

Алгоритмы многопараметрического анализа состояния многоотраслевого производственного комплекса регионов на основе интеллектуальных технологий

Б.Г. Ильясов

Факультет информатики и робототехники
Уфимский государственный авиационный
технический университет
Уфа, Россия
e-mail: ilyasov@tc.ugatu.ac.ru

Е. А. Макарова

Факультет информатики и робототехники
Уфимский государственный авиационный
технический университет
Уфа, Россия
e-mail: ea-makarova@mail.ru

Э.Р. Габдуллина

Факультет информатики робототехники
Уфимский государственный авиационный
технический университет
г. Уфа, Россия
e-mail: gabdullina_er@mail.ru

М. Т. Мансурова

Факультет информатики и робототехники
Уфимский государственный авиационный
технический университет
Уфа, Россия
e-mail: margaritamansurova1996@gmail.com

Аннотация¹

Проведен компонентный, кластерный и нейросетевой анализ состояния региональных многоотраслевых производственных комплексов (РМПК) России с учетом состояния добывающей и обрабатывающей отраслей. Построены диаграммы рассеивания в пространстве главных компонент. Сформулированы правила отнесения объектов к кластерам. Скорректированы границы кластеров и сформулированы уточненные после проведения кластерного анализа правила. Построены карты Кохонена. По результатам анализа сделан вывод о неравномерном развитии РМПК России в рассматриваемых направлениях.

1. Введение

Реальный сектор национальной экономики отличает зависимость от нефтегазовых доходов и неоднородная отраслевая структура, в которой доминантными являются сырьевые отрасли. Стабилизация реального сектора, его диверсификация и обеспечение роста производства в обрабатывающих отраслях являются ключевыми факторами создания устойчивого экономического роста [1-3]. Ввиду сложившейся ситуации для поддержания конкурентоспособности национальной экономики необходимо уделять пристальное

внимание развитию обрабатывающей промышленности, а также поддержанию добывающей промышленности. Особо актуальными являются вопросы мониторинга состояния этих отраслей и разработки программ государственной поддержки. По данным опроса Российского союза промышленников и предпринимателей чуть менее половины компаний-членов РСПП из всех федеральных округов прибегают к услугам институтов развития или инструментам государственной поддержки [1].

Исследования состояния реального сектора регионов России, в том числе и региональных многоотраслевых производственных комплексов, проводятся на кафедре технической кибернетики УГАТУ [4-7].

Под сложным региональным многоотраслевым производственным комплексом (РМПК) понимается множество взаимосвязанных по ресурсным потокам (материальным, финансовым, трудовым) производственных комплексов, функционирование которых может рассматриваться на федеральном или региональном уровнях. Процесс функционирования сложного РМПК рассматривается в тесной взаимосвязи с функционированием других секторов экономики, формирующим кругооборот финансовых и материальных потоков и образующим в целом воспроизводственный процесс. Для исследования состояния РМПК применяются различные методы: методы имитационного динамического моделирования и интеллектуальные технологии. Методы интеллектуального анализа данных (ИАД) являются одним из применяемых инструментариев

Труды Седьмой всероссийской научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 28-30 мая, Уфа-Ставрополь, Ханты-Мансийск, Россия, 2019

VII Всероссийская научная конференция "Информационные технологии интеллектуальной поддержки принятия решений", Уфа-Ставрополь-Ханты-Мансийск, Россия, 2019

поддержки принятия решений в рассматриваемой предметной области [8,9].

Для проведения анализа данных о состоянии предприятий РМПК России отобраны исходные данные на основе сборника «Регионы России. Социально-экономические показатели» за 2017, который опубликован на сайте Федеральной службы государственной статистики [2].

При проведении ИАД применены методы компонентного, кластерного и нейросетевого анализа, при этом использовались следующие признаки, касающиеся добывающей (ДП) и обрабатывающей (ОП) промышленности в регионах РФ: распределение числа предприятий по видам экономической деятельности на конец 2017 г. (число предприятий); оборот организаций по видам экономической деятельности к началу 2017 г.; объем отгруженных товаров собственного производства, выполненных работ и услуг собственными силами по видам экономической деятельности в 2017 году (объем отгруженных товаров); индекс производства по видам экономической деятельности в 2017 году.

