

Нейронная экспертная система для анализа и картирования процессов засоления почв по данным дистанционного зондирования

Х.Р. Исмагова

*Институт Космических Исследований Природных Ресурсов
(ИКИПР АНАКА НАНА)*

AZ1106 г. Баку, пр. Азадлыг, 159

E-mail: khasiyat@box.az

Предлагаемая нейронная экспертная система является своеобразным мостом между системой обработки данных дистанционного зондирования Земли и геоинформационной системой. Реализован многоступенчатый итеративный процесс распознавания и интерпретации изображений на основе методов, закладываемых пользователем в базу знаний. С этой целью использовались оценки пригодности земель по классификации LCCS FAO (Land Cover Classification System FAO UN) и вовлечение их в обученные искусственные нейронные сети, а также частные шкалы оценок почв исследуемого региона.

Введение

В последние десятилетия дистанционное зондирование Земли из космоса является интенсивно развивающейся областью исследований. Было получено большое количество интересных научных результатов, разработана обширная научно-методическая база, сконструированы и реализованы приборы, установленные на космических аппаратах, успешно работающих на орбите. Накоплен огромный объем фактических данных, для которых, с одной стороны, требуется обеспечение средств хранения, представления и обеспечения доступа, а с другой — осмысление, интерпретация и включение в научный аппарат исследования нашей планеты.

В настоящее время исследователи перешли к этапу применения полученных знаний в прикладных целях: отслеживание состояния природных и антропогенных объектов на поверхности Земли, оценка и прогнозирование процессов, имеющих значение важное, как для отдельных государств, так и для планеты в целом. Наступил этап создания гибридных интеллектуальных систем на основе интеграции различных технологий обработки информации.

В данной работе развивается подход, основанный на использовании нейронных экспертных систем в задачах распознавания, интерпретации изображений и оценки процессов засоления тестового участка. С этой целью использовались оценки пригодности земель по классификации LCCS FAO (Land Cover Classification System FAO UN) и вовлечение их в обученные искусственные нейронные сети [1], а также частные оценки почв исследуемого региона [2]. Использовалась искусственная модель нейронных сетей с обратным распространением ошибки. Обученные нейронные сети были сохранены в базе данных, которая представляла базу знаний. Эта база знаний использовалась в оценке засоления земель при изучении процессов опустынивания территории Абшеронского полуострова. Схема обучения нейронной экспертной системы построена на взаимосвязи с интегрируемой с данными дистанционного зондирования геоинформационной системой.

Тематическая ориентация нейронной экспертной системы непосредственно связана с логикой выделения территориальных операционных единиц, характеризующихся различными видами измерений информации об объектах исследований, то есть локальными ландшафтными структурами (ЛЛС) исследуемого региона.

Таким образом, одной из главных задач является задача тематической ориентации нейронной экспертной системы к задачам распознавания и интерпретации процессов засоления почв по данным дистанционного зондирования, организации её логической структуры, способов функционирования и экспертно-информационного обеспечения, а также организации простого доступа к информации.

Задачи построения системы

Основным элементом любой задачи распознавания изображений является ответ на вопрос: относятся ли данные (входные) изображения к классу изображений, который представляет данный эталон? Казалось бы, ответ можно получить, сравнивая непосредственно изображение с эталонами (или их признаки). Однако возникает ряд трудностей и проблем:

1. Изображения предъявляются на сложном фоне;
2. Изображения эталона и входные изображения отличаются положением в поле зрения;
3. Входные изображения не совпадают с эталонами за счет случайных помех;
4. Отличия входных и эталонных изображений возникает за счет изменения освещенности, подсветки, локальных помех;
5. Эталоны и изображения могут отличать геометрические преобразования, включая такие сложные как аффинные и проективные;
6. Интерпретация изображений зависит от опыта дешифровщика, полноты дополнительной информации по конкретной задаче (условия неопределённости).

Перечисленные обстоятельства обуславливают появление неопределённости в условиях задачи распознавания изображений и нечётких высказываний при их интерпретации, когда привлекаемая дополнительная информация может быть неполной и дешифровщик принимает решение на основании своего опыта. Иначе, при любых способах обработки аэрокосмической информации (визуальном, интерактивном или с помощью нейросетевых технологий), все они в конечном итоге неизбежно сводятся к экспертной интерпретации [3]. Наилучший выход в этой ситуации находят в интеграции с данными дистанционного зондирования экспертных знаний, геоинформационных систем и технологий искусственных нейронных сетей (ИНС).

Эти обстоятельства обуславливают обращение к теории экспертных систем, где одним из важных ограничений при их использовании является трудность формулировки правил для машинной обработки. А при использовании совместно с геоинформационной системой - это трудность формулировки правил передачи и преобразования экспертизы оценки территорий от источника знания до программы. В этом виде в экспертной системе для оперативной выборки знаний должна быть разработана эффективная методология фиксации, сохранения и использования экспертных знаний [3].

