ОСНОВЫ ТЕОРИИ СИСТЕМ УПРАВЛЕНИЯ БАЗАМИ ДАННЫХ

(Приложения теории СУБД к задачам оценки земельных ресурсов)

Голозубов Олег Модестович,
Южный федеральный университет,
Кафедра почвоведения

Цель лекции:

определить задачи почвоведа в проектировании и управлении функционированием почвенно-географических БД.

Связь задач оценки земельных ресурсов и задач государственного мониторинга

Министерство сельского хозяйства Российской федерации
К О Н Ц Е П Ц И Я
развития государственного мониторинга
земель сельскохозяйственного назначения …
и формирования государственных информационных ресурсов …
Приказ от 4 мая 2010 г. №150
Об утверждении Порядка государственного учета показателей состояния плодородия земель сельскохозяйственного назначения

МИНИСТЕРСТВО ЭКОНОМИЧЕСКОГО РАЗВИТИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ПРИКАЗ от 20 сентября 2010 г. N 445
ОБ УТВЕРЖДЕНИИ МЕТОДИЧЕСКИХ УКАЗАНИЙ ПО ГОСУДАРСТВЕННОЙ КАДАСТРОВОЙ ОЦЕНКЕ ЗЕМЕЛЬ СЕЛЬСКОХОЗЯЙСТВЕННОГО НАЗНАЧЕНИЯ

Постановление Правительства РФ от 9 августа 2013 г. N 681
"О государственном экологическом мониторинге (государственном мониторинге окружающей среды) и государственном фонде данных государственного экологического мониторинга (государственного мониторинга окружающей среды)"

Участники мониторинга:

1. Министерство природных ресурсов и экологии
2. Министерство сельского хозяйства
3. Федеральная служба по гидрометеорологии и мониторингу окружающей среды
4. Федеральная служба государственной регистрации, кадастра и картографии
5. Федеральное агентство лесного хозяйства
6. Федеральное агентство по недропользованию
7. Федеральное агентство водных ресурсов
8. Федеральное агентство по рыболовству
9. Органы исполнительной власти субъектов Российской Федерации

Задачи мониторинга:

- Учреждения осуществляют формирование и ведение первичной базы данных;
- Обеспечивают возможность сохранения первичного состояния исходных данных, а также внесение изменений и дополнений.

Основные виды информации, формируемой на основе государственных информационных ресурсов о с/х землях с использованием современных информационных технологий

- **информация о границах сельскохозяйственных земель** (участков, сельскохозяйственных полигонов, контуров), их площади, состоянии, виде разрешенного и хозяйственного использования, потенциальной продуктивности;
- **показатели для земельного участка определяются как средневзвешенные по площади почвенных разновидностей в составе земельного участка;**
- **информация о состоянии плодородия почв, включая показатели, характеризующие морфогенетические свойства почв, их гранулометрический состав, кислотность, содержание гумуса, макро- и микроэлементов, тяжелых металлов и радионуклидов, степени эродированности, переувлажнения, заболачивания, засоления, опустынивания, каменистости**, а также характеристики произрастающей на них растительности по геоботаническому составу, урожайности сельскохозяйственных культур, установленной при проведении наземных обследований;
- другая информация с различными степенями агрегации (Российская Федерация, субъект Российской Федерации, муниципальный район/городской округ, сельское/городское поселение), подготовленная в соответствии с потребностями пользователей, а также программные продукты.

Данные и метаданные: номинальные, количественные, порядковые…
Значения: текст, целые, вещественные, логические,…

Образец почвы
Морфометрия
Химические свойства
Физические свойства
Водно-физические свойства
Показатели плодородия
Показатели засоления
Микроагрегатный состав
Экологические показатели

Земельный участок
Кадастровые показатели:
(ОКАТО,ИНН,координаты,шифры)
Природные показатели:
(уклон, тип угодья, намытость,каменистость …)

Разрез

Координаты (долгота, широта, высота над у.м.)
Мощность пах. горизонта

Почвенный контур
- Координаты границ
- Природная зона
- Почвенная провинция
- Вид
- Разновидность

Данные мониторинга:

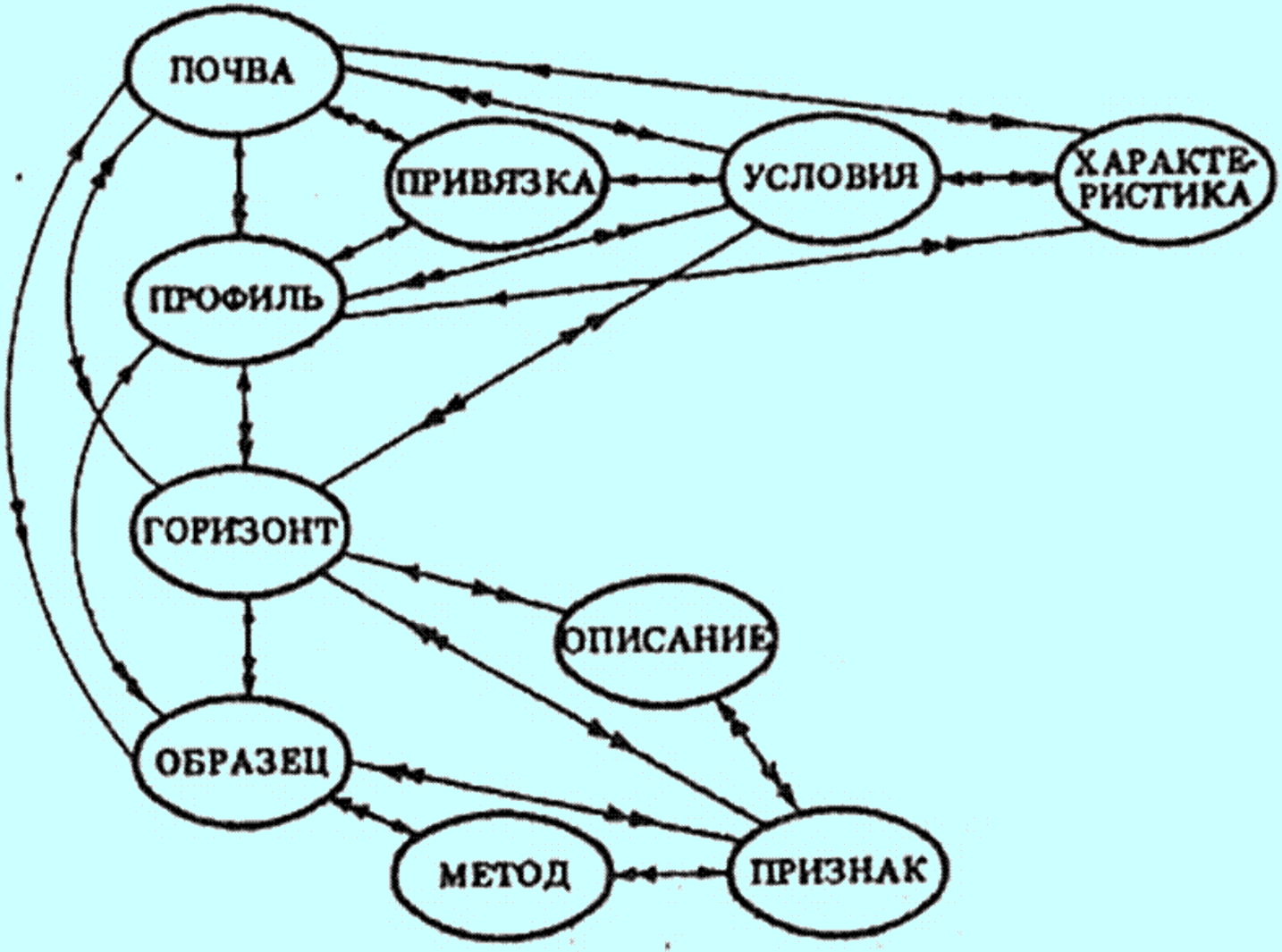
Культура произрастания

Параметры загрязнения

Агрохимические мероприятия

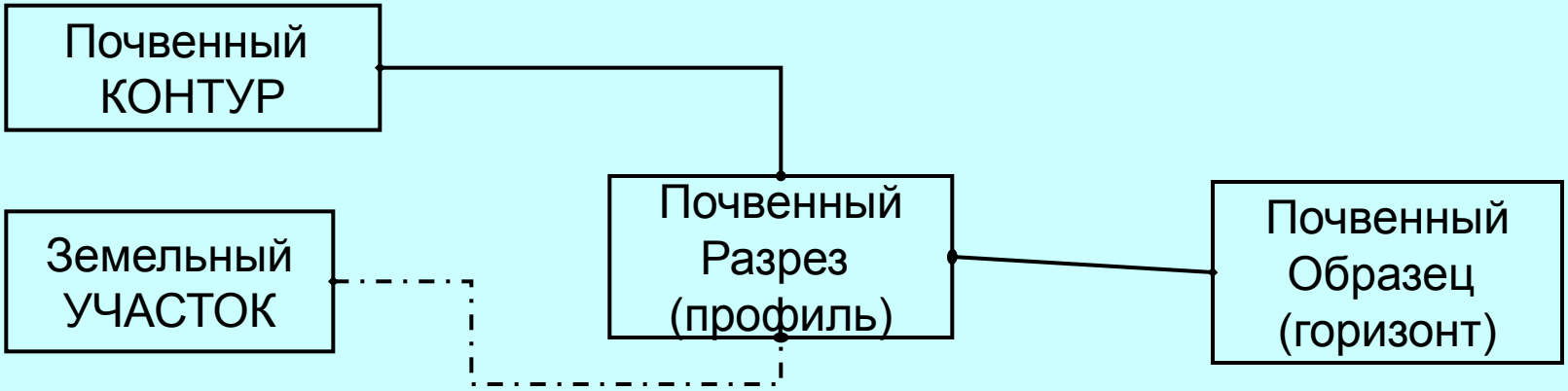
Иерархическая модель

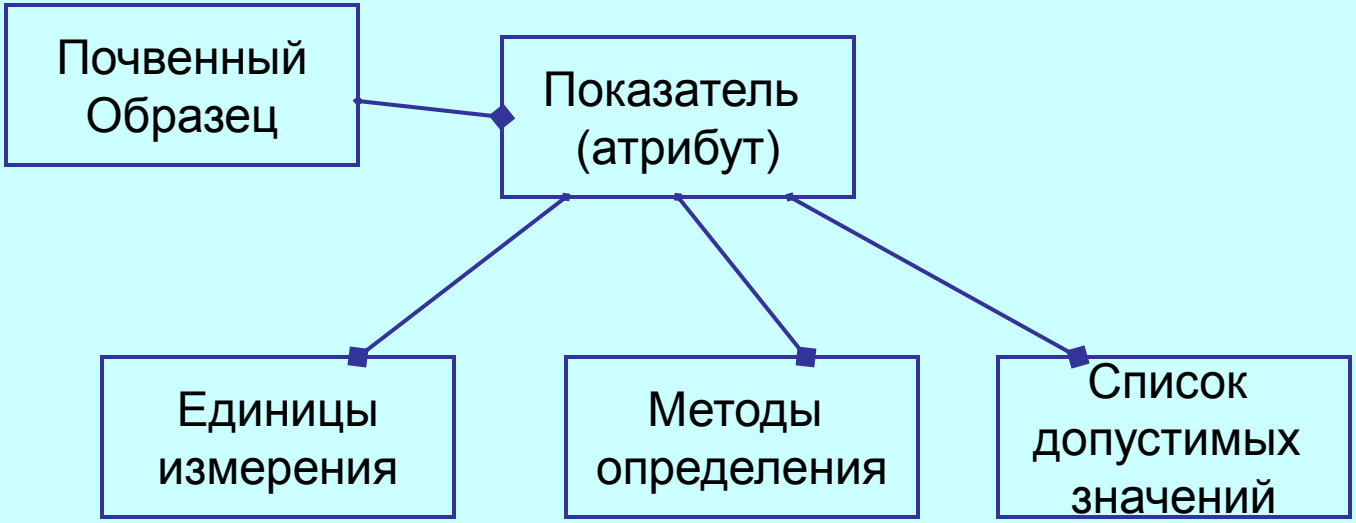
Сетевая модель



Реляционная структура данных

Реляционная структура данных





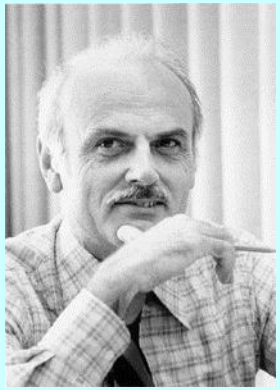
Реляционная СУБД — *СУБД*, управляющая реляционными базами данных.

Эти модели характеризуются простотой структуры данных, удобным для пользователя табличным представлением и возможностью использования формального аппарата алгебры отношений и реляционного исчисления для обработки данных.

Реляционная модель ориентирована на организацию данных в виде двумерных таблиц. Каждая реляционная таблица представляет собой двумерный массив и обладает следующими свойствами:

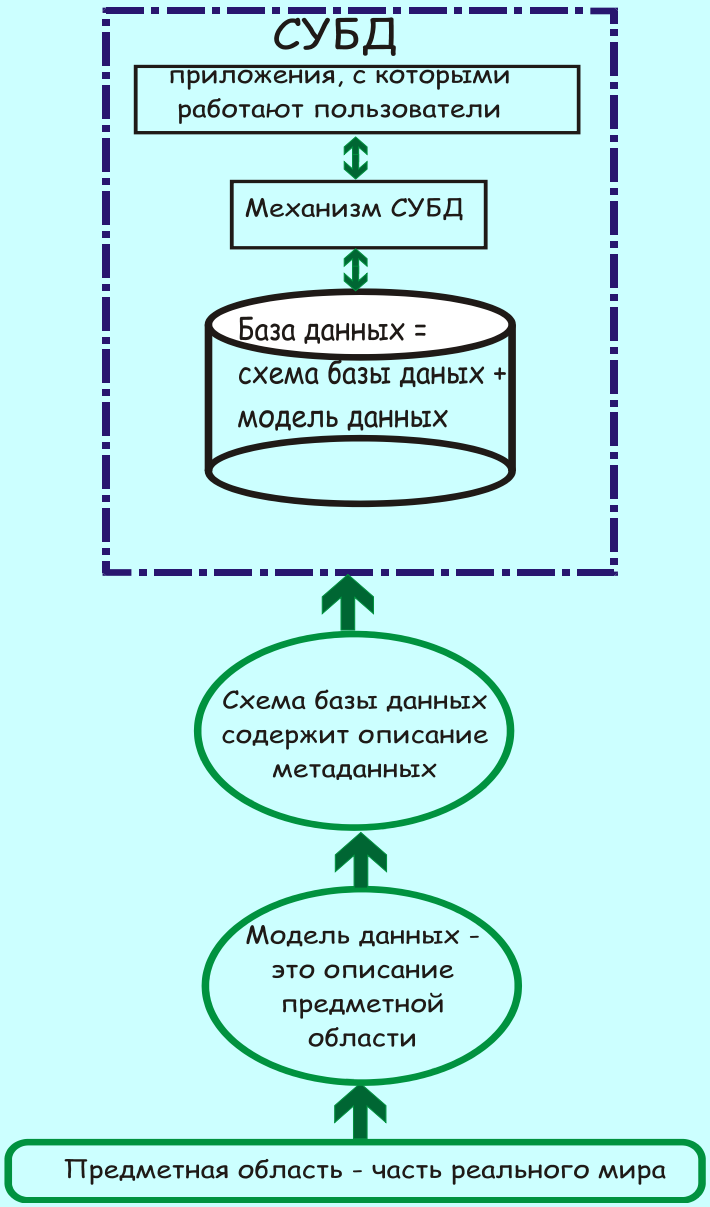
- каждый элемент таблицы — один элемент данных
- все ячейки в столбце таблицы однородные, то есть все элементы в столбце имеют одинаковый тип (числовой, символьный и т. д.)
- каждый столбец имеет уникальное имя
- одинаковые строки в таблице отсутствуют
- порядок следования строк и столбцов может быть произвольным