2. Результаты компонентного анализа

В процессе проведения компонентного анализа исходной выборки в пакете *STATGRAPHICS* выбраны первые три главные компоненты (ГК), так как они описывают более 85% дисперсии [8,9]. В таблице 1 представлены весовые коэффициенты ГК.

По таблице 1 определены значащие признаки, участвующие в названии первых двух ГК: объем отгруженных товаров ОП, число предприятий ДП, оборот организаций ОП, число предприятий ОП, оборот организаций ДП, объем отгруженных товаров ДП.

По значащим признакам сформулированы следующие названия главных компонент: ГК 1 – уровень развития ОП в регионе с учетом распределения числа ДП; ГК 2 – оборот организаций и объем отгруженных товаров ДП.

Диаграмма рассеивания в пространстве ГК 1 и ГК 2, полученная при проведении компонентного анализа исходной выборки о состоянии добывающей и обрабатывающей промышленности регионов России, представлена на рисунке 1.

На диаграмме рассеивания особенно выделяется объект «г. Москва». Он отличается большим значением первой ГК и может быть выделен в отдельный кластер.

Значения признаков, которыми характеризуется объект «г. Москва»: очень большой объем отгруженных товаров ОП, очень большое число предприятий ДП, очень большой оборот организаций ОП, очень большое число предприятий ОП, малый оборот организаций ДП, малый объем отгруженных товаров ДП.

Алгоритмы многопараметрического анализа состояния многоотраслевого производственного комплекса регионов на основе интеллектуальных технологий

Для лучшей интерпретируемости структуры всего множества РМПК регионов объект «г. Москва» исключен из исходной выборки, в результате получена диаграмма рассеивания в пространстве ГК 1 и ГК 2, представленная на рисунке 2.

Табл. 1. Таблица весовых коэффициентов главных компонент

| Весовые коэффициенты | | | |
|------------------------------|----------|----------|----------|
| | ГК 1 | ГК 2 | ГК 3 |
| Число предприятий ДП | 0,458527 | 0,041014 | -0,02998 |
| Число предприятий ОП | 0,448811 | -0,28096 | 0,062596 |
| Оборот организаций ДП | 0,228446 | 0,656935 | 0,047239 |
| Оборот организаций ОП | 0,458075 | -0,25968 | 0,064201 |
| Объем отгруженных товаров ДП | 0,279455 | 0,613062 | 0,058713 |
| Объем отгруженных товаров ОП | 0,471201 | -0,1761 | 0,067239 |
| Индекс производства ДП | -0,14578 | 0,10628 | 0,241816 |
| Индекс производства ОП | -0,07016 | -0,04727 | 0,96041 |

По диаграмме рассеивания, представленной на рисунке 3, сформулированы правила отнесения объектов к кластерам в пространстве главных компонент.

Правило 1: ЕСЛИ ГК 1 = малая И ГК 2 = средняя, ТО Кластер = 1;

Правило 2: ЕСЛИ ГК 1 = ниже среднего И ГК 2 = средняя, ТО Кластер = 2;

Правило 3: ЕСЛИ ГК 1 = средняя И ГК 2 = средняя, ТО Кластер = 3;

Правило 4: ЕСЛИ ГК 1 = выше среднего И ГК 2 = средняя, ТО Кластер = 4;

Правило 5: ЕСЛИ ГК 1 = высокая И ГК 2 = низкая, ТО Кластер = 5;

Правило 6: ЕСЛИ ГК 1 = высокая И ГК 2 = высокая,
ТО Кластер = 6.

