. Разработана модель формирования компонент временного ряда в условиях множества источников их возникновения, отличающаяся от традиционных методов тем, что данные оцениваются экспертами дважды, используя модифицированную шкалу Харрингтона, что позволяет повысить качество исходных данных для прогнозирования развития территории муниципального образования.

Для цитирования в научных исследованиях

Павленков И.М., Парамонов А.В. Формирование информационной базы стратегического планирования муниципального образования в условиях некомплектности данных // Экономика: вчера, сегодня, завтра. 2023. Том 13. № 11А. С. 466-473. DOI: 10.34670/AR.2023.11.60.051

Ключевые слова

Муниципальное образование, управление, прогнозирование, ретроспективные данные, некомплектные данные, экспертные методы.

Введение

В деятельности муниципального образования задействованы государственные, региональные и частные предприятия, поэтому проблемой управления является координация их деятельности, а для этого требуются качественные данные и эффективная организация информационного взаимодействия при прогнозировании показателей развития территории.

Для прогнозирования используются ретроспективные данные. В практических расчетах эти данные обычно формируются в форме таблиц. Для оценки достоверности и полноты используется метод дисперсии: определяется дисперсия строки и дисперсия столбца таблицы. Для обеспечения информационной поддержки процесса прогнозирования показателей развития территории разработан новый метод оценки достоверности данных и формирования временного ряда.

Предлагаемый метод, используя синтез оценок по периоду и источнику возникновения, обеспечивает интегрированную оценку достоверности исходных данных, что позволяет адекватно оценить исходные данные, обеспечить информационную базу достоверными данными для прогнозирования показателей развития территории муниципального образования.

Основная часть

Для реализации метода разработаны две модели: модель MV «Восстановление данных» и модель MF «Формирование временного ряда».

Формализовано модель MV представим в следующем виде:

$$F(X,t) = Y,$$

где

X – исходные данные;

t – период учета данных;

Y- восстановленные данные;

F – оператор преобразования данных.

Данные для прогнозирования поступают из множества источников I (i=1,I) за T (t=1,T) периодов учета данных (лет).

Среди основных причин некачественных и недостоверных исходных данных можно выделить некомплектность данных: пропуски; сокрытие данных; недостоверность данных.

Некомплектность этих данных приводит к искажению статистических результатов, поэтому предлагается подход к обработке данных для использования результатов в прогнозировании.

Вариант первый: исключение пропусков данных можно рекомендовать при малом количестве таких пропусков.

Вариант второй: заполнение пропущенных данных, но оценки результатов по реальным и новым данным могут иметь сильное смещение.

Имеется ряд методов заполнения пропущенных данных в исходной информации.

Методы подбора. Эти методы реализуются легко, но имеют значительный разброс дисперсии.

Методы регрессии. Эти методы зависят от качества модели регрессии.

Сплайн-интерполяция. Если имеются пропуски, которые следуют подряд, то оценки могут быть точными не всегда.

МП-оценивание. Прим еняются функции правдоподобия для заполнения пропущенных данных.

Кластерный анализ. При восстановлении позволяет определить предпочтительный порядок.

Факторный анализ. Этот метод достаточно трудоемкий в реализации.

Локальные алгоритмы. Для алгоритмов Zet (Wanga), чтобы подтвердить восстановленные значения их правдоподобность, требуется задать ряд важных параметров.

Нейросетевые методы. Эти методы требуют дополнительных научных исследований для восстановления данных.

Некоторые практические результаты заполнения пропущенных данных приведены в работе.

Рассмотренные методы могут быть рекомендованы исследователю, если он выбрал метод решения задачи. При решении практических задач следует понимать, что имеются разные источники и оценить некомплектность данных сложно и более того невозможно в том числе из-за большого объема информации. Для решения этих задач разработаны модели MV и MF.

Модель MV. Восстановление данных.

Разработана модель и алгоритм восстановления данных, используя полиграммную оценку. Общая структура модели показана на рисунке 1.