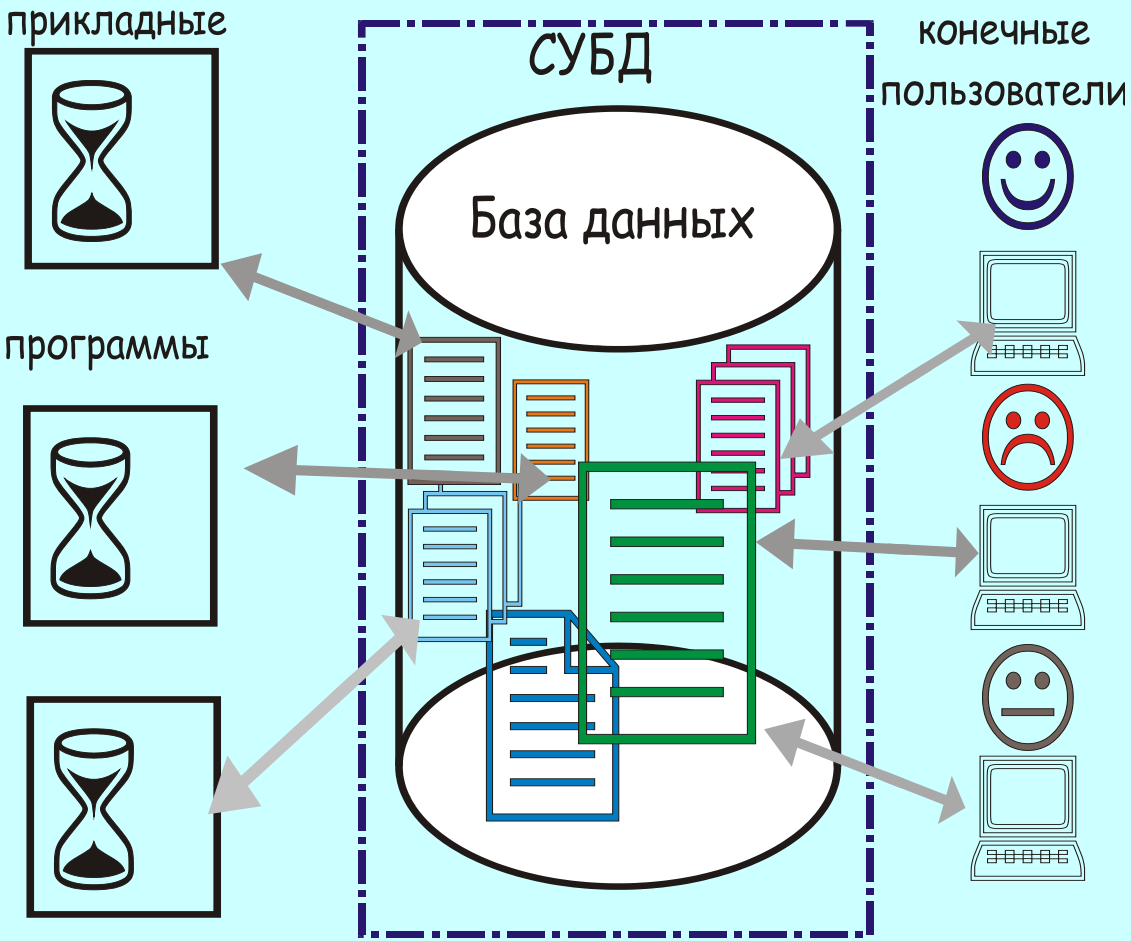
1985 г. : Сформулированы 12 правил Кодда



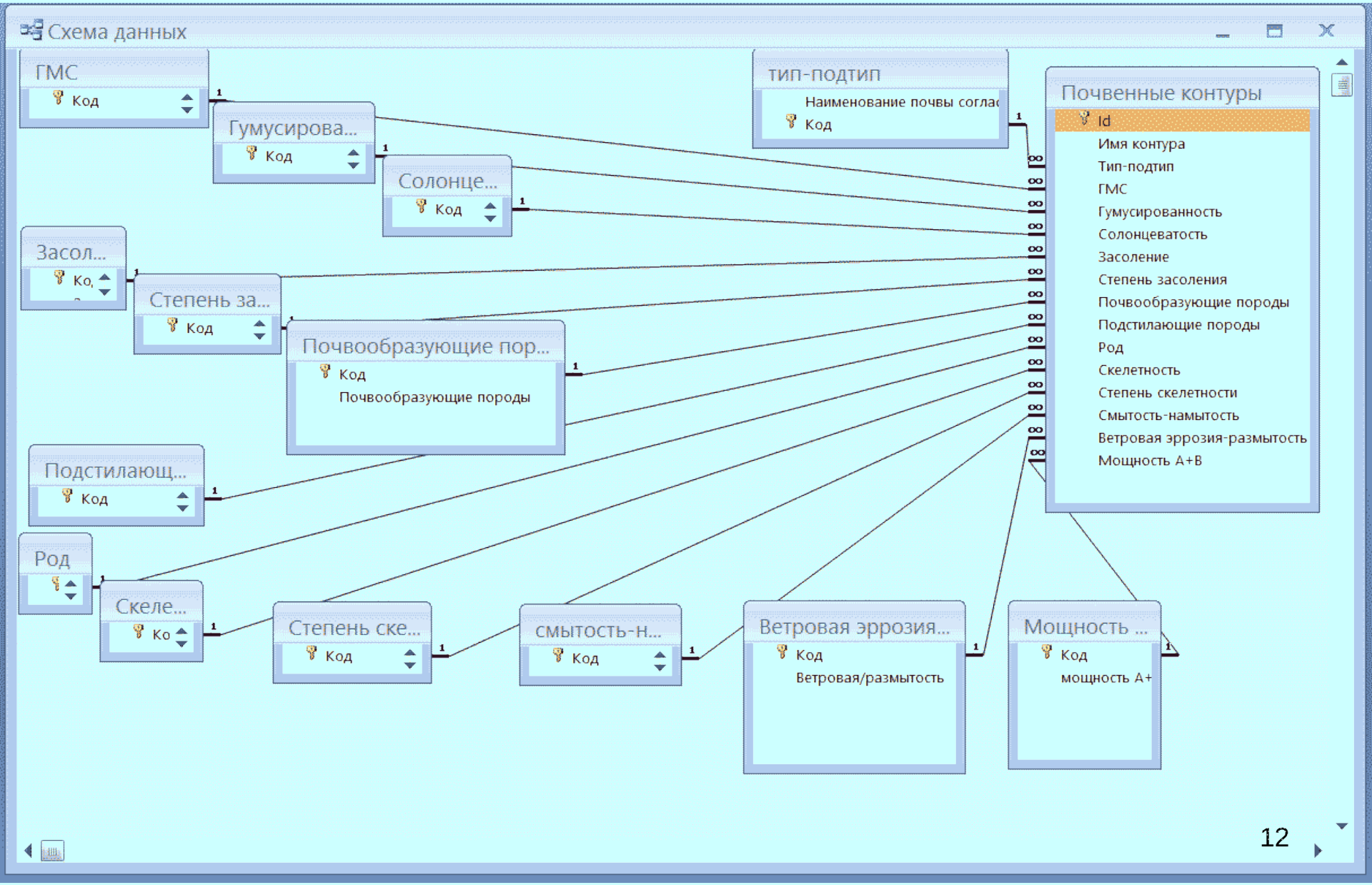
Понятие реляционный (англ. *relation* — отношение) связано с разработками известного английского специалиста в области систем баз данных **Эдгара Франка Кодда (**Edgar Frank Codd)

Правило 0: Раздельное существование данных и программ, множественность программ, независимость данных от программ.