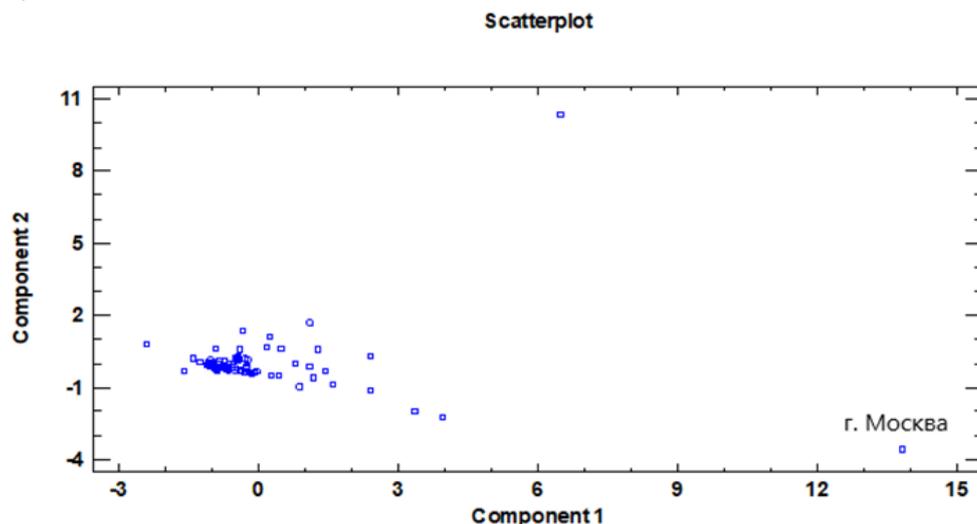


Рис. 1. Диаграмма рассеивания в пространстве первой и второй ГК (полная выборка)

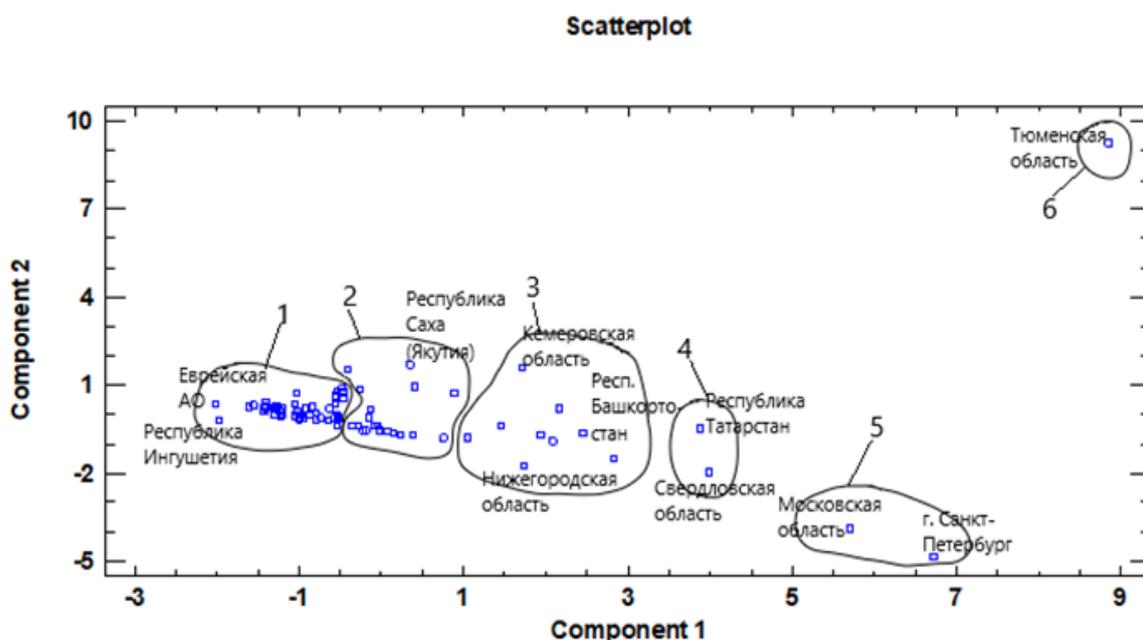


Рис. 2. Диаграмма рассеивания в пространстве первой и второй ГК (усеченная выборка)

Кроме того, сформулированы уточненные правила отнесения объектов к кластерам в признаковом пространстве. Уточненное правило для наиболее многочисленного кластера 1 представлено ниже.

Правило 1: ЕСЛИ объем отгруженных товаров ОП = малый И число предприятий ДП = малое И оборот организаций ОП = малый И число предприятий ОП = малое И оборот организаций ДП = средний И объем отгруженных товаров ДП = средний, ТО Кластер = 1.

