

Применение метода главных компонент и метода собственных состояний для анализа и прогнозирования социально-экономических систем

**Мокеев В.В., заведующий
кафедрой информационных
систем**

Метод главных компонент

Задачи, решаемые МГК, можно разделить на 4 класса:

1. Аппроксимация данных линейными многообразиями меньшей размерности;
2. Поиск подпространства меньшей размерности, в ортогональной проекции на которые разброс данных максимален;
3. Поиск подпространства меньшей размерности, в ортогональной проекции на которые среднееквадратичное расстояние между точками максимально;
4. Построение для многомерной случайной величины такого ортогонального преобразования координат, в результате которого корреляции между отдельными координатами обратятся в ноль.

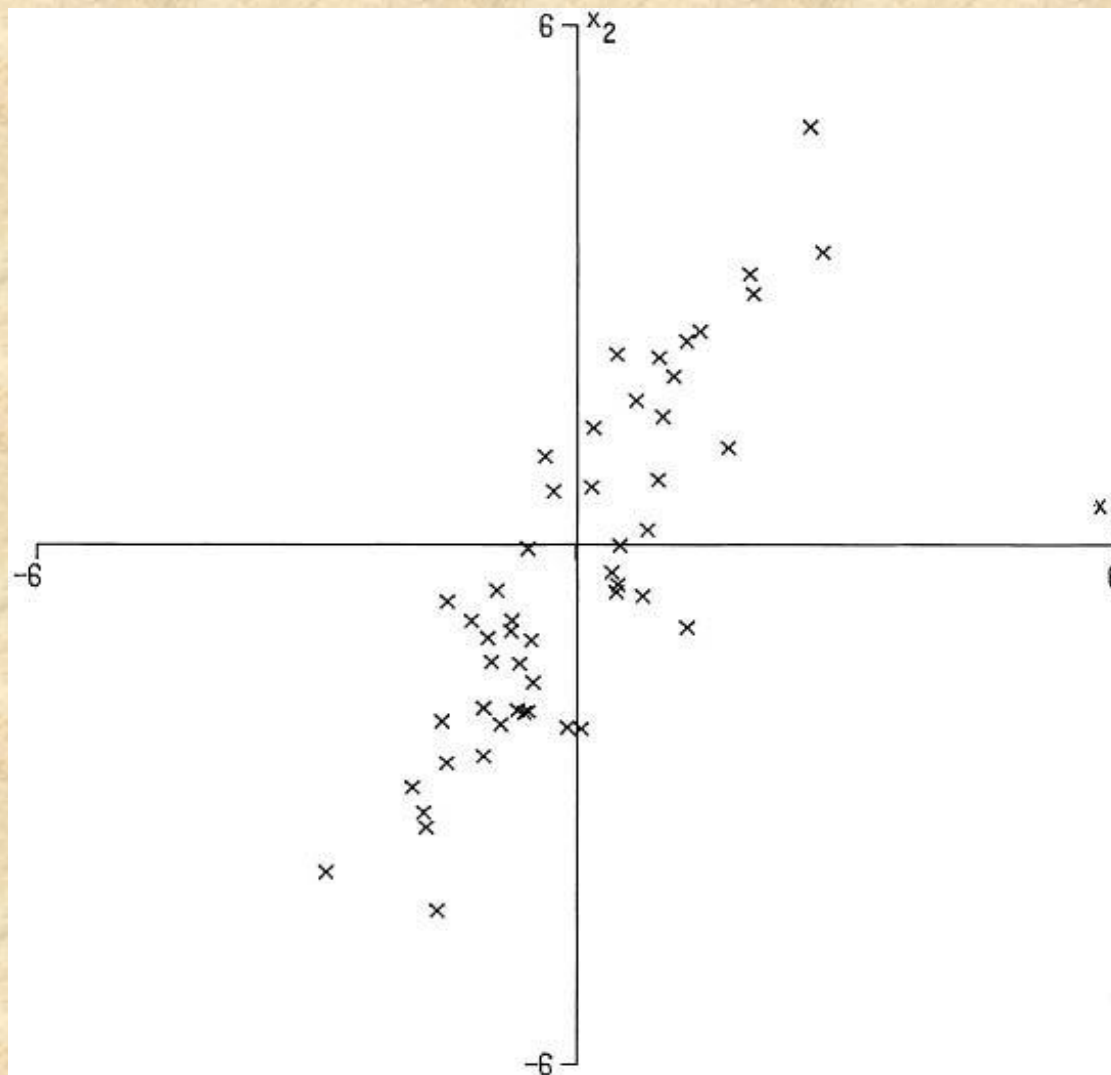
ГЛАВНЫЕ КОМПОНЕНТЫ

Основная идея решения задач этих четырех классов является поиск таких линейных комбинации показателей, которые бы имели бы имеющей максимальную дисперсию (разброс данных)

$$z_1 = \alpha_1^T x = \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1p}x_p = \sum_{j=1}^p \alpha_{1j}x_j$$

Найденные комбинации показателей представляют новые показатели и называются главными компонентами

ПРОСТОЙ ПРИМЕР



50 наблюдений переменных X_1 И X_2 .