Правило 5: Консервативные (справочники) и оперативные данные, понятие связей «один-многие…



Id	Название таблицы	Название поля
1	тип	Type
2	ГМС	Granularity
3	подтип	SubType
4	Подстилающие_породы	Background_Material
5	Почвообразующие_породы	Parent_Material
6	содержание гумуса	Humus_Content
7	смытость_намытость	Washed_off
8	род	Kind
9	засоление	Salinity
10	степень засоления	Salinity_degree
11	каменистость	Stoniness
12	ветровая_размытость	Wind_Errosion
13	мощность A+B	AB_Horizon_Power
14	солонцеватость	Saltness

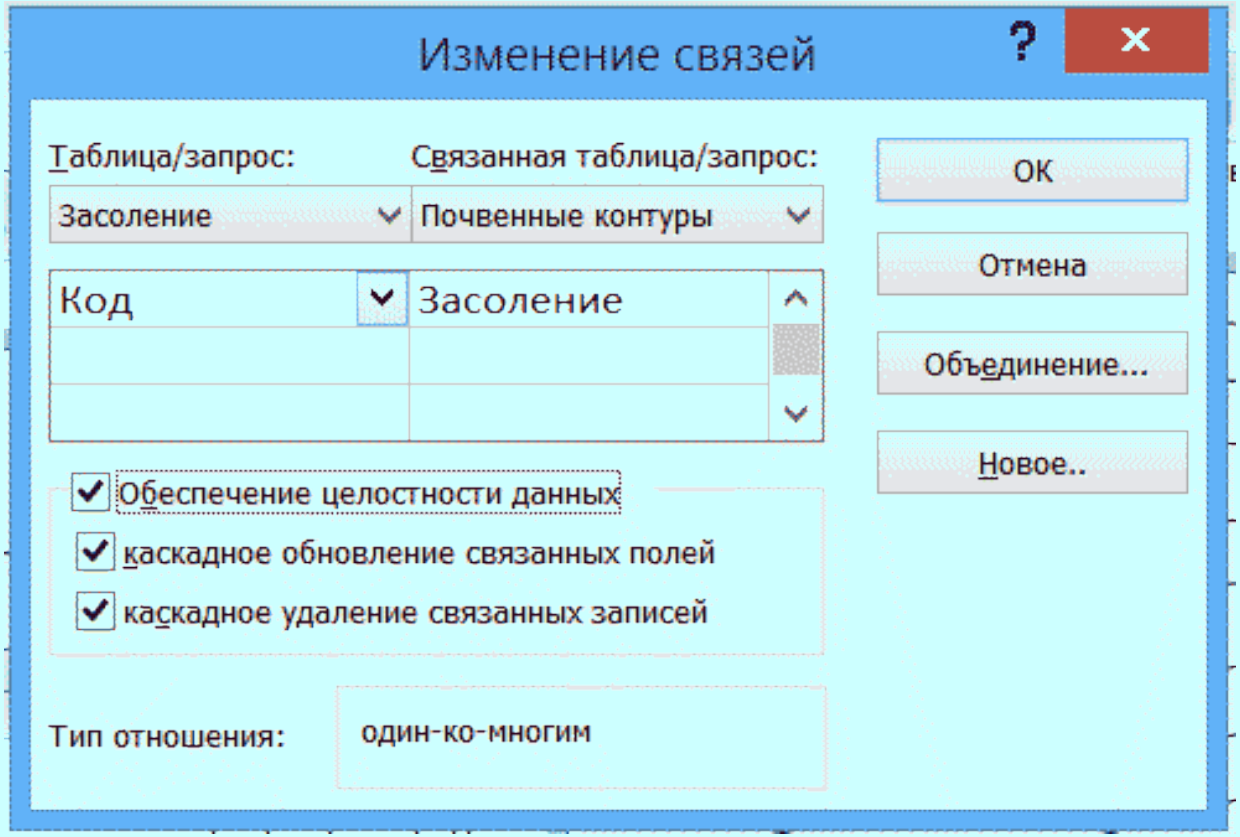
Код	Type
1	не определено
2	Аллювиально-делювиальные
3	Аллювиально-луговые карбонатные
4	Аллювиальные болотные иловато-перегнойно-глеевые
5	
6	
7	
8	
9	

Granularity	Код
не определено	1
глинистый	2
тяжелоглинистый	3
среднеглинистый	4
легкоглинистый	5
	6
	7
	8
	9
	10
	11
	12
	13
	14
несуглинистый	15
ый	16
	17
	18
, среднесуглинистый	19
	20
тяжелосуглинистый, среднесуглинистый	21
тяжелосуглинистый, легкосуглинистый	22
среднесуглинистый, легкосуглинистый	23
разный	24

Код	Background_Material
1	не определено
2	аргиллиты
3	бучакские пески
4	древнеаллювиальные отложения
5	желто-бурые глины
6	мергель
7	пески
8	пески, супеси
9	третичные пески
10	элювий известняков
11	элювий песчаника
12	элювий песчаника, мела
13	элювий плотных пород

Правило 10,12: Понятие целостности БД.

- Ссылочная целостность,
- каскадное обновление и удаление в связанных таблицах.
- Диаграмма связей.
- Понятие транзакции как цепочки операций над БД.



Правило 6: Обновление **представлений** (**View**).

Фрагмент отчета 5-ДДЗ: Агрегирование данных в «Куб»

Параметры мониторинга почвенного плодородия по видам угодий

	Параметры	№ строки	Площадь почв				
			Пашни, га	Пастбища, га	Сенокосы, га	Многолетние насаждения, га	Залежь, га
	1	2	3	4	5	6	7
	Содержание гумуса (Апах, А1) %	01					
1	<1	02	**0**	**0**	**0**	**0**	**0**
2	1,01 – 1,5	03	**500**	**1500**	**500**	**200**	**0**
3	1,51 – 2,0	04	**900**	**200**	**300**	**0**	**0**
4	2,01 – 2,5	05	**1600**	**2000**	**1000**	**200**	**0**
5	2,51 – 3,0	06	**2000**	**1300**	**700**	**0**	**0**
6	3,01 – 4,0	07	**800**	**1800**	**1400**	**300**	**0**
7	4,01 – 6,0	08	**800**	**1700**	**1100**	**0**	**0**
8	6,01 – 8,0	09	**0**	**0**	**0**	**0**	**0**
9	8,01 – 10,0	10	**0**	**0**	**0**	**0**	**0**
10	> 10,0	11	**0**	**0**	**0**	**0**	**0**

- Вывод об однократном (уникальном) хранении введенных значений,
- Вывод о хранении только исходных (первичных) данных.
- Вывод о оперативном (on-line) перевычислении расчетных показателей
- Вывод об отличии отчёта (Report) от представления.

Правило 3: Разреженные и неполные данные, понятие Null как неопределенного значения. Пример Soil_DB.

Полная обработка неизвестных значений (Systematic Treatment of Null Values):

Неизвестные значения NULL, отличные от любого известного значения, должны поддерживаться для всех типов данных при выполнении любых операций. Например, для числовых данных неизвестные значения не должны рассматриваться как нули, а для символьных данных — как пустые строки.

№ контура	Наименование почвы согласно легенде	Род	солонцеватость	Засоление	степень	Ветровая / размытость	смытость /намытость	Скелетность	степень	мощность A+B	Гумусированность	Почвообразующие породы	Подстилающие породы

Результат в БД:

8	33	5	null	null	null	null	1	3	1	3	2	13	8

Правило 1-2: независимость и уникальность строк в таблице

правило 1: **Явное представление данных** (The Information Rule):

Информация должна быть представлена в виде данных, хранящихся в ячейках. Данные, хранящиеся в ячейках, должны быть атомарны. Порядок строк в реляционной таблице не должен влиять на смысл данных.

правило 2: **Гарантированный доступ к данным** (Guaranteed Access Rule):

Доступ к данным должен быть свободен от двусмысленности. К каждому элементу данных должен быть гарантирован доступ с помощью комбинации имени таблицы, первичного ключа строки и имени столбца.

Особенности технологий СУБД при работе с почвенными данными:

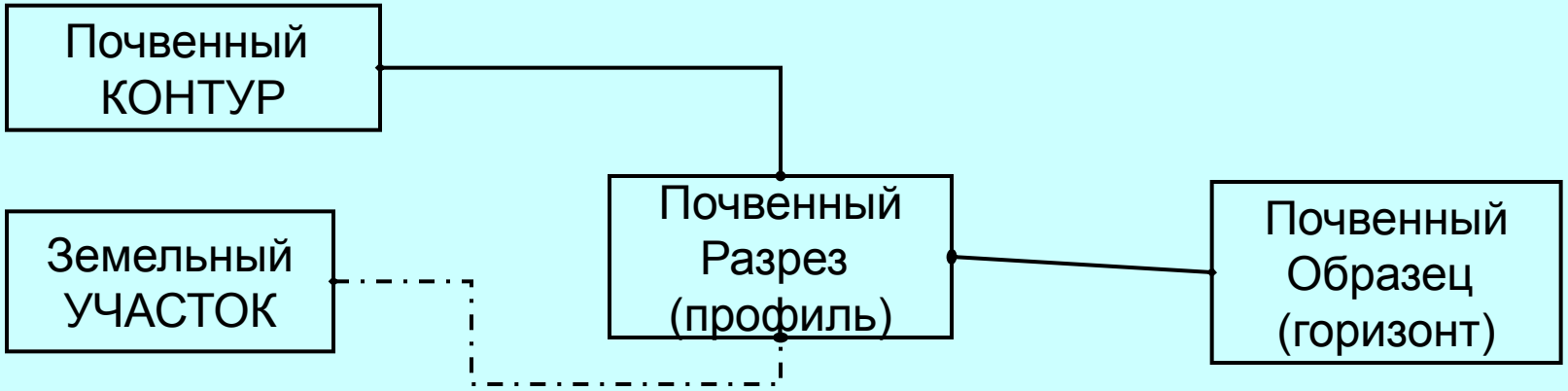
• хранение «сырых» данных (вечное);
• сложные процедуры обработки «сырых» данных;
• отслеживание изменяемости исходных данных;
• модель описания данных – многомерный вложенный массив с неровными краями и разреженностью;
• поддержка версионности и истории изменения данных;
• поддержка работы с данными, имеющими погрешность измерений;
• …