3. Результаты кластерного анализа

Для уточнения результатов, полученных при проведении компонентного анализа, в пакете *STATGRAPHICS* проведен кластерный анализ усеченной выборки. При проведении кластерного

анализа использовался метод Варда, метрика *City-Block*. Координаты центроидов кластеров представлены на рисунке 3. Дендрограмма представлена на рисунке 4.

По координатам центроидов кластеров видно, что признаки индекс производства ДП и индекс производства ОП не являются значимыми и при формировании правил отнесения объектов к кластерам не участвуют

По результатам кластерного анализа определены уточненные границы кластеров, диаграмма рассеивания в пространстве ГК 1 и ГК 2 представлена на рисунке 5..

Centroids

| Cluster | number_of_ent_(DP) | number_of_ent_(OP) | turnover_(DP) | turnover_(OP) | volume_(DP) |
|---------|--------------------|--------------------|---------------|---------------|-------------|
| 1 | 121,542 | 3641,58 | 22,8875 | 369,35 | 15241,3 |
| 2 | 77,4839 | 1236,13 | 24,8645 | 67,3903 | 32443,4 |
| 3 | 411,5 | 24156,0 | 64,4 | 2720,25 | 16943,5 |
| 4 | 347,462 | 1791,85 | 315,454 | 138,569 | 313867, |
| 5 | 428,778 | 8326,11 | 253,411 | 1095,88 | 231786, |
| 6 | 1004,0 | 5441,0 | 5663,3 | 810,1 | 5,06891E6 |

| Cluster | volume_(OP) | index_(DP) | index_(OP) |
|---------|-------------|------------|------------|
| 1 | 453477, | 104,938 | 105,129 |
| 2 | 97765,8 | 114,577 | 110,958 |
| 3 | 2,27469E6 | 107,75 | 109,3 |
| 4 | 175408, | 101,823 | 101,238 |
| 5 | 1,20335E6 | 99,8889 | 104,633 |
| 6 | 1,56861E6 | 102,1 | 103,6 |