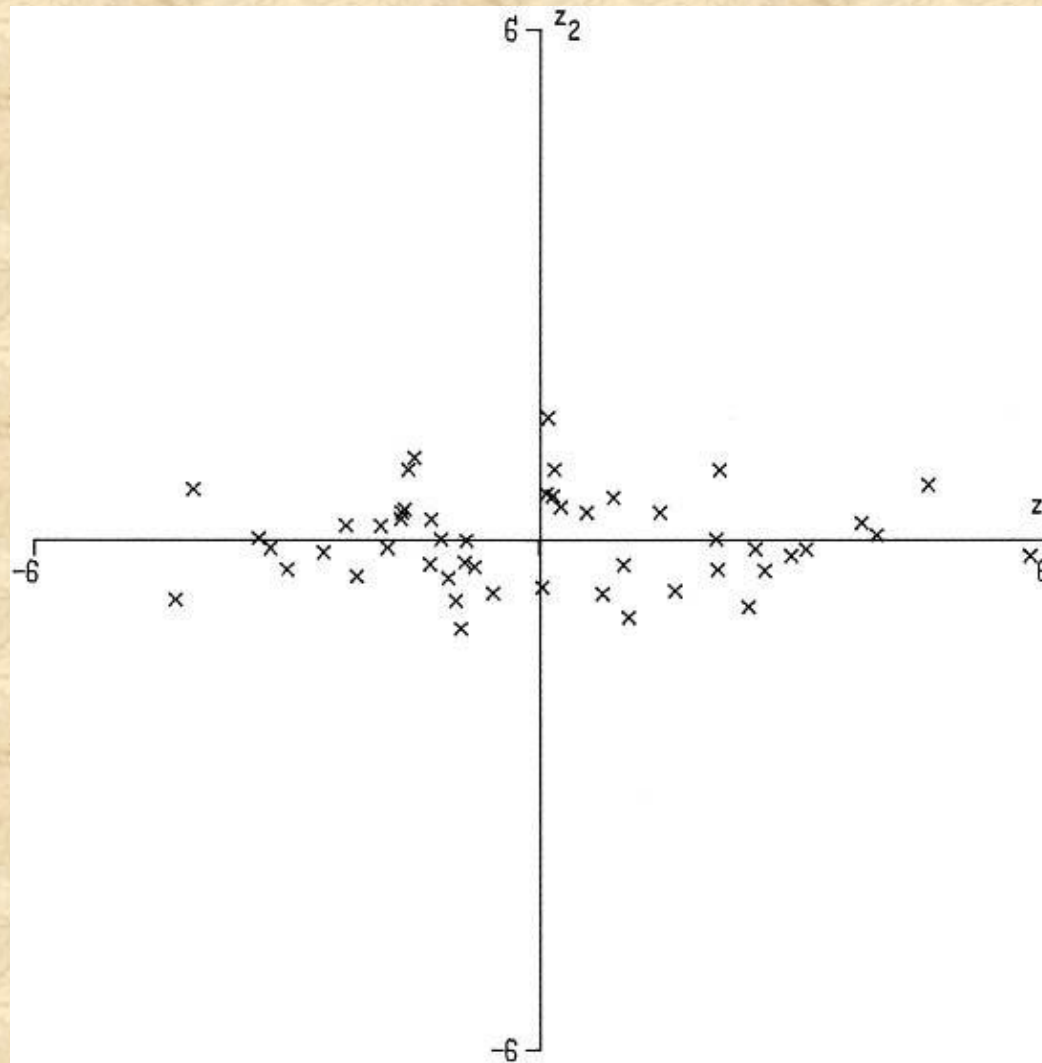
ГЛАВНЫЕ КОМПОНЕНТЫ

Комбинируя переменные, мы можем найти две комбинации, имеющих разные дисперсии

$$z_1 = \alpha_{11}(x_1 - \bar{x}_1) + \alpha_{12}(x_2 - \bar{x}_2)$$

$$z_2 = \alpha_{21}(x_1 - \bar{x}_1) + \alpha_{22}(x_2 - \bar{x}_2)$$

ГЛАВНЫЕ КОМПОНЕНТЫ



50 наблюдений переменных Z_1 И Z_2 .

ГЛАВНЫЕ КОМПОНЕНТЫ

В рамках решения задачи поиска подпространства меньшей размерности, в ортогональной проекции на которые разброс данных максимален, переменные x можно описать с помощью первой главной компоненты

$$\begin{aligned}x_1 &= \bar{x}_1 + \alpha_{11}z_1 \\x_2 &= \bar{x}_2 + \alpha_{12}z_1\end{aligned}\tag{1}$$

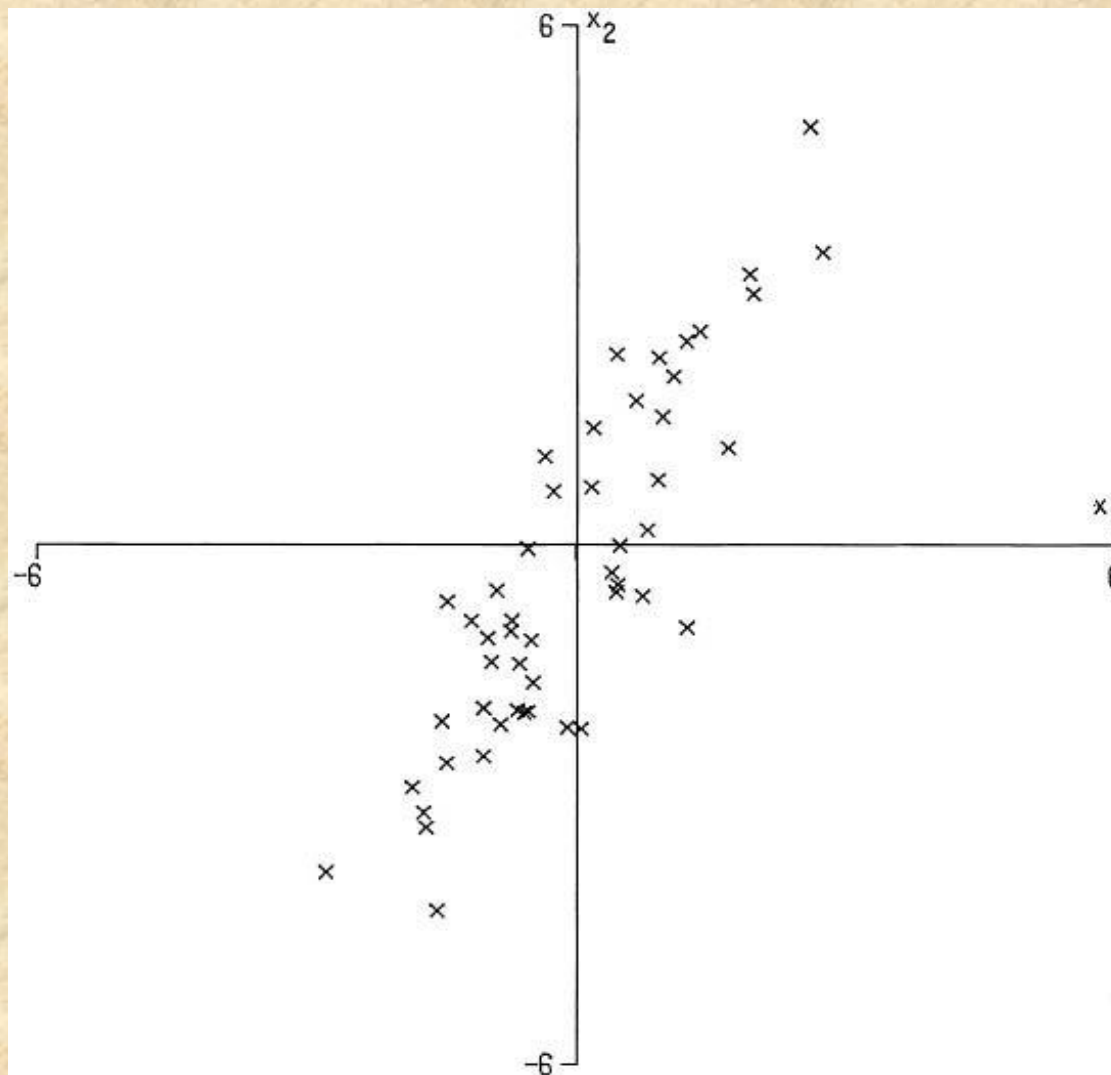
ГЛАВНЫЕ КОМПОНЕНТЫ

Используя описания переменных в пространстве меньшей размерности можно прогнозировать значение показателей.

Например, если известно значение x_1
то используя первое уравнение системы (1)
получим

$$z_1 = -(x_1 - \bar{x}_1) / \alpha_{11}$$

Главные компоненты



50 наблюдений переменных x_1 и x_2 .