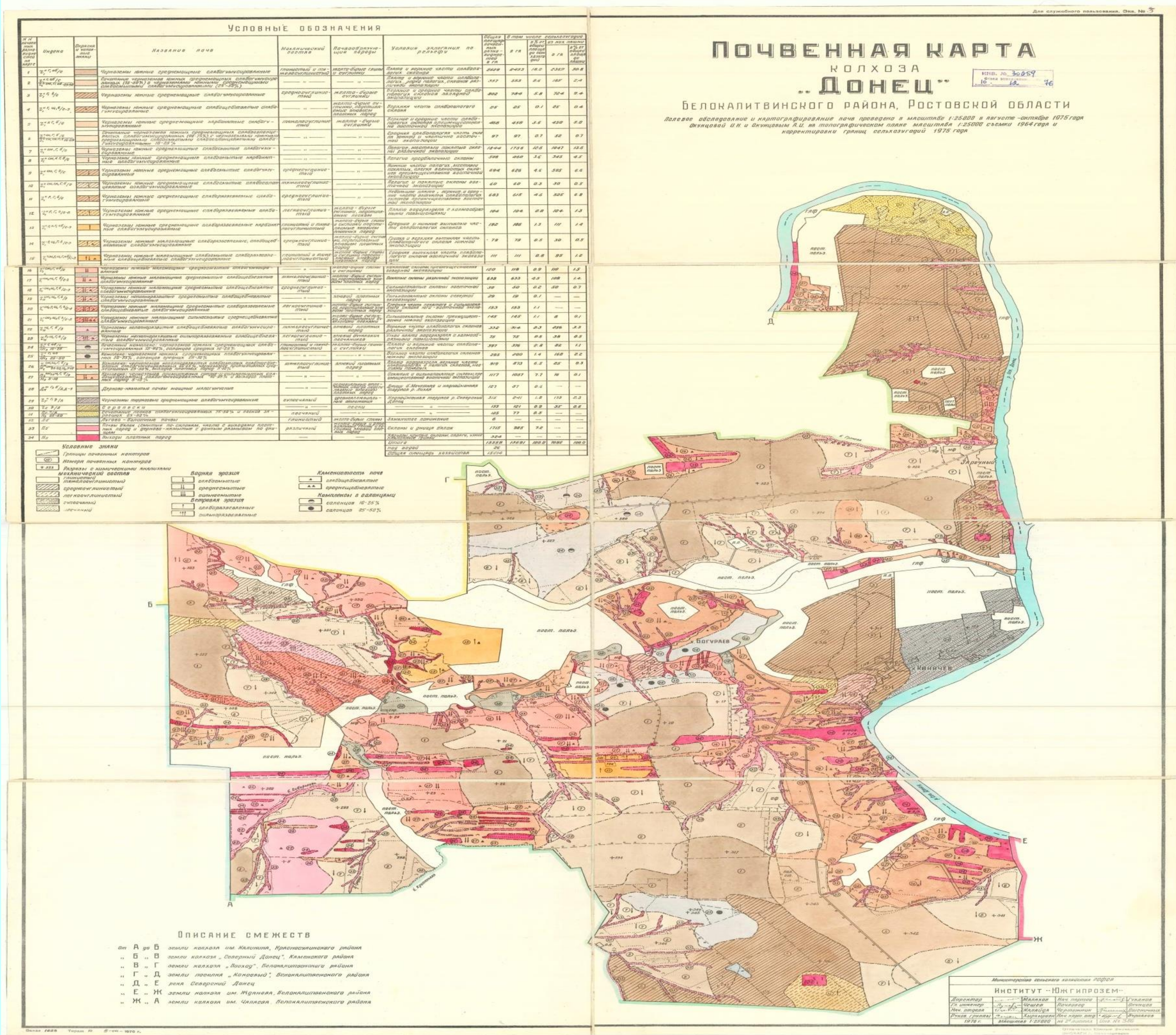
Расширение правил Кодда

для задач почвоведения в современных условиях

•Вывод о многообразии, приблизительности, пространственной и временной зависимости почвенных данных.
•Определение объекта оценки, необходимость его однозначной идентификации (на момент актуализации).
•Принцип идентификации по пространственно-временным координатам.

показатели для земельного участка определяются как средневзвешенные по площади почвенных разновидностей в составе земельного участка;




Для служебного пользования. Экз. № 3
Почвенная карта
колхоза
„Донец"
Белокалитвинского района, Ростовской области
Полевое обследование и картографирование почв проведено в масштабе 1:25000 в августе–октябре 1975 года
Окунцовой И.Н. и Окунцовым А.И. на топографическом плане масштаба 1:25000 съемки 1964 года и
корректировки границ сельхозугодий 1975 года
Условные обозначения
Условные знаки
Описание смежеств
от А до Б земли колхоза им. Калинина, Красносулинского района
„ Б „ В земли колхоза „Северный Донец", Каменского района
„ В „ Г земли колхоза „Восход", Белокалитвенского района
„ Г „ Д земли поселка „Коксовый", Белокалитвенского района
„ Д „ Е река Северский Донец
„ Е „ Ж земли колхоза им. Жданова, Белокалитвенского района
„ Ж „ А земли колхоза им. Чапаева, Белокалитвенского района
Министерство сельского хозяйства РСФСР
Институт „Южгипрозем"

УСЛОВНЫЕ ОБОЗНАЧЕНИЯ

N N почвенных разновидностей на карте	Индекс	Окраска и условные знаки	Название почв	Механический состав	Почвообразующие породы
1	$\tau_1''^{\,г,\,аб}/с$		Черноземы южные среднемощные слабогумусированные	глинистый и тяжелосуглинистый	желто-бурые глины и суглинки
2	$\tau_1''^{\,г,\,аб}/с$ / $\tau_1''^{\,см,\,г,\,аб}$ -25-50		Сочетание черноземов южных среднемощных слабогумусированных (75-50%) с черноземами южными среднемощными слабосмытыми слабогумусированными (25-50%)	—— " ——	—— " ——
3	$\tau_1''^{\,г,\,в}/с$		Черноземы южные среднемощные слабогумусированные	среднесуглинистый	желто-бурые суглинки
4	$\tau_1''^{\,г,\,щ,\,в}/с$-э	▲	Черноземы южные среднемощные слабощебневатые слабогумусированные	—— " ——	желто-бурые суглинки, подстилаемые элювием плотных пород
5	$\tau_1''^{\,к\,г,\,б}/с$		Черноземы южные среднемощные карбонатные слабогумусированные	тяжелосуглинистый	желто-бурые суглинки

Условные знаки

- Границы почвенных контуров
- (10) Номера почвенных контуров
- + 333 Разрезы с химическими анализами

Механический состав

- глинистый тяжелосуглинистый
- среднесуглинистый
- легкосуглинистый
- супесчаный
- песчаный