Рис. 3. Координаты центроидов кластеров

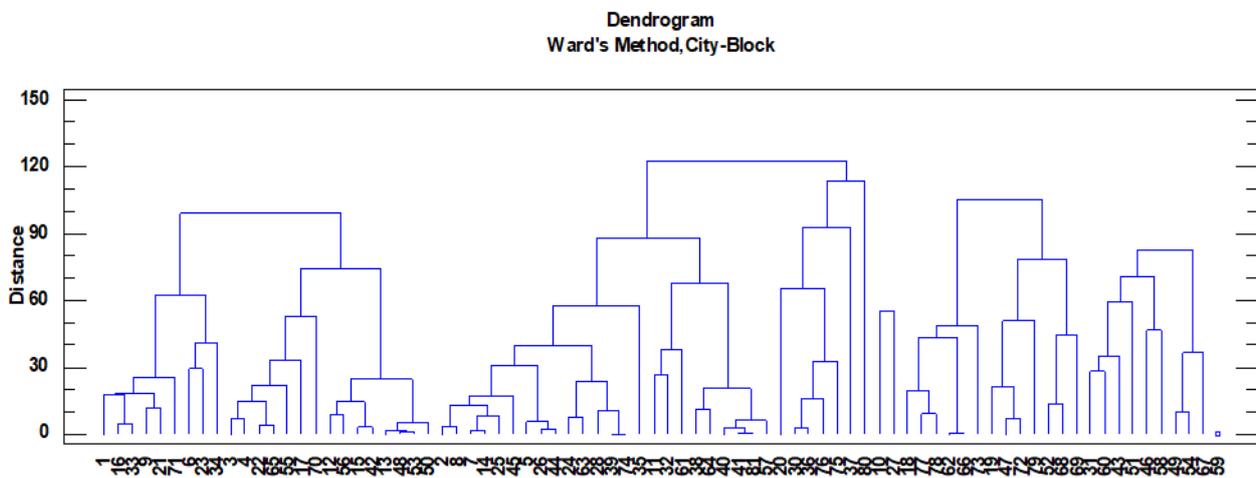


Рис. 4. Дендрограмма

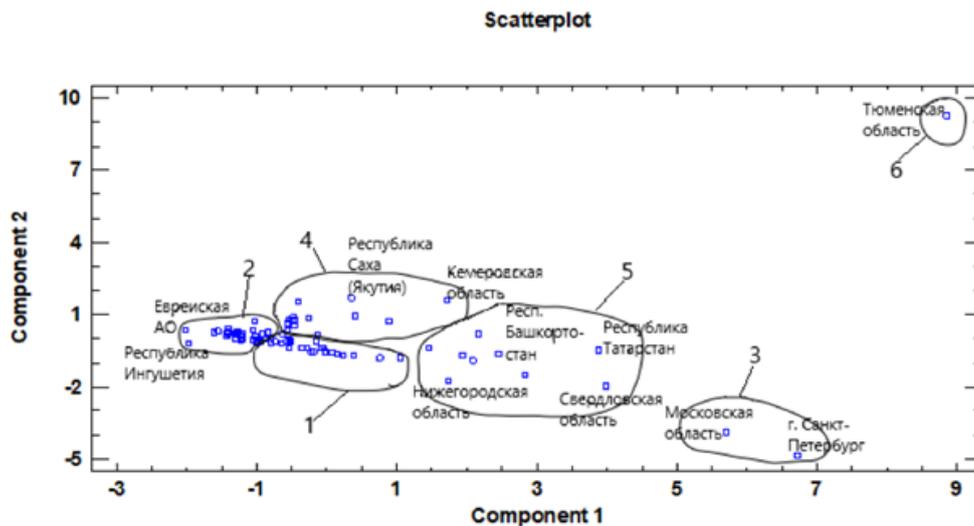


Рис. 5. Диаграмма рассеивания, иллюстрирующая уточненные в ходе проведения кластерного анализа границы кластеров

Алгоритмы многопараметрического анализа состояния многоотраслевого производственного комплекса регионов на основе интеллектуальных технологий

Уточненные правила отнесения объектов к кластерам в признаковом пространстве, сформулированные при проведении кластерного анализа, представлены ниже.

Правило 1: ЕСЛИ объем отгруженных товаров ОП = ниже среднего И число предприятий ДП = ниже среднего И оборот организаций ОП = ниже среднего И число предприятий ОП = ниже среднего И оборот организаций ДП = малый И объем отгруженных товаров ДП = малый, ТО Кластер = 1.

Правило 2: ЕСЛИ объем отгруженных товаров ОП = малый И число предприятий ДП = малое И оборот организаций ОП = малый И число предприятий ОП = малое И оборот организаций ДП = малый И объем отгруженных товаров ДП = малый, ТО Кластер = 2.

Правило 3: ЕСЛИ объем отгруженных товаров ОП = большой И число предприятий ДП = большое И оборот организаций ОП = большой И число предприятий ОП = большое И оборот организаций ДП = малый И объем отгруженных товаров ДП = малый, ТО Кластер = 3.

Правило 4: ЕСЛИ объем отгруженных товаров ОП = ниже среднего И число предприятий ДП = ниже среднего И оборот организаций ОП = ниже среднего И число предприятий ОП = ниже среднего И оборот организаций ДП = малый И объем отгруженных товаров ДП = малый, ТО Кластер = 4.

организаций ДП = средний И объем отгруженных товаров ДП = средний, ТО Кластер = 4.

Правило 5: ЕСЛИ объем отгруженных товаров ОП = средний И число предприятий ДП = среднее И оборот организаций ОП = средний И число предприятий ОП = среднее И оборот организаций ДП = малый И объем отгруженных товаров ДП = малый, ТО Кластер = 5.

Правило 6: ЕСЛИ объем отгруженных товаров ОП = высокий И число предприятий ДП = высокое И оборот организаций ОП = высокий И число предприятий ОП = высокое И оборот организаций ДП = высокий И объем отгруженных товаров ДП = высокий, ТО Кластер = 6.

Результаты кластерного анализа незначительно отличаются от результатов компонентного по составам кластеров и извлеченным правилам. Однако в целом извлеченные правила сходны.