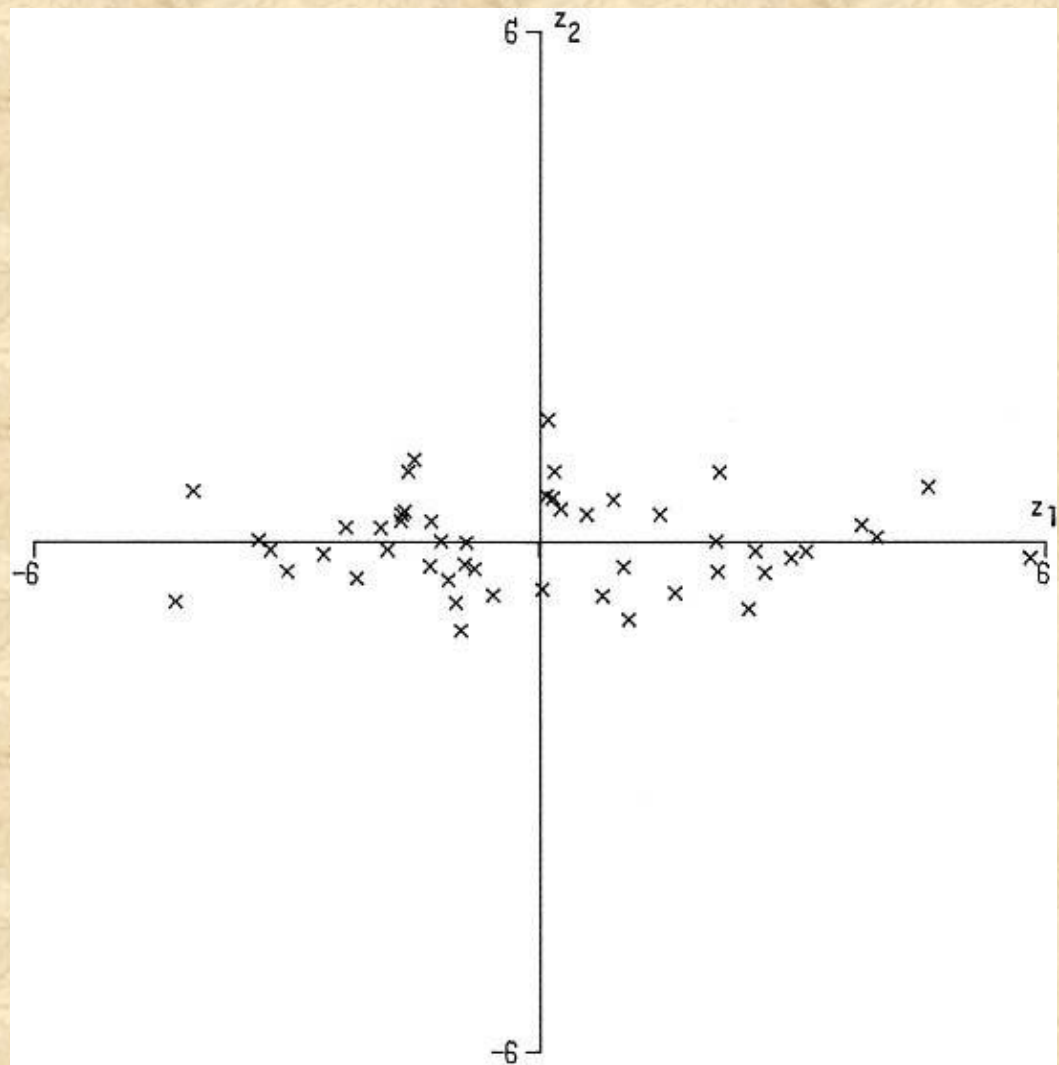
Главные компоненты

Комбинируя переменные, мы можем найти две комбинации имеющих разные дисперсии

$$z_1 = \alpha_{11}(x_1 - \bar{x}_1) + \alpha_{12}(x_2 - \bar{x}_2)$$

$$z_2 = \alpha_{21}(x_1 - \bar{x}_1) + \alpha_{22}(x_2 - \bar{x}_2)$$

Главные компоненты



50 наблюдений переменных z_1 и z_2 .

Главные компоненты

Переменные x можно достаточно хорошо описать с помощью первой главной компоненты.

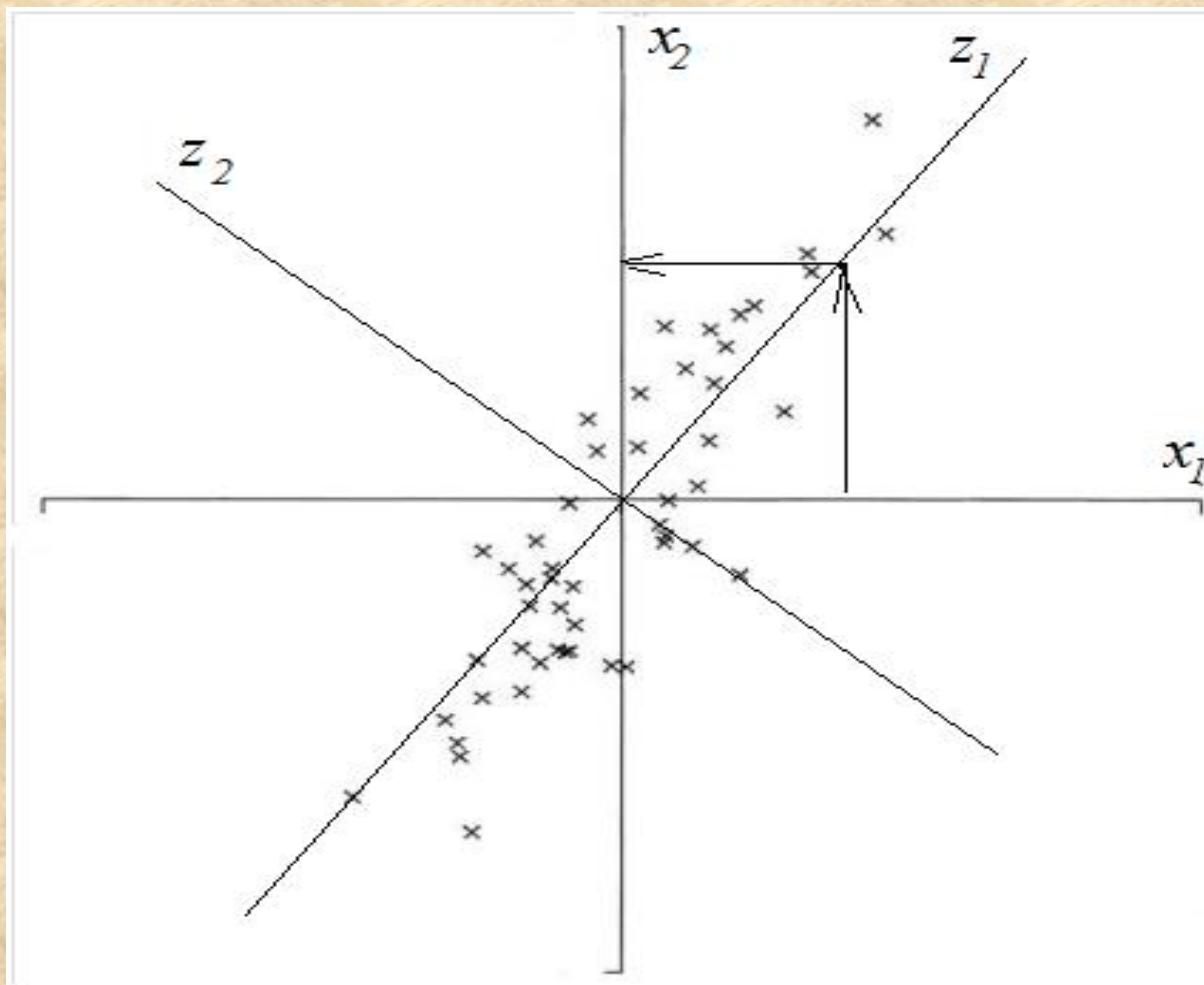
$$\begin{aligned}x_1 &= \bar{x}_1 + \alpha_{11}z_1 \\x_2 &= \bar{x}_2 + \alpha_{12}z_1\end{aligned}\tag{1}$$

Если мы знаем значение переменной x_1

то используя первое уравнение системы (1) получим

$$z_1 = -(x_1 - \bar{x}_1) / \alpha_{11}$$

Главные компоненты



Прогнозирование по главным компонентам

Вычисление главных компонент

Основной задачей при **вычислении главных компонент** является определение весовых коэффициентов новых переменных.

Для этого решается задача собственных значений

$$(\mathbf{A} - \lambda \mathbf{I}) \mathbf{v}_0 = 0$$

Здесь $\mathbf{A}$ – матрица разброса / ковариационная матрица

$\mathbf{V}_0 = [\mathbf{v}_{01}, \mathbf{v}_{02}, \dots, \mathbf{v}_{0p}]$ – матрица собственных векторов

Главные компоненты вычисляются по формуле

$$z_i = \mathbf{v}_{0i}^T \mathbf{x} = \sum_{j=1}^n \mathbf{v}_{0ij} x_j$$

МЕТОД СОБСТВЕННЫХ СОСТОЯНИЙ

Метод собственных состояний представляет развитие метода главных компонент применительно к анализу социально-экономических систем.

Однако в отличие от метода главных компонент **основной целью** метода собственных состояний является построение эталонной модели.

МЕТОД СОБСТВЕННЫХ СОСТОЯНИЙ

Для построения эталонной модели необходимо сформулировать **требования к эталонной деятельности** социально-экономической системы.

Эти требования могут накладывать ограничение на изменения некоторых (базовых) показателей.

Эталонная деятельность социально-экономической системы представляет деятельности, в рамках которой изменение базовых показателей полностью соответствует сформулированным требованиям.