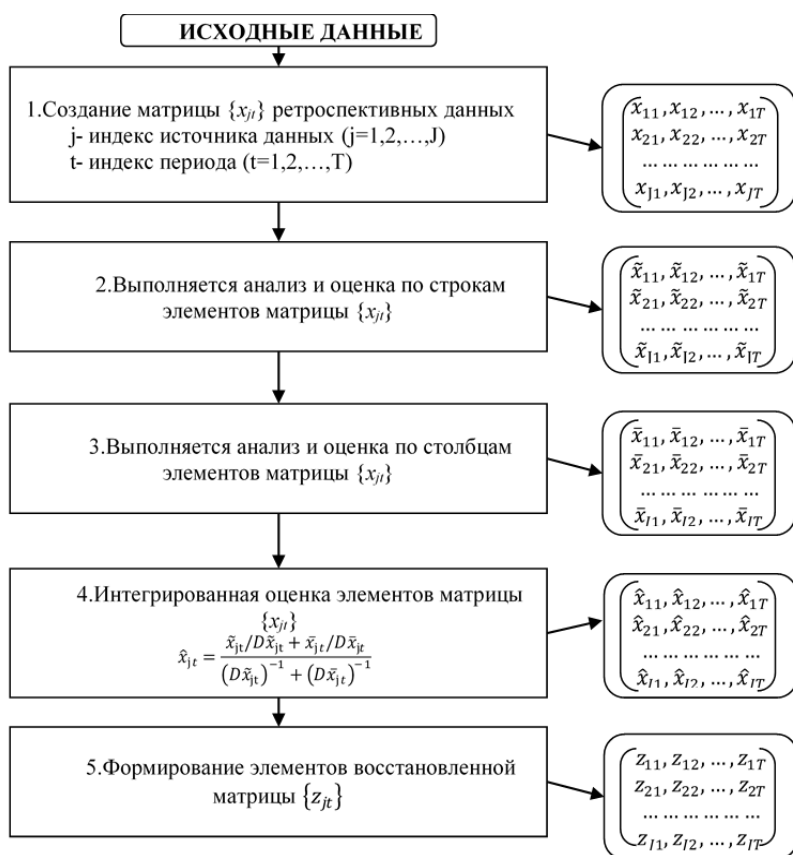


Рисунок 1 - Структура модели восстановления данных

Разработанная модель восстановления данных отличается тем, что оценка выполняется по параметру времени « t » и информационному источнику « j ».

Точность оценивания пропущенных значений снижается из-за того, что с помощью традиционных методов значения элементов строк и столбцов оцениваются индивидуальной дисперсией.

В модели, которая разработана, выполняются оценки 1-го рода и 2-го, а дисперсия зависит от некомплектных данных в строках и столбцах соответственно. Модель позволяет качественно обработать некомплектность данных и сформировать матрицу $\{z_{jt}\}$ $j \in \overline{1, J}$; $t \in \overline{1, T}$ восстановленных данных.

Модель MF «Формирование временного ряда».

Для формирования временного ряда, применяются экспертные методы с использованием модифицированной шкалы Харрингтона. Структура модели MF приведена на рисунке 2.