4. Результаты нейросетевого анализа (построение карт Кохонена)

Для усеченной выборки, которая не включает в себя объект «г. Москва», в программе *Deductor Studio Academic* построены карты Кохонена, представленные на рисунке 6.

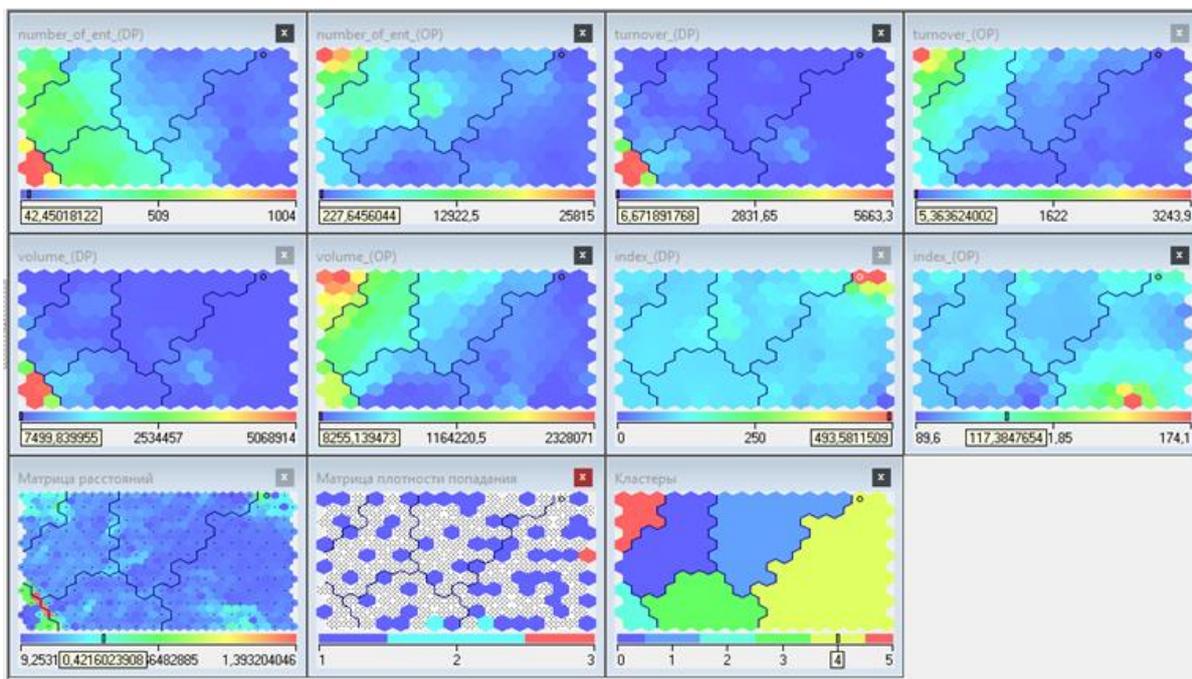


Рис. 6. Карты Кохонена

По картам Кохонена сформулированы правила отнесения объектов к кластерам. Признаки индекс производства ДП и индекс производства ОП в построении правил не участвуют, так как они являются незначимыми.

Правило 1: ЕСЛИ число предприятий ДП = среднее И число предприятий ОП = ниже среднего И оборот

организаций ДП = малый И оборот организаций ОП = ниже среднего И объем отгруженных товаров ДП = малый И объем отгруженных товаров ОП = средний, ТО Кластер = 1.

Правило 2: ЕСЛИ число предприятий ДП = малое И число предприятий ОП = малое И оборот организаций ДП = малый И оборот организаций ОП =

малый И объем отгруженных товаров ДП = ниже среднего И объем отгруженных товаров ОП = ниже среднего, ТО Кластер = 2.

Правило 3: ЕСЛИ число предприятий ДП = большое И число предприятий ОП = ниже среднего И оборот организаций ДП = большой И оборот организаций ОП = ниже среднего И объем отгруженных товаров ДП = большой И объем отгруженных товаров ОП = средний, ТО Кластер = 3.