МЕТОД СОБСТВЕННЫХ СОСТОЯНИЙ

Метод собственных состояний также как и в метод главных компонент формирует главные компоненты.

Но каждая главная компонента интерпретируется как **однофакторная модель**, описывающая некоторую тенденцию или процесс развития социально-экономической системы.

Таким образом, весовые коэффициенты главных компонент интерпретируются как характеристики собственного состояния социально-экономической системы.

МЕТОД СОБСТВЕННЫХ СОСТОЯНИЙ

Собственные состояния обладают двумя свойствами.

- **Свойство 1.** При описании состояния социально-экономической системы в виде взвешенной **суммы собственных состояний**, главные компоненты собственных состояний являются **независимыми**.
- **Свойство 2.** Изменения состояния социально-экономической системы, связанное с изменением главной компоненты i -го собственного состояния, может происходить только пропорционально коэффициентам i -го собственного состояния.

МЕТОД СОБСТВЕННЫХ СОСТОЯНИЙ

Метод собственных состояний выполняет построение модели по схеме аналогичной той, что использует метод главных компонент.

Однако схема **метода собственных состояний** имеет особенности.

При вычислении весовых коэффициентов собственных состояний наряду с ковариационной матрицей может использоваться матрица **начальных вторых моментов** и матрицы **меж- и внутриклассовых различий**.

ПОСТРОЕНИИ ЭТАЛОННОЙ МОДЕЛИ

Эталонная модель социально-экономической системы формируется в виде суммы собственных состояний, удовлетворяющим требованиям к эталонной деятельности.

Эталонные модели можно разделить на модели с постоянным и переменным числом собственных состояний.

- **Модель с постоянным числом собственных состояний** использует при вычислении значений показателей для каждого наблюдения одно и то же число собственных состояний

ПОСТРОЕНИИ ЭТАЛОННОЙ МОДЕЛИ

Модель с переменным числом собственных состояний меняет число собственных состояний для каждого наблюдения в зависимости от знака главной компоненты данного наблюдения.

Различаются три группы собственных состояний

- 1) группа постоянных собственных состояний;
- 2) группа собственных состояний с положительными главными компонентами;
- 3) группа собственных состояний с отрицательными главными компонентами.

Классификация социально-экономических систем

Классификация состояний социально-экономических систем осуществляется путем сравнения эталонной и фактической деятельности социально-экономической системы.

Для классификации используется комплексный индикатор, который вычисляется методом штрафных функций

Принципы сравнения делятся на

- нормативный,
- затратно - результатный.

Кластеризация социально-экономических систем

Кластеризация состояний социально-экономических систем предлагается осуществлять в пространстве штрафных функций, полученных при сравнения фактической и эталонной развития социально-экономической системы.

Для кластеризации может использоваться **метод k-средних**.

ПОСТРОЕНИЕ РЕЙТИНГА ПРЕДПРИЯТИЙ металлургической отрасли

Рейтинговый анализ развития предприятий выполняется для следующих предприятий

- Волжский трубный завод,
- Синарский трубный завод,
- Северский трубный завод,
- Таганрогский металлургический завод,
- Челябинский трубопрокатный завод,
- Первоуральский новотрубный завод,
- Выксунский металлургический завод,
- Альметьевский трубный завод,
- Борский трубный завод

ПОСТРОЕНИЕ РЕЙТИНГА ПРЕДПРИЯТИЙ металлургической отрасли

Обобщенный показатель финансовой устойчивости, который рассчитывается по формуле

$$\Phi_y = \sqrt[9]{K_{мл} * K_{фз} * K_n * K_{ав} * K_{фр} * K_m * K_{па} * M_{ск} * K_{осс}}$$

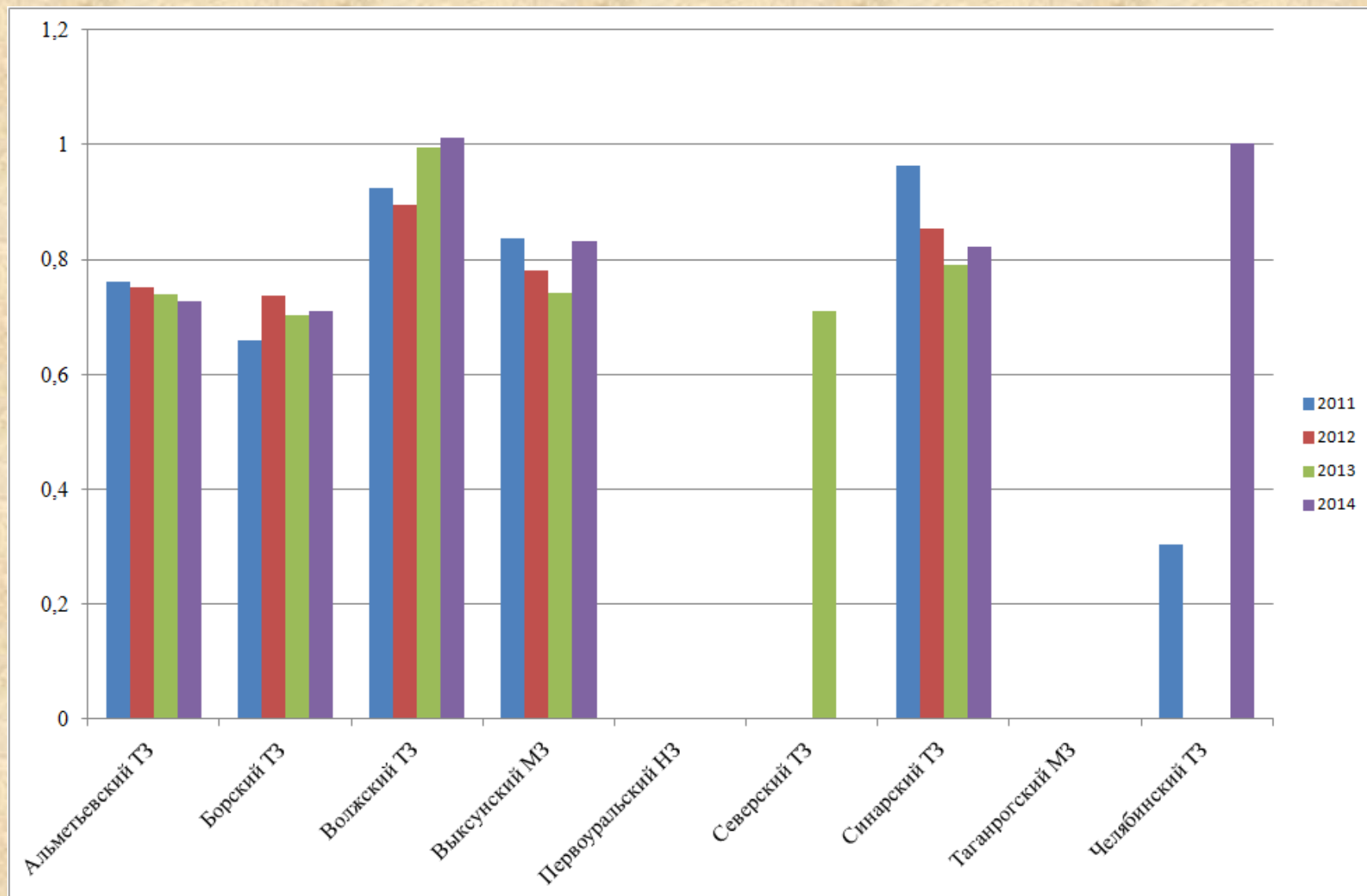
Для оценки риска банкротства в среднесрочной перспективе используется пятифакторную модель Кадыкова -Сайфулина

$$K = 2 * K_{осс} + 0,1 * K_{мл} + 0,08 * K_{инт} + 0,45 * K_{мен} + P_{ск}$$

Обобщенный показатель финансовой устойчивости фактической деятельности предприятий