Водная эрозия

- ↓ слабосмытые
- ↓↓ среднесмытые
- ↓↓↓ сильносмытые

Ветровая эрозия

- ↑ слаборазвеваемые
- ↑↑↑ сильноразвеваемые

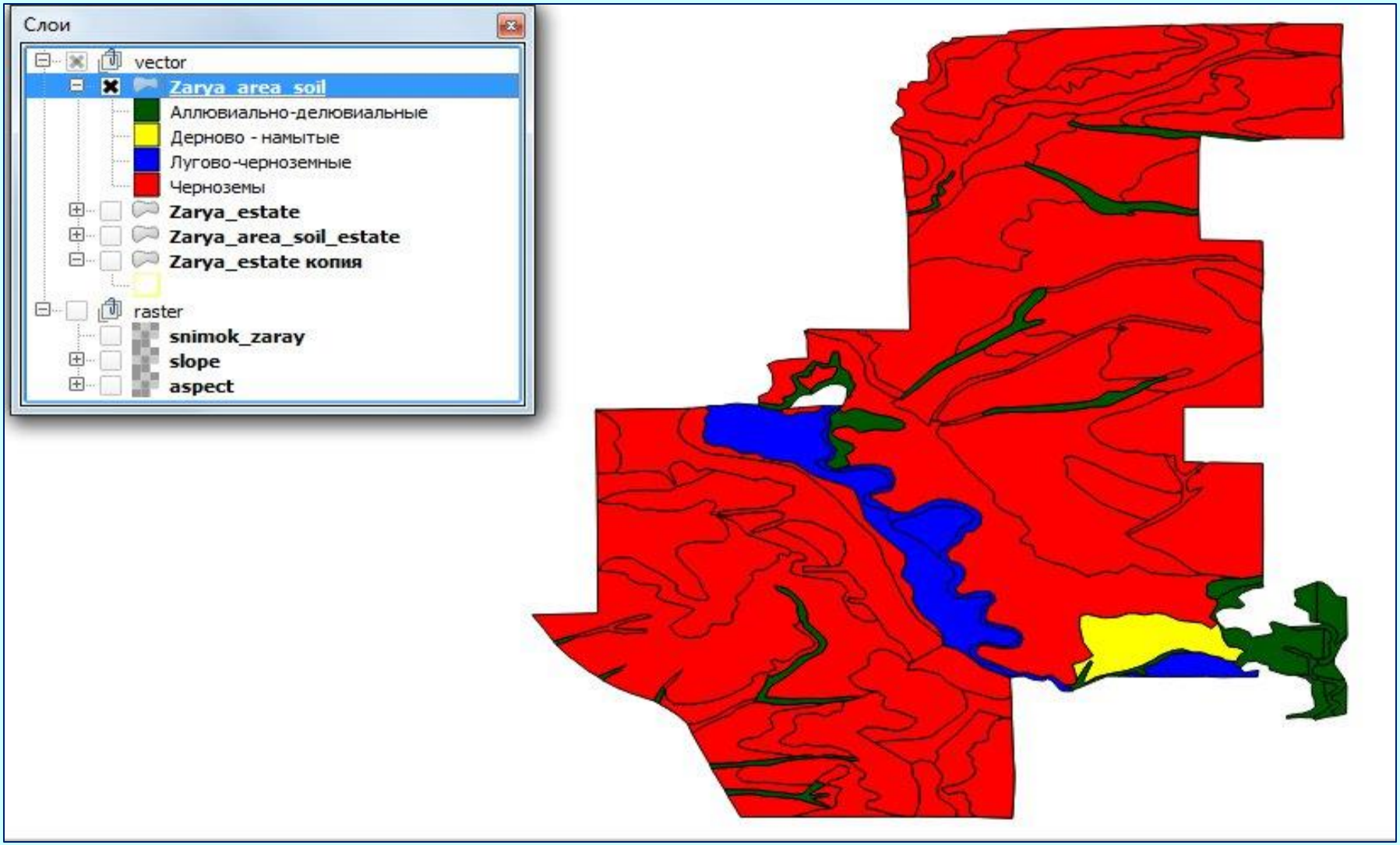
Каменистость почв

- ▲ слабощебневатые
- ▲▲ среднещебневатые

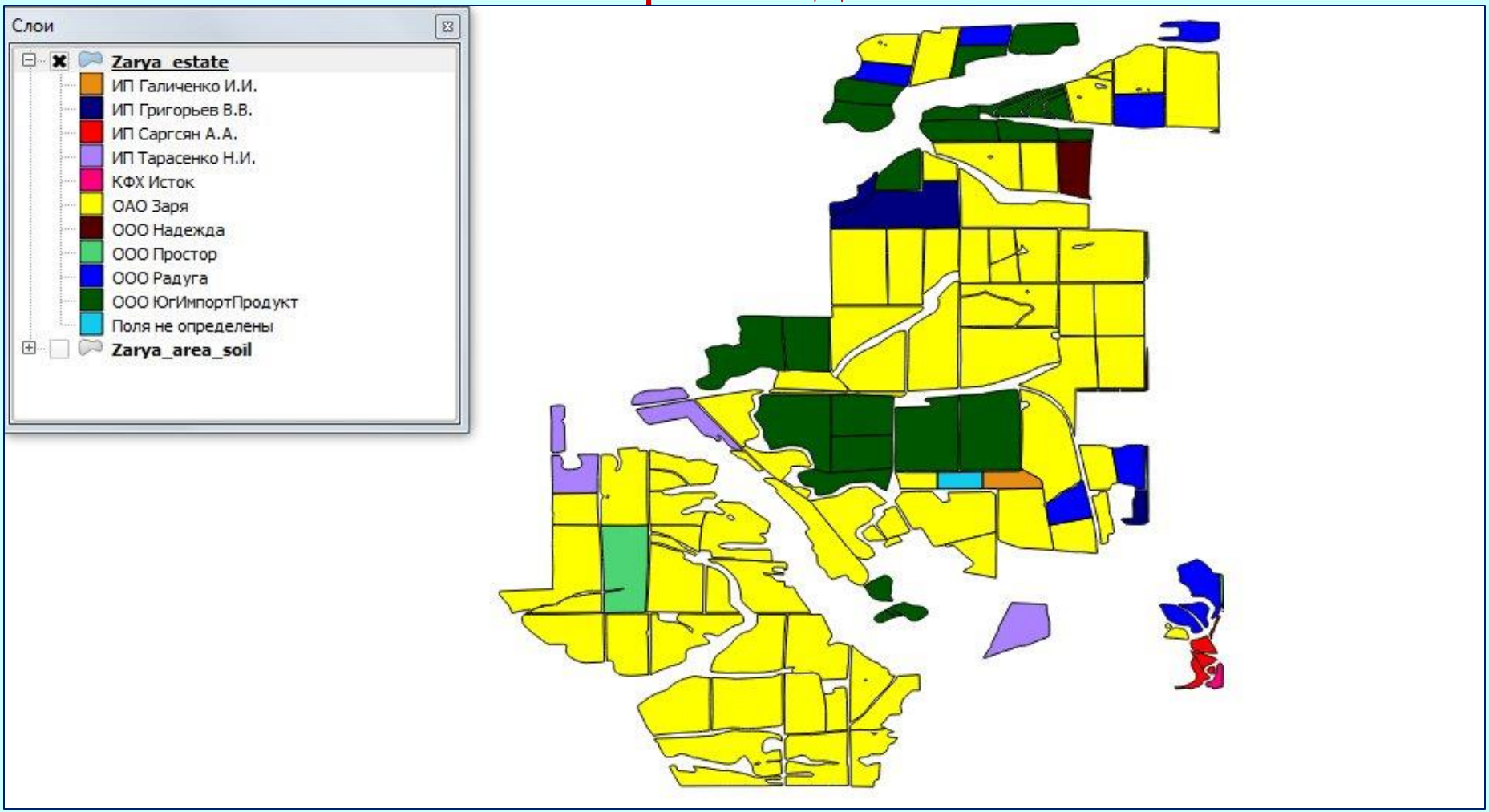
Комплексы с солонцами

- солонцов 10-25%
- солонцов 25-50%

Карта типов почв хозяйства Заря



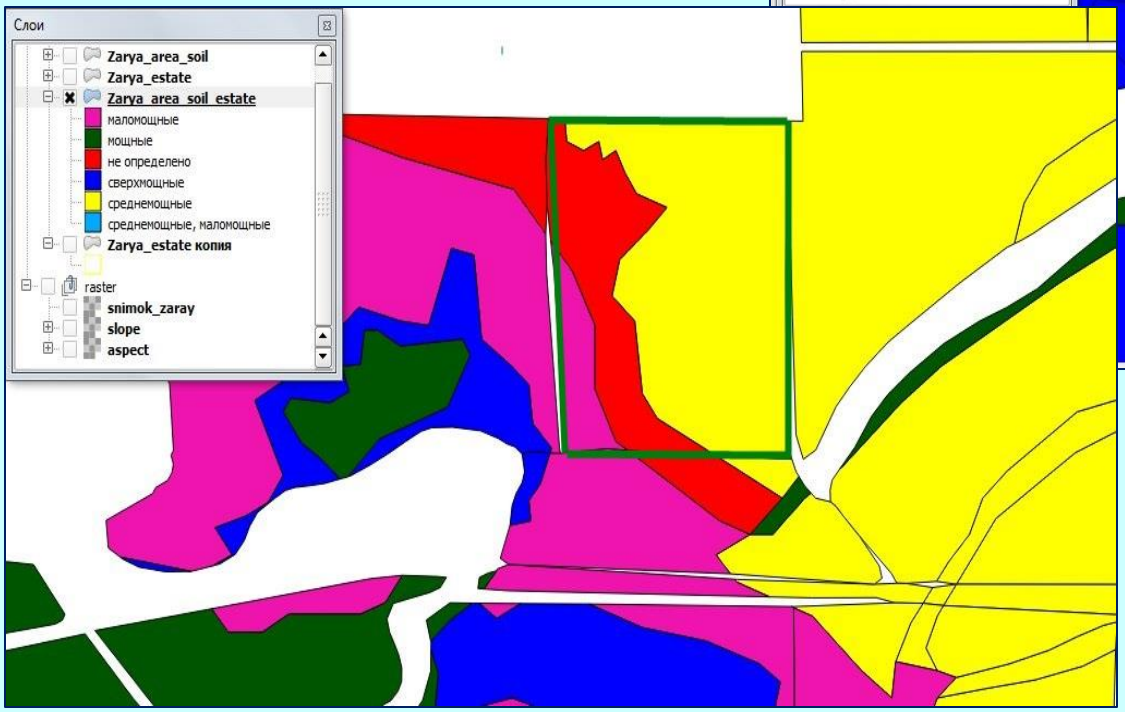
Распределение территории хозяйства по современным сельскохозяйственным производителям



Анализ пестроты почвенного покрова на территории выбранного поля

Пестрота поля по смытости

Пестрота поля по мощности горизонта А+В



Новый тип данных – пространственные данные.

ГИС

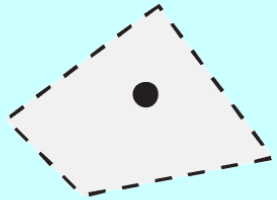
• разделение на атрибутивные и пространственные данные;
• поддержка географических и проекционных систем координат;
• наглядное представление данных в виде нескольких слоев;
• преимущественно визуальные методы пространственного анализа

СУБД

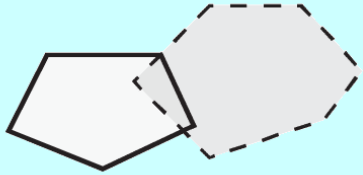
• разделение на атрибутивные и пространственные данные;