Правило 4: ЕСЛИ число предприятий ДП = среднее И число предприятий ОП = малое И оборот организаций ДП = малый И оборот организаций ОП = малый И объем отгруженных товаров ДП = малый И объем отгруженных товаров ОП = малый, ТО Кластер = 4.

Правило 5: ЕСЛИ число предприятий ДП = малое И число предприятий ОП = малое И оборот организаций ДП = малый И оборот организаций ОП = малый И объем отгруженных товаров ДП = малый И объем отгруженных товаров ОП = малый, ТО Кластер = 5.

Правило 6: ЕСЛИ число предприятий ДП = среднее И число предприятий ОП = выше среднего И оборот организаций ДП = малый И оборот организаций ОП = выше среднего И объем отгруженных товаров ДП = малый И объем отгруженных товаров ОП = высокий, ТО Кластер = 6.

На картах Кохонена некоторые объекты представляют особый интерес.

Объект «г. Санкт-Петербург» отличается наибольшим числом предприятий ОП и наибольшим значением оборота организаций ОП при большом значении объема отгруженных товаров ОП и низком уровне развития ДП. Объект «Тюменская область» отличается наибольшим числом предприятий ДП, наибольшим значением оборота организаций ДП и наибольшим значением объема отгруженных товаров при уровне развития ОП ниже среднего.

5. Заключение

В результате проведения компонентного, кластерного анализов и построения карт Кохонена выявлены следующие особенности РМПК России: обрабатывающая промышленность наиболее развита в г. Москва, в г. Санкт-Петербург и в Московской области; наивысший уровень развития добывающей промышленности достигнут в Тюменской области; небольшая группа регионов (Республика Башкортостан, Республика Татарстан, Свердловская область и др.) отличаются средним уровнем развития добывающей и обрабатывающей промышленности; большая часть регионов характеризуется невысоким уровнем развития обрабатывающей промышленности и средним уровнем развития добывающей промышленности.

Полученные результаты целесообразно применять при формировании интеллектуальной системы поддержки принятия решений при управлении развитием региональных производственных систем.

Благодарности

Работа выполнена при финансовой поддержке РФФИ, проект № 17-08-01155.

Список используемых источников

1. Российский союз промышленников и предпринимателей [Электронный ресурс] – URL: <http://www.rspp.ru> (дата обращения: 10.04.2019).
 2. Федеральная служба государственной статистики [Электронный ресурс] – URL: <http://www.gks.ru> (дата обращения: 07.02.2019).
 3. Замаев Б., Маршова Т. Инвестиционные процессы и структурная перестройка российской экономики // Вопросы экономики. 2017. № 12. С. 40-62. издат., 1996. – 558 с.
 4. Димов Э. М., Ильясов Б. Г., Макарова Е. А., Закиева Е. Ш., Ефтонова Т. А., Гиздатуллина Э. С. Методология системного динамического моделирования и управления функционированием многоотраслевого производственного комплекса в рамках воспроизводственного процесса макроуровня. // Инфокоммуникационные технологии. Самара: Изд. ПГУТИ, 2018. Том 16, №1. С. 81-96.
 5. Е. А. Макарова. Динамические модели функционирования экономических агентов и их взаимодействия в рамках воспроизводственного процесса с учетом запасов капитала / Инфокоммуникационные технологии. / Том 13, №2. 2015. С. 164-176. (№ 1042 перечня российских рецензируемых научных журналов)
 6. Ильясов Б. Г., Димов Э. М., Макарова Е. А., Ефтонова Т. А. Динамическая модель функционирования сложного многоотраслевого производственного комплекса с учетом запасов капитала и воспроизводственных взаимосвязей. // Инфокоммуникационные технологии. Том 13, №3. 2015. С. 281-290.
 7. Экономическая политика государства: новые факторы и механизмы реализации: монография / колл. авт.; отв. ред. О.Е. Малых. – Сыктывкар: ГОУ ВО КРАГСИУ, 2015. – 246 с.
 8. Кулаичев А. П. Методы и средства комплексного анализа данных: учебное пособие. 4-е изд., перераб. и доп. – М.: ФОРУМ: ИНФРА-М, 2013. – 312 с.
- Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям. Учебное пособие / 2-е изд., исправ.. – СПб.: Питер, 2013. – 704 с.