Наименование	2011	2012	2013	2014
Альметьевский ТЗ	0,76	0,75	0,738	0,728
Борский ТЗ	0,66	0,74	0,704	0,711
Волжский ТЗ	0,93	0,89	0,994	1,013
Выксунский МЗ	0,84	0,78	0,742	0,831
<i>Первоуральский НЗ</i>	-	-	-	-
<i>Северский ТЗ</i>	-	-	0,71	-
Синарский ТЗ	0,96	0,85	0,791	0,823
<i>Таганрогский МЗ</i>	-	-	-	-
<i>Челябинский ТЗ</i>	0,304	-	-	1,002

Обобщенный показатель финансовой устойчивости фактической деятельности предприятий



Рейтинговый показатель Кадыкова-Сайфулина фактической деятельности предприятий

Наименование	2011	2012	2013	2014
Альметьевский ТЗ	2,17	1,39	1,30	1,23
Борский ТЗ	1,37	1,89	1,55	1,62
Волжский ТЗ	2,22	1,69	2,02	2,13
Выксунский МЗ	3,09	2,12	1,62	2,06
<i>Первоуральский НЗ</i>	0,39	-3,04	-2,30	-2,57
<i>Северский ТЗ</i>	0,48	0,47	1,22	-0,49
Синарский ТЗ	1,73	1,47	1,42	1,82
<i>Таганрогский МЗ</i>	-0,71	-1,63	-2,12	-1,64
<i>Челябинский ТЗ</i>	1,83	0,67	-0,19	1,40

Рейтинговый показатель Кадыкова-Сайфулина фактической деятельности предприятий



ПОСТРОЕНИЕ РЕЙТИНГА ПРЕДПРИЯТИЙ

методом собственных состояний

Финансовая устойчивость отражает такое состояние финансовых ресурсов, при котором организация, свободно маневрируя денежными средствами, способна путём их эффективного использования обеспечить бесперебойный процесс производства и реализации продукции (работ, услуг)

Эталонная модель строится из собственных состояний, вычисленных по матрице внутриклассовых различий пяти предприятий, которые показатели наилучшие финансовые результаты.

ПОСТРОЕНИЕ РЕЙТИНГА ПРЕДПРИЯТИЙ металлургической отрасли

В результате экспериментальных исследований для эталонной модели были выбраны пять собственных состояний. Из них

- одно собственное состояние относится к группе постоянных собственных состояний,
- три собственных состояния относятся к группе собственных состояний с положительными главными компонентами
- одно собственное состояние — к группе собственных состояний с отрицательными главными компонентами.

Обобщенный показатель финансовой устойчивости эталонной деятельности предприятий

Наименование	2011	2012	2013	2014
Альметьевский ТЗ	0,74	0,85	0,93	0,94
Борский ТЗ	0,74	0,74	0,74	0,81
Волжский ТЗ	0,96	0,96	0,99	1,01
Выксунский МЗ	0,80	0,80	0,80	0,81
Первоуральский НЗ	0,94	1,00	1,03	1,01
Северский ТЗ	0,86	0,86	0,93	0,93
Синарский ТЗ	0,86	0,86	0,87	0,90
Таганрогский МЗ	0,86	0,86	0,86	0,88
Челябинский ТЗ	1,02	1,09	1,10	1,12

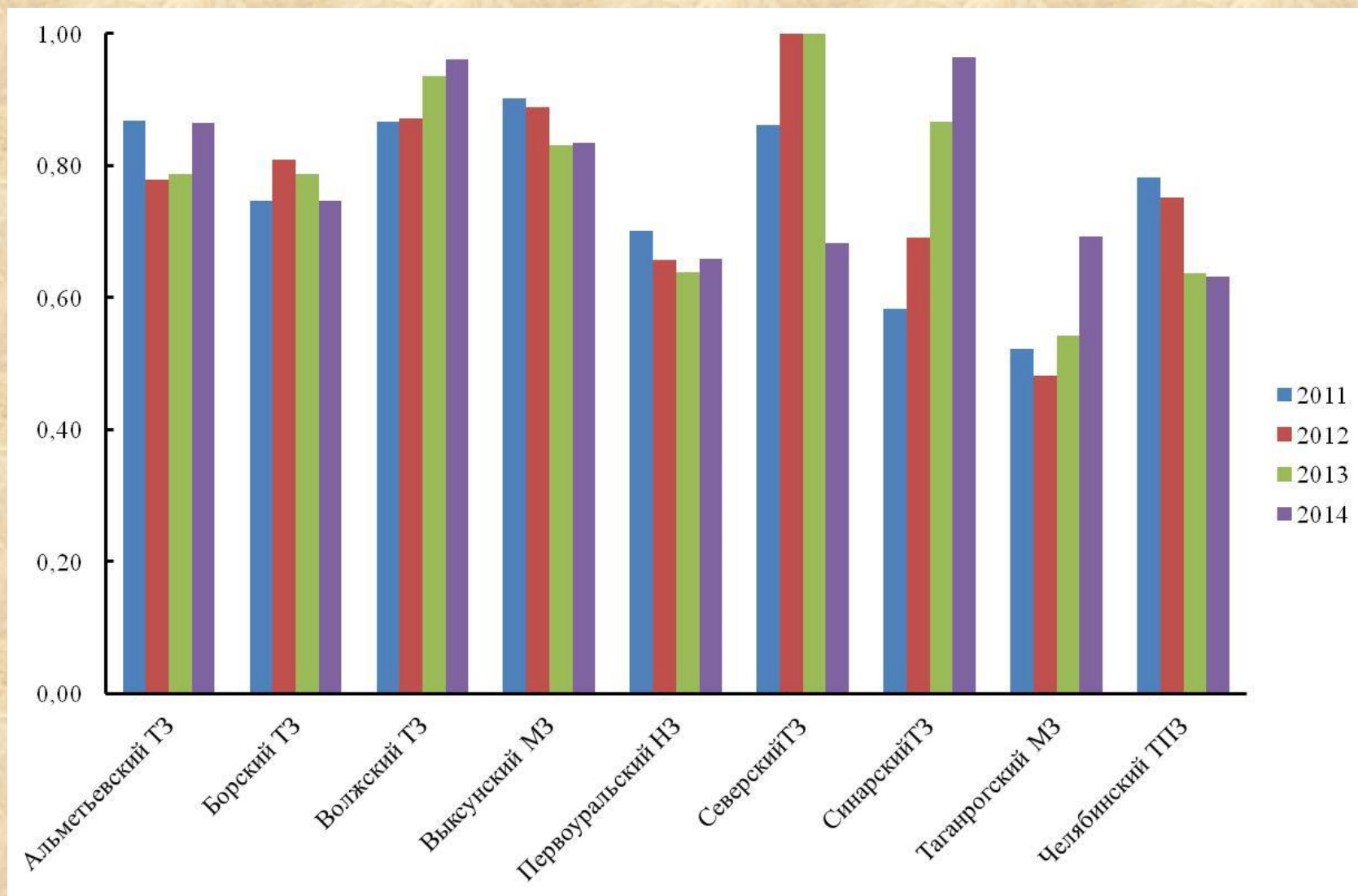
Рейтинговый показатель Кадыкова-Сайфулина фактической деятельности предприятий

Наименование	2011	2012	2013	2014
Альметьевский ТЗ	1,49	1,92	2,25	2,25
Борский ТЗ	1,49	1,49	1,49	1,76
Волжский ТЗ	2,02	2,02	2,14	2,22
Выксунский МЗ	2,24	2,23	2,23	2,27
Первоуральский НЗ	1,98	2,24	2,38	2,31
Северский ТЗ	1,63	1,64	1,92	1,93
Синарский ТЗ	1,63	1,64	1,66	1,79
Таганрогский МЗ	1,63	1,63	1,63	1,70
Челябинский ТЗ	2,24	2,54	2,56	2,66

Интегральный рейтинговый показатель (метод собственных состояний)