• поддержка географических и проекционных систем координат;
• поддержка реляционных свойств пространственных данных – нарушение принципов топологичности ГИС;
• численные методы пространственного анализа с любым количеством «слоёв»

Современные ГИС имеют в своей основе поддержку СУБД,
Современные СУБД работают с пространственными данными

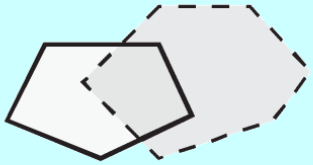
Типы пространственных данных, точка, линия, полигон, мультиполигон, растр. Реляционная природа пространственных данных. Операции пересечения, объединения, нахождения внутри, дополнительные операции



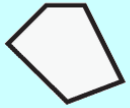
Point.STIntersects(Polygon) = 1



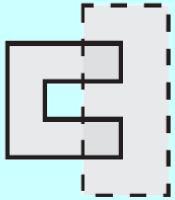
Polygon.STIntersects(Polygon) = 1



Polygon.STIntersection(Polygon)



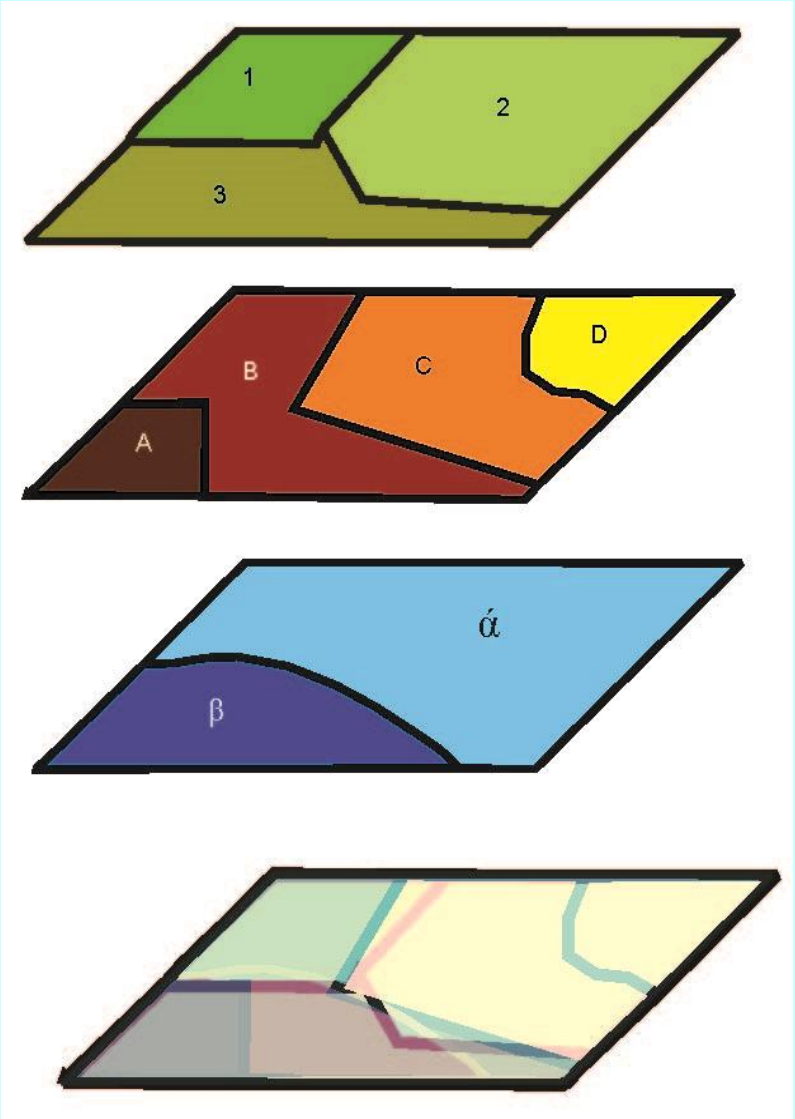
Polygon



Polygon.STIntersection(Polygon)

MultiPolygon

Примеры нарушения правил Кодда – осколочные полигоны

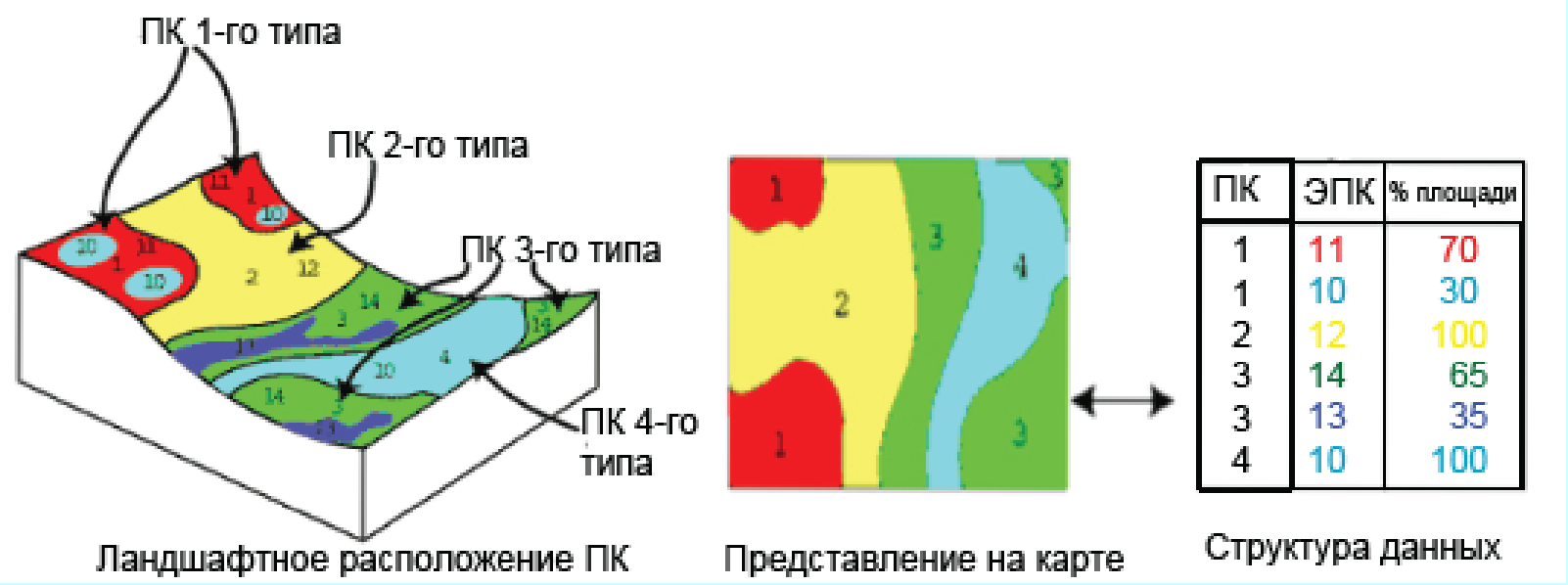


Соотнесение административно-кадастровых и естественно-природных границ.