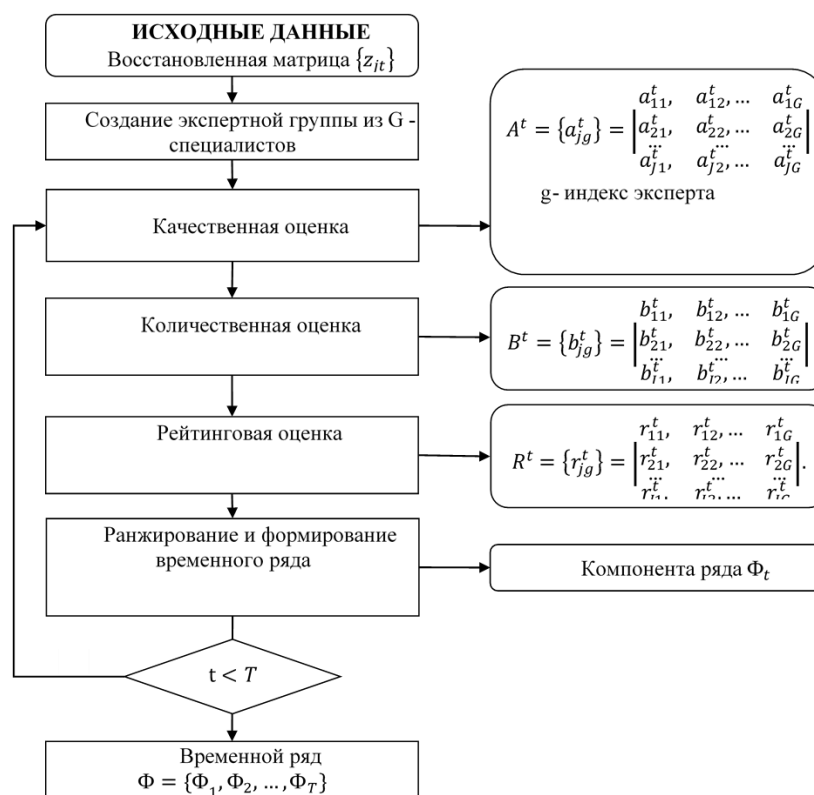


Рисунок 2 - Общая структура модели

Для вычисления компонент временного ряда имеем матрицу $Z = \{z_{jt}\}$:

$$Z = \{z_{jt}\} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1T} \\ z_{21} & z_{22} & \dots & z_{2T} \\ \dots & \dots & \dots & \dots \\ z_{j1} & z_{j2} & \dots & z_{jT} \end{bmatrix}.$$

Сформировали экспертную группу из специалистов численностью G .

Экспертные оценки носят качественный характер:

0-нулевой уровень;

1-уровень низкий;

2-уровень ниже среднего;

3-уровень средний;

4-уровень выше среднего;

5-высокий уровень.

Эксперты оценивают источники информации в матрице j ($j=\overline{1, J}$) по каждому периоду возникновения t ($t=\overline{1, T}$). Источники после оценки сравниваются, ранжируются и определяется рейтинг. Далее числовыми из шкалы Харрингтона (Таблица 4.1.1, 4.1.2) значениями заменяются, полученные после ранжирования.

1. Формирование матрицы A^t .

Для каждого периода t ($t \in \overline{1, T}$) берется вектор-столбец Z_t из матрицы Z

$$Z_t = \begin{pmatrix} z_{1t} \\ z_{2t} \\ \ddots \\ z_{jt} \end{pmatrix}.$$

Эксперты g ($g=1, 2, \dots, G$) качественно оценивают данные вектор-столбца Z_t . Результаты оценки записываются в виде матрицы A^t :

$$A^t = \{a_{jg}^t\} = \begin{pmatrix} a_{11}^t & a_{12}^t & \dots & a_{1G}^t \\ a_{21}^t & a_{22}^t & \dots & a_{2G}^t \\ \ddots & \ddots & \ddots & \ddots \\ a_{j1}^t & a_{j2}^t & \dots & a_{jG}^t \end{pmatrix}.$$

2. Формирование матрицы B^t . Элементы матрицы A^t заменяются численными значениями из таблицы 2:

$$B^t = \{b_{jg}^t\} = \begin{pmatrix} b_{11}^t & b_{12}^t & \dots & b_{1G}^t \\ b_{21}^t & b_{22}^t & \dots & b_{2G}^t \\ \ddots & \ddots & \ddots & \ddots \\ b_{j1}^t & b_{j2}^t & \dots & b_{jG}^t \end{pmatrix}.$$

3. Формирование рейтинговой матрицы R^t .