Наименование	2011	2012	2013	2014
Альметьевский ТЗ	0,87	0,78	0,79	0,87
Борский ТЗ	0,75	0,81	0,79	0,75
Волжский ТЗ	0,87	0,87	0,94	0,96
Выксунский МЗ	0,90	0,89	0,83	0,84
Первоуральский НЗ	0,70	0,66	0,64	0,66
Северский ТЗ	0,86	1,00	1,00	0,68
Синарский ТЗ	0,58	0,69	0,87	0,96
Таганрогский МЗ	0,52	0,48	0,54	0,69
Челябинский ТЗ	0,78	0,75	0,64	0,63

Интегральный рейтинговый показатель (метод собственных состояний)



Штрафные функции 2011

	Альме тьевск ий ТЗ	Борск ий ТЗ	Волжск ий ТЗ	Выксун ский МЗ	Первоу ральск ий НЗ	Северс кий ТЗ	Синарс кий ТЗ	Таганр огский МЗ	Челяби нский ТЗ
Производство ТП	0	0	0,12	0,222	0,686	0	0	0,710	0
Отгрузка ТП	0	0	0	0,265	0,68	0	0	0,705	0,005
Выбросы ЗВ	0	0	0	0	0	0,014	0	0	0,248
Средняя ЗП	0,053	0,01	0,42	0,070	0,05	0,107	0,026	0,082	0
Производ-ть труда	0,475	0	0	0	0,473	0	0	0	0,405
Запасы	0	0	0	0	0	0	0	0	0
Выручка	0	0,30	0,1	0	0	0,114	0,224	0,136	0,212
Прибыль от продаж	0	0,18	0,039	0	0	0,120	0,626	0,609	0
Себестоимость продаж	0,014	0	0	0	0	0	0	0	0
Валовая прибыль	0	0,46	0,16	0	0	0,224	0,437	0,328	0
Чистая прибыль	0	0,62	0	0	0	0,278	0,9	0,895	0

Анализ устойчивого развития регионов РФ в рамках концепции «Давление Состояние Реакция»

Объекты анализа представляют собой 45 регионов Российской Федерации. Для описания эколого-экономического развития регионов используются 13 показателей, которые в соответствии с концепцией ДСР объединены в три группы.

Показатели по всем группам собраны за 2011 год. Для того чтобы оценить эффективность реакции к показателям 2011 года добавляются показатели группы «Давления» и «Состояние» за 2012 год.

Анализ устойчивого развития регионов РФ в рамках концепции «Давление Состояние Реакция»

Группа	Показатели	Сокращение
Давление	Обрабатывающие производства (млн. руб.)	ОП
	Мощность электростанций (мегаватт)	ВРП
	Отправление грузов а/м транспортом, тонн	ООРТ
Состояние	Выбросы загрязняющих веществ (тыс. тонн.)	ВЗВ
	Сброс загрязненных сточных вод (млн. м)	СЗСВ
	Использование чистой воды (млн. м)	ИЧВ
	Болезни крови и сопутствующие заболевания (на млн. чел.)	БК
	Болезни эндокринной системы и сопутствующие заболевания (на 1 млн. чел.)	БЭС
	Врожденные аномалии (на млн. чел.)	ВА
	Младенческая смертность (умерших до 1 года на млн. родившихся)	МС
	Повторно использованная вода (м)	ПИВ
	Улавливание загрязняющих атмосферу веществ (тыс. тонн.)	УЗВ
Реакция	Затраты на охрану окружающей среды (млн. руб.)	ЗООС

Анализ устойчивого развития регионов РФ в рамках концепции «Давление Состояние Реакция»

В основе модели будет лежать следующее условие: затраты на охрану окружающей среды в 2011 году должны снижать темпы роста выбросов загрязняющих веществ и сброс загрязненных сточных вод в 2012 году.

$$K_{2012} = \frac{\left(\text{ВЗВ}_{2011} \frac{\text{ОП}_{2012}}{\text{ОП}_{2011}} - \text{ВЗВ}_{2012} \right)}{\text{ЗООС}_{2011}}$$

$$\tau_{2012} = \frac{\left(\text{СЗСВ}_{2011} \frac{\text{ОП}_{2012}}{\text{ОП}_{2011}} - \text{СЗСВ}_{2012} \right)}{\text{ЗООС}_{2011}}$$

Анализ устойчивого развития регионов РФ в рамках концепции «Давление Состояние Реакция»

Регионы Российской федерации делятся на три класса в пространстве трех показателей: выбросы загрязняющих веществ, сброс загрязненных сточных вод, использование чистой воды.

В первый класс (низкий уровень загрязнений) входят 26 регионов.

Во второй класс (средний уровень загрязнений) входят 13 регионов.

Третий класс (высокий уровень загрязнений) составляют 6 регионов.

Анализ устойчивого развития регионов РФ в рамках концепции «Давление Состояние Реакция»

Основным требованием экологически-устойчивого (эталонного) развития регионов является условие: *относительные коэффициенты удельных выбросов загрязняющих веществ и удельных сбросов загрязненных сточных вод должны лежать выше допустимых значений.*

Первый класс: $\kappa_{2012} > 15 \text{ кг/руб}$ $\tau_{2012} > 7 \text{ м}^3/\text{руб}$

Второй класс: $\kappa_{2012} > 5 \text{ кг/руб}$ $\tau_{2012} > 4 \text{ м}^3/\text{руб}$

Третий класс: $\kappa_{2012} > 2,5 \text{ кг/руб}$ $\tau_{2012} > 4 \text{ м}^3/\text{руб}$

Анализ устойчивого развития регионов РФ в рамках концепции «Давление Состояние Реакция»

Эталонная модель формируется из трех собственных состояний.

Первое собственное состояние принадлежит к группе постоянных собственных состояний, двенадцатое и тринадцатое собственные состояния — к группе собственных состояний с положительными главными компонентами.

Коэффициенты эталонного развития регионов 1-го класса

Регион	K_{2012}	τ_{2012}
Брянская область	27,0	15,5
Владимирская область	16,3	11,3
Курская область	17,6	9,5
Архангельская область	20,4	11,4
Мурманская область	30,7	16,7
Новгородская область	27,8	16,8
Псковская область	33,2	10,7
Астраханская область	38,1	11,0
Кабардино-Балкарская Республика	36,9	7,8
Карачаево-Черкесская Республика	59,9	13,2
Ставропольский край	74,2	33,7
Оренбургская область	16,3	11,3
Пензенская область	18,9	11,1
Саратовская область	16,3	11,3