При выполнении операции пересечения **разнородных** пространственных фигур атрибутивные таблицы **складываются!!!**

Примеры нарушения правил Кодда:

Особенности представления почвенных комбинаций

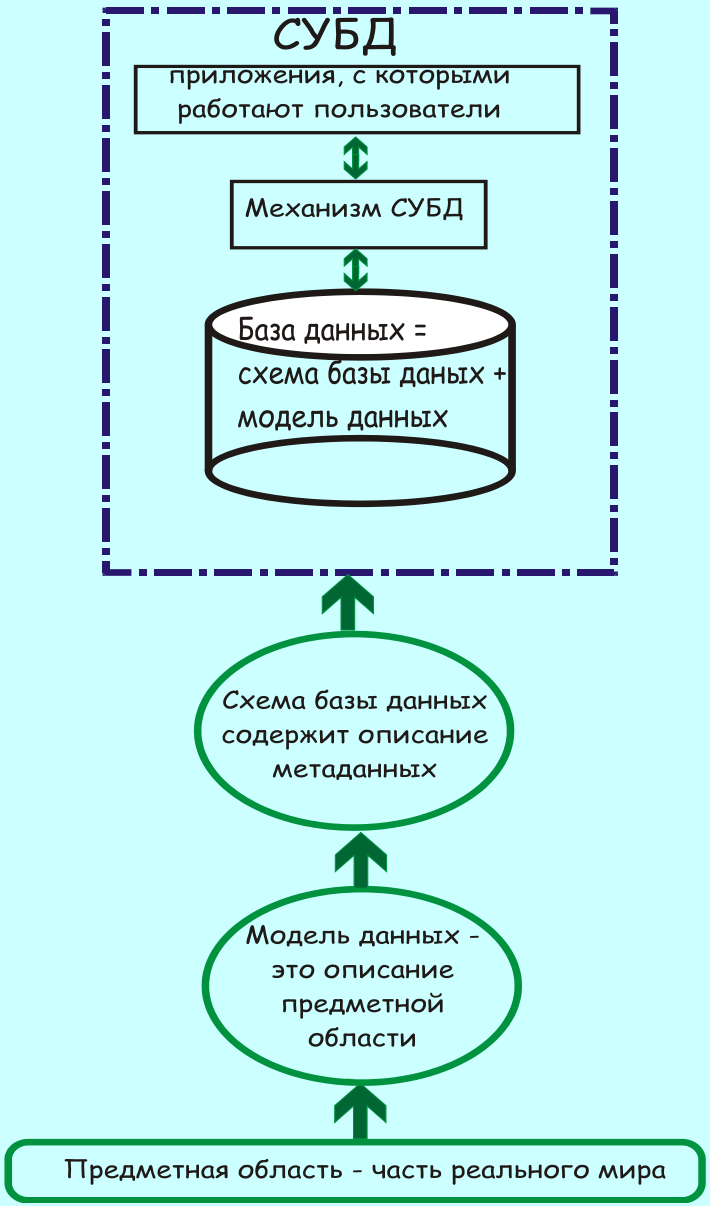
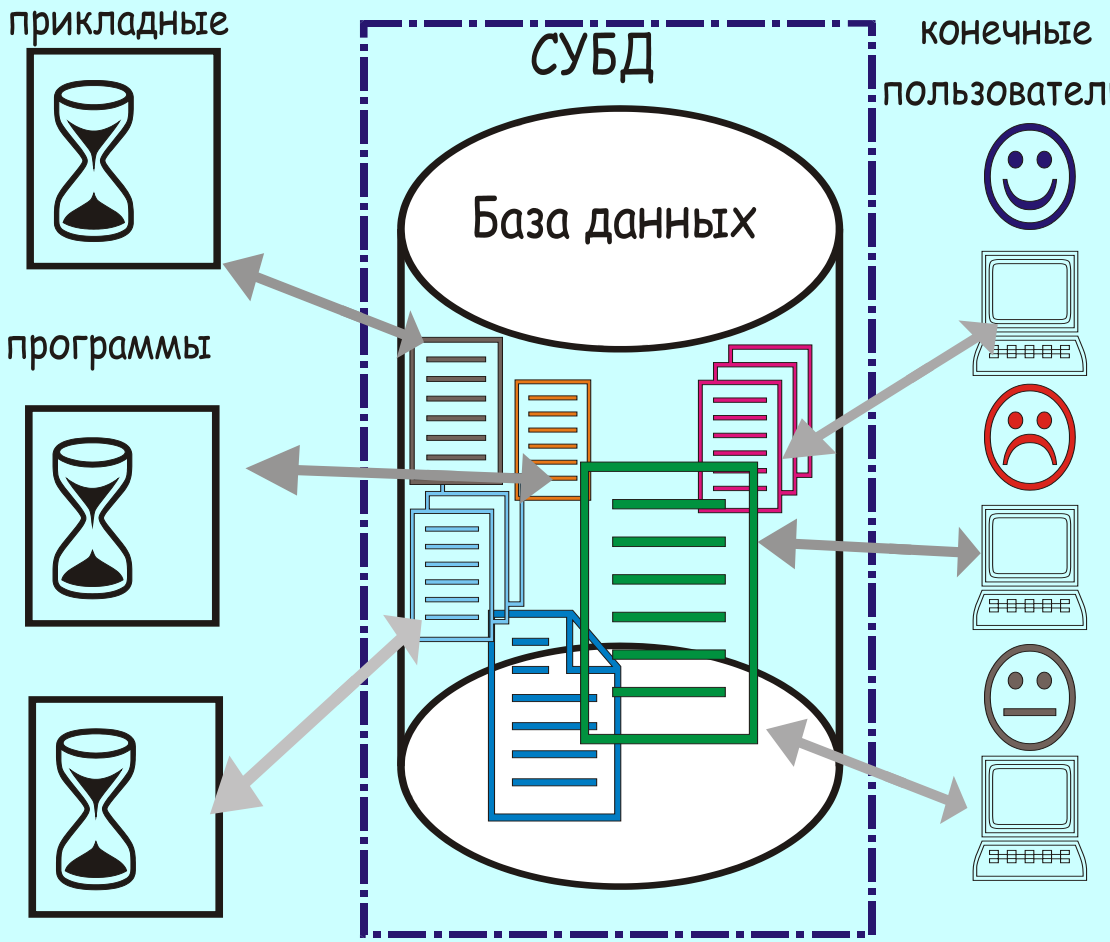


ПК	ЭПК	% площади
1	11	70
1	10	30
2	12	100
3	14	65
3	13	35
4	10	100

Схема представления данных при описании комплексов. Почвенные контуры 1,3 содержат комплексы. В итоге карта из 4 контуров представлена шестью контурами с наложением и весовыми коэффициентами. Это схема представления почвенного контура с комплексами, в виде слоев с весом.

В случае если значения некоторых атрибутов в комплексе совпадают, веса складываются.

Теорема о «неполноте»
доказана Куртом Гёделем в 1930 году
(опубликована в 1931)

Вывод: невозможность построения почвенно-географической СУБД, полностью соответствующей правилам Кодда.

Жизненный цикл почвенно-географических баз данных, программ, поддержка и сопровождение