Расчет по каждому источнику обобщенной оценки s_j^t по следующей формуле

$$s_j^t = \sum_{g=1}^G b_{jg}^t.$$

Определяется лучшая оценка m_j^t по следующей формуле

$$m_j^t = \max_g b_{jg}^t.$$

По каждому источнику эксперты устанавливают качественные значения самый высокий - 1 и т.д. самый низкий - 5.

$$R^t = \{r_{jg}^t\} = \begin{pmatrix} r_{11}^t & r_{12}^t & \dots & r_{1G}^t \\ r_{21}^t & r_{22}^t & \dots & r_{2G}^t \\ \ddots & \ddots & \ddots & \ddots \\ r_{j1}^t & r_{j2}^t & \dots & r_{jG}^t \end{pmatrix}.$$

Определяется нормированный рейтинг r_j^t :

$$r_j^t = 1 - \min_g \left(r_{jg}^t / \max_j r_{jg}^t \right).$$

Формирование вектора \bar{r}_t , который по показателям m_j^t , s_j^t , r_j^t лексикографическим методом ранжируется.

$$\bar{r}_t = \begin{pmatrix} r_{1t} \\ r_{2t} \\ \vdots \\ r_{jt} \end{pmatrix}.$$

Количественное значение ранжированного ряда j ($j=\overline{1, J}$) в t -м периоде становится компонентой ряда $\Phi = \{\Phi_t\}$.

Данный расчет проводится по всем периодам T . Так временный ряд $\Phi = \{\Phi_1, \Phi_2, \dots, \Phi_T\}$ формируется для прогнозирования данного показателя.

Заключение

Разработан метод оценки достоверности данных, отличающийся от традиционных методов тем, что, используя синтез оценок по периоду и источнику возникновения, обеспечивается интегрированная оценка достоверности исходных данных, что позволяет адекватно оценить исходные данные, обеспечить информационную базу достоверными данными для прогнозирования показателей развития территории муниципального образования.

Усовершенствован метод формирования компонент временного ряда в условиях множества источников их возникновения, отличающийся от традиционных методов тем, что данные оцениваются экспертами дважды, используя модифицированную шкалу Харрингтона. Метод позволяет повысить качество исходных данных для прогнозирования развития территории муниципального образования.

Библиография

1. Демиденко Е.З. Линейная и нелинейная регрессия. М.: Финансы и статистика, 1981. 304 с.
2. Загоруйко Н. Г., Елкина В. Н., Тимеркаев В. С. Алгоритм заполнения пропусков в эмпирических таблицах (алгоритм Zet) // Эмпирическое предсказание и распознавание образов. Вып. 61: Вычислительные системы. Новосибирск, 1975. С. 3-27.
3. Захарова А. А. Новые модели принятия решений о социально-экономическом развитии города // Менеджмент в России и за рубежом. 2006. № 1.
4. Круглов В. В., Абраменкова И. В. Методы восстановления пропусков в массивах данных // Программные продукты и системы. 2005.- № 2. С. 59-63.
5. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. М.: Горячая линия - Телеком. 2002. 382 с.
6. Литтл Р.Дж.А., Рубин Д. Б. Статистический анализ данных с пропусками. М.: Финансы и статистика, 1990. 336 с.
7. Мажукин В.И., Королева О.Н. Математическое моделирование в экономике. – М.: «Флинта: МГУ», 2004 – 232 с.
8. Павленков М. Н., Маева Л. С. Алгоритм оценки полноты и достоверности информационного обеспечения прогнозирования объемов твердых коммунальных отходов // Контроллинг. 2018, №4(70), С. 64-71.
9. Павленков М.Н., Парамонов А.В., Павленков И.М. Анализ показателей среднесрочного прогнозирования// Инновации и инвестиции. 2023, №10. С.554-557.
10. Россиев А. А. Моделирование данных при помощи кривых для восстановления пробелов в таблицах. Методы нейроинформатики / Под. ред. А.Н. Горбаня. – Красноярск: КГТУ, 1998. С. 6-22.
11. Черепанов Е. В. Нетрадиционные вероятностно-статистические методы для социально-экономических и социологических исследований: Монография. – М.: Издательство «Спутник+», 2012. 215 с.
12. Liao T. W. Clustering of time series data—a survey / T. W. Liao // Pattern Recognition. 2005, Vol. 38(11), P.1857-1874.