Коэффициенты эталонного развития регионов

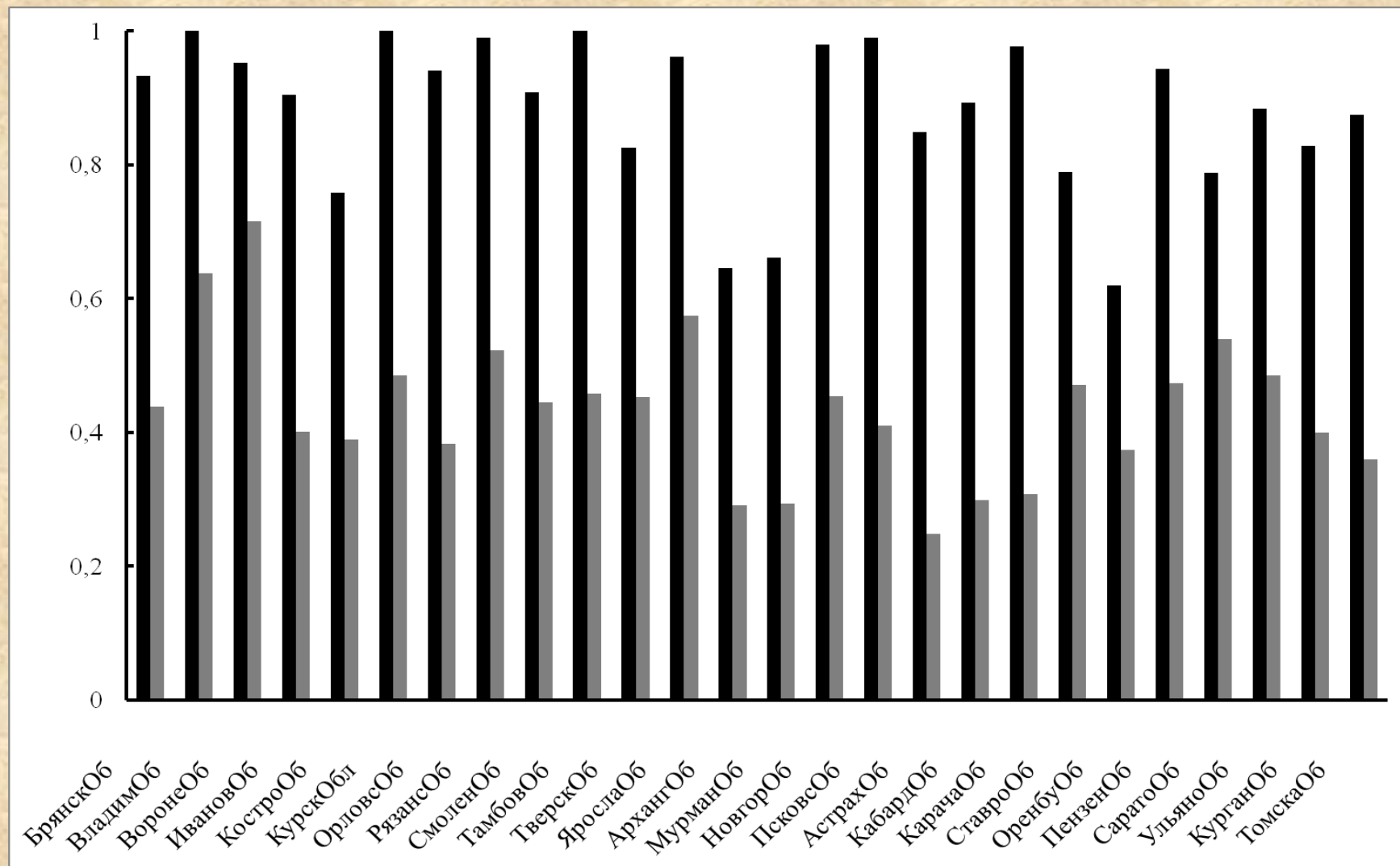
2-го класса

Регион	K_{2012}	τ_{2012}
Калужская область	9,5	5,5
Липецкая область	7,7	4,5
Тульская область	7,5	2,5
Вологодская область	6,8	4,9
Калининградская область	10,1	2,6
Ленинградская область	7,9	5,5
Краснодарский край	38,2	16,2
Волгоградская область	7,3	5,9
Ростовская область	6,6	5,3
Кемеровская область	31,6	12,8
Новосибирская область	10,7	2,6
Омская область	7,1	5,7

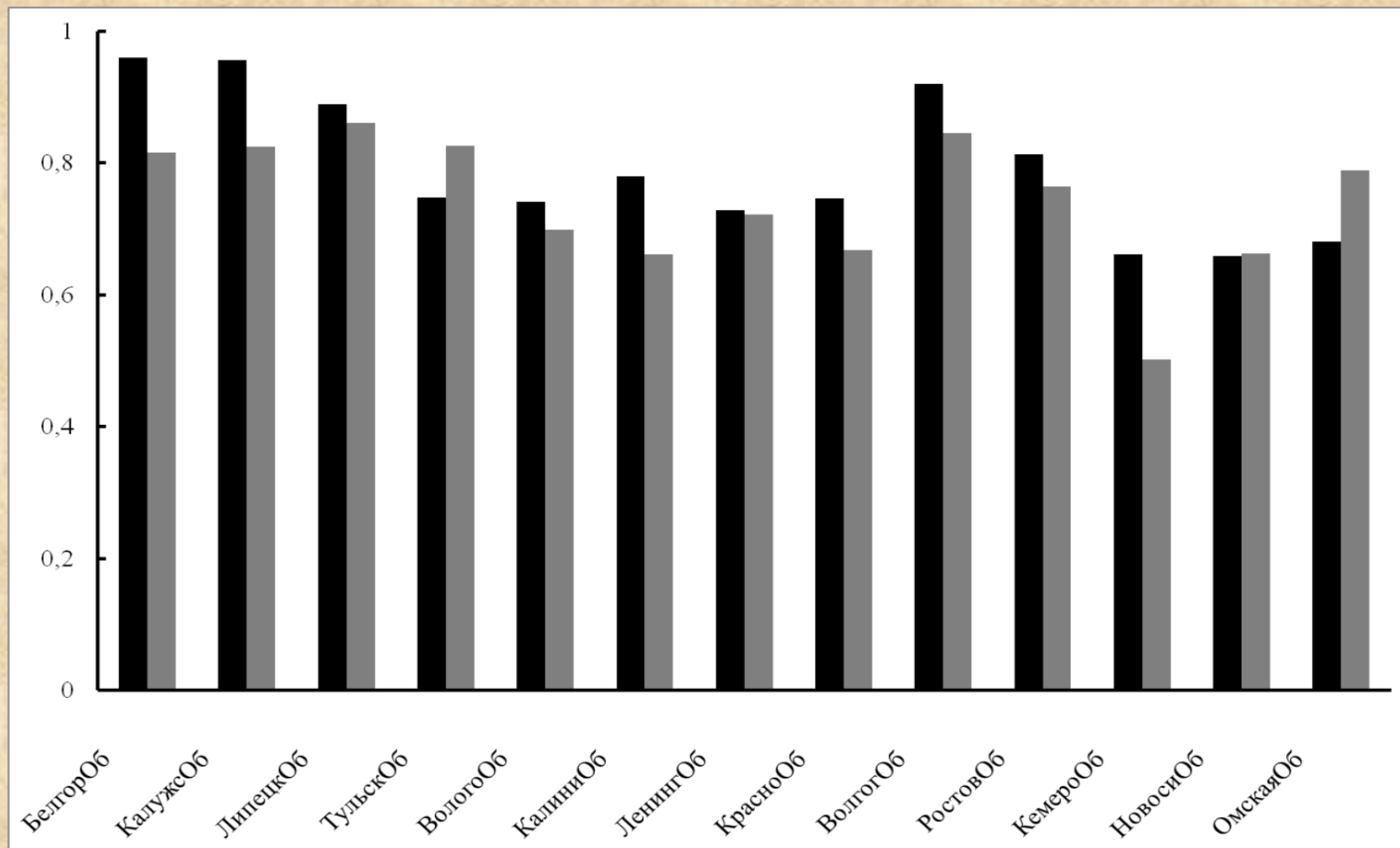
Коэффициенты эталонного развития регионов 3-го класса

Регион	κ_{2012}	τ_{2012}
Московская область	3,4	7,9
Нижегородская область	3,8	6,9
Самарская область	3,0	5,8
Свердловская область	4,1	8,2
Тюменская область	3,4	7,2
Челябинская область	3,5	7,1

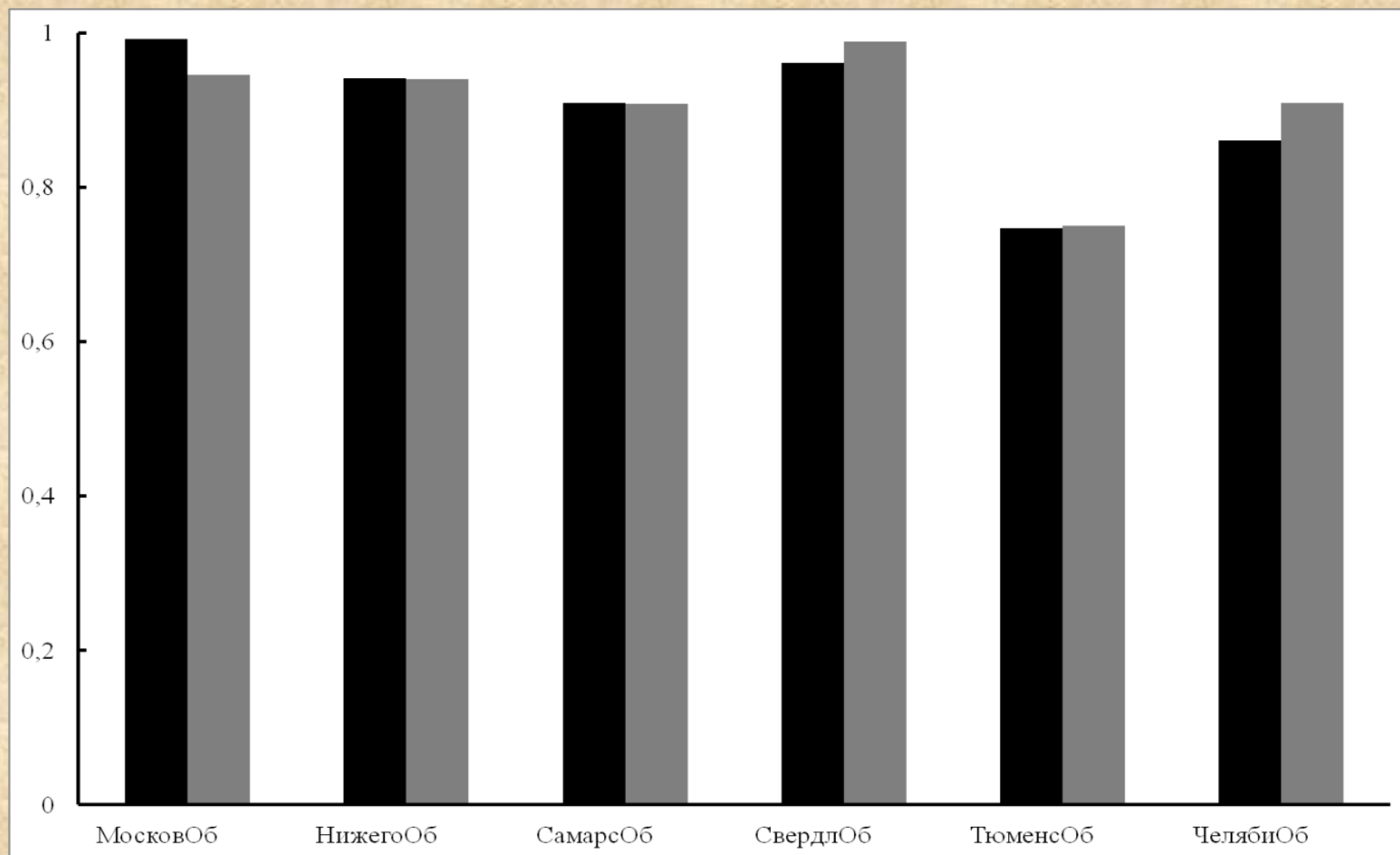
Комплексный индикатор эколого-устойчивого развития регионов 1-го класса



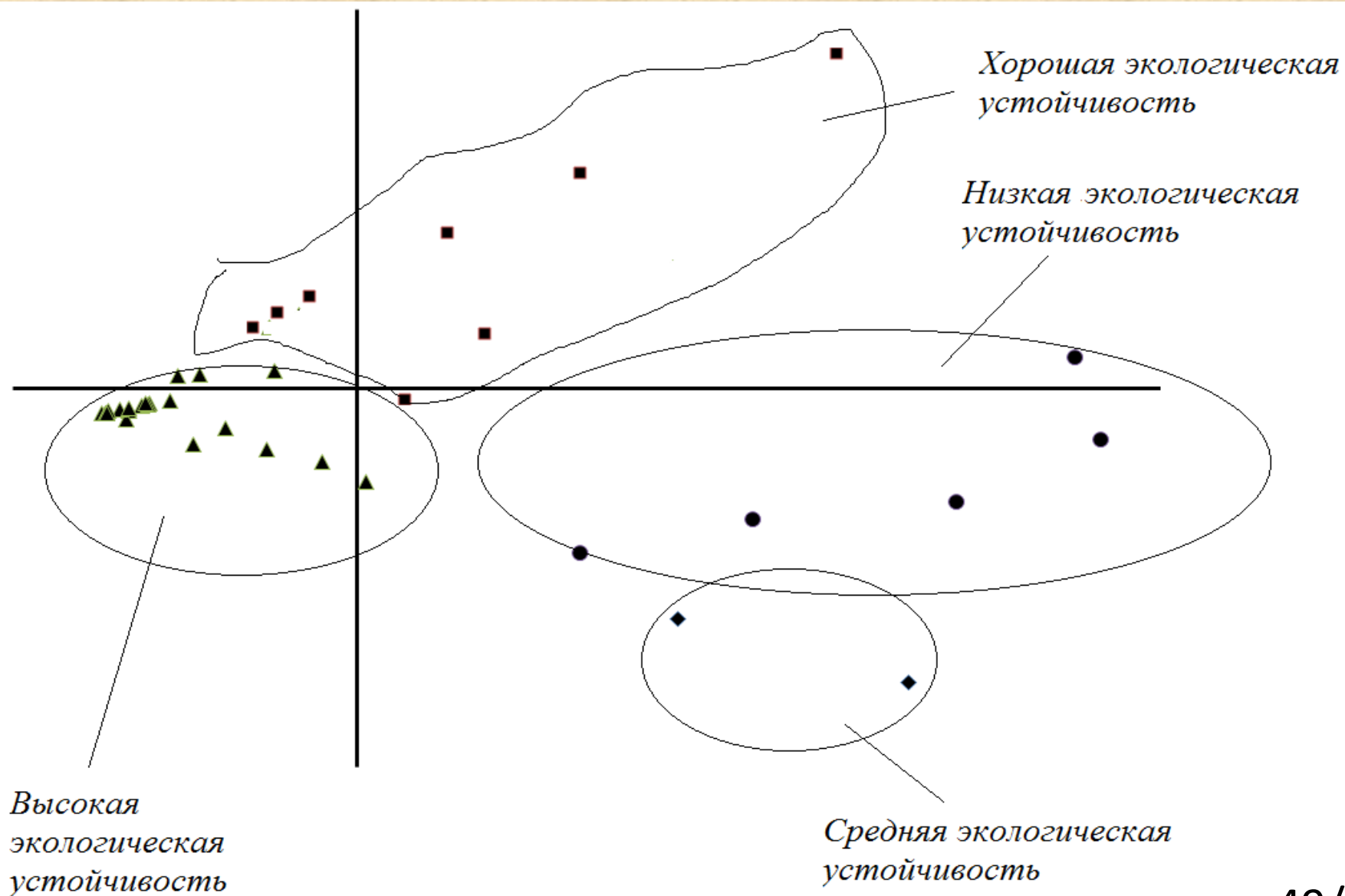
Комплексный индикатор эколого-устойчивого развития регионов 2-го класса



Комплексный индикатор эколого-устойчивого развития регионов 1-го класса



Кластеризация регионов



3

Рейтинг эколого-устойчивого развития регионов			
Регион	Класс	Индикатор	Рейтинг
Астраханская область	Хорошая экологическая устойчивость	0,85	B+
Тверская область		0,82	B+
Ростовская область		0,81	B+
Ставропольский край		0,79	B
Костромская область		0,76	B
Тюменская область		0,75	B
Ленинградская область		0,73	B
Оренбургская область		0,62	B-
Тульская область	Средняя экологическая устойчивость	0,75	C+
Новосибирская область		0,66	C
Калининградская область	Низкая экологическая устойчивость	0,78	D+
Краснодарский край		0,75	D+
Кемеровская область		0,66	D
Мурманская область		0,66	D
Архангельская область		0,65	D
50/50			