Formation of the information base of strategic planning of the municipality in the conditions of incomplete data

Mikhail N. Pavlenkov

Doctor of Economics, Associate Professor,
Nizhny Novgorod State University named after. N.I. Lobachevsky,
Dzerzhinsky branch,
606037, 17, Budennogo str., Dzerzhinsk, Russian Federation;
e-mail: kaf-fin-df@yandex.ru

Aleksandr V. Paramonov

Candidate of Economic Sciences, Associate Professor,
Nizhny Novgorod Institute of Management of the Russian Academy
of National Economy and Public Administration,
603950, 46, Gagarin ave., Nizhny Novgorod, Russian Federation;
e-mail: paramonov-nn@mail.ru

Abstract

The article discusses the issues of forming an information base for strategic planning of a municipality in the conditions of incompleteness of retrospective data. In managing the activities of a municipality, high-quality data is required, including when forecasting indicators of territory development. The main reasons for incomplete data are: omissions; data hiding; unreliability of data. To provide information support for forecasting indicators of territory development, a method for assessing the reliability of data has been developed. The method, using a synthesis of estimates by period and source of occurrence, provides an integrated assessment of the reliability of the source data, which allows an adequate assessment of the source data. To implement the method, the “Data Recovery” model and the “Time Series Formation” model were developed. A model and algorithm for data recovery using polygram evaluation have been developed. The model allows you to qualitatively process data incompleteness and generate a matrix of reconstructed data. To form a time series, expert methods are used. A model for the formation of time series components in the context of multiple sources of their occurrence has been developed, which differs from traditional methods in that the data is assessed by experts twice using the modified Harrington scale, which improves the quality of the initial data for predicting the development of the territory of the municipality.

For citation

Pavlenkov I.M., Paramonov A.V. (2023) Formirovanie informatsionnoi bazy strategicheskogo planirovaniya munitsipal'nogo obrazovaniya v usloviyakh nekomplektnosti dannykh [Formation of the information base of strategic planning of the municipality in the conditions of incomplete data]. *Ekonomika: vchera, segodnya, zavtra* [Economics: Yesterday, Today and Tomorrow], 13 (11A), pp. 466-473. DOI: 10.34670/AR.2023.11.60.051

Keywords

Municipal formation, management, forecasting, retrospective data, incomplete data, expert methods.

Referecens

1. Cherepanov E.V. Non-traditional probabilistic and statistical methods for socio-economic and sociological research: Monograph. – M.: Publishing house “Sputnik+”, 2012. 215 p.
2. Demidenko E.Z. Linear and nonlinear regression. M.: Finance and Statistics, 1981. 304 p.
3. Kruglov V.V., Abramenkova I.V. Methods for restoring gaps in data arrays // Software products and systems. 2005.- No. 2. P. 59-63.
4. Kruglov V.V., Borisov V.V. Artificial neural networks. Theory and practice. M.: Hotline - Telecom. 2002. 382 p.
5. Liao T. W. Clustering of time series data—a survey / T. W. Liao // Pattern Recognition. 2005, Vol. 38(11), P.1857-1874.
6. Little R.J.A., Rubin D.B. Statistical analysis of missing data. M.: Finance and Statistics, 1990. 336 p.
7. Mazhukin V.I., Koroleva O.N. Mathematical modeling in economics. – M.: “Flint: MSU”, 2004 – 232 p.
8. Pavlenkov M.N., Maeva L.S. Algorithm for assessing the completeness and reliability of information support for forecasting volumes of solid municipal waste // Controlling. 2018, No. 4(70), pp. 64-71.
9. Pavlenkov M.N., Paramonov A.V., Pavlenkov I.M. Analysis of medium-term forecasting indicators // Innovations and investments. 2023, no. 10. P.554-557.
10. Rossiev A. A. Data modeling using curves to restore gaps in tables. Methods of neuroinformatics / Ed. ed. A.N. Gorbanya. – Krasnoyarsk: KSTU, 1998. P. 6-22.
11. Zagoruiko N. G., Elkina V. N., Timerkaev V. S. Algorithm for filling gaps in empirical tables (Zet algorithm) // Empirical prediction and pattern recognition. Vol. 61: Computing systems. Novosibirsk, 1975. P. 3-27.
12. Zakharova A. A. New models of decision-making on the socio-economic development of the city // Management in Russia and abroad. 2006. No. 1.