Альтернативный метод фиксирования экспертных знаний, не используя правила, заключается в применении искусственных нейронных сетей, используя их способность к обобщению, самообучению и переобучению. Их преимущество также заключается в возможности работы в «on-line» режиме, адаптивности и быстрой приспособляемости в конкретных ситуациях.

Таким образом, исходя из того факта, что изучение и оценка процессов засоления почв - это процесс сложный и многофакторный, а специфика ландшафта оказывает большое влияние на возникновение и расположение засоленных участков ставятся следующие задачи:

1. Используя стандартные процедуры предварительной и тематической обработки изображений выделить локальные ландшафтные структуры на исследуемой территории по космоснимку Ландсат ТМ (использовать ENVI 3.2);
2. Собрать и систематизировать данные контактных традиционных исследований: системы частных шкал оценок почв Кавказа и оценки пригодности земель по системе LCCS ФАО в единой экспертно-информационной системе для оценки процессов засоления (использовать MS ACCESS);
3. Создать картографическую базу знаний в виде векторных слоёв ArcView 3.2 и привязанных к ним атрибутивных таблиц картографических показателей;
4. Интегрировать полученные при выполнении трёх первых этапов знания, используя технологии нейронных экспертных систем;
5. Создать карты процессов засоления почв Апшеронского полуострова на базе интегрированных знаний.

Интеграция знаний о ландшафте исследуемой территории

Знание особенностей почв и форм рельефа может помочь в идентификации типичных засоленных участков или для определения площадей, требующих срочного вмешательства, и разработке мер по повышению плодородия земель.

В соответствии с этими требованиями по спутниковым и самолетным снимкам методами визуального и цифрового дешифрирования (использовались стандартные процедуры ENVI3.2) были выделены локальные ландшафты, отвечающие легенде землепользования, на каждую ландшафтную структуру собрана информация о форме рельефа, литологии и типе почв района.

В геоинформационной системе полученные карты использовались в качестве базовой информации для выработки многоуровневой схемы сбора дополнительных наземных данных и в качестве маски в процессе классификации изображений. Карты также использовались при моделировании связей между пространственным расположением засоленных земель и формами рельефа, для отображения степени опасности засоления почв. На рис.1,2 приведены карты, полученные при дешифрировании космоснимка Ландсат 1998г., оцифровки топографической карты масштаба 1:100 000 (система измерений 1942г, Пулково, проекция Гаусса-Крюгера).

При создании карты типов деградации земель были привлечены экспертные знания. Карта локальных ландшафтов Апшеронского полуострова, полученная при дешифрировании снимка содержит 63 наименований ландшафтов, с указанием типов почв и растительного покрова.

Таким образом, необходимую информацию для оценки засоления почв можно собрать в единую информационную среду посредством:

1. Дешифрирования аэрокосмической информации;
2. Оцифровки тематических и топографических карт;
3. Привлечения знаний экспертов и измерений при полевых исследованиях.

Частные оценки экспертов, характеризующие свойства и типы почв

Почвенные характеристики Апшеронского полуострова изучены и знания хранятся в табличном и картографическом видах. На основе этих источников информации была составлена база данных частных шкал оценок: во-первых – это система частных оценок почв Кавказа, во – вторых, оценки, принятые в системе LCCS [2].

База данных системы частных шкал оценок содержит следующие оценки в табличном виде [3]:

Балл бонитета почв с учётом всех экологических условий;

1. Градации по плотному остатку;
2. По $0 < T < 10^{\circ} \text{C}$;
3. По величине Md;
4. По Kp;
5. По КУ
6. По величине БКГ;
7. По осадкам;
8. По солонцеватости;
9. По степени эродированности;
10. По коэффициенту континентальности;

Исходная информация задавалась в виде таблицы 1, где приведены две категории оценок.

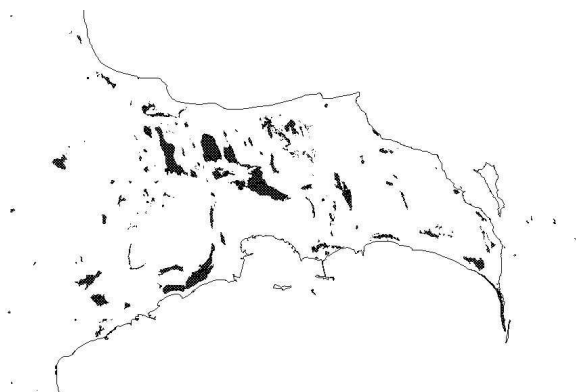


Рис. 1. Карта-схема озёр, озерков, солончаков и различных озёрных понижений (масштаб 1:100 000, полученная по космоснимку Ландсат ТМ, 1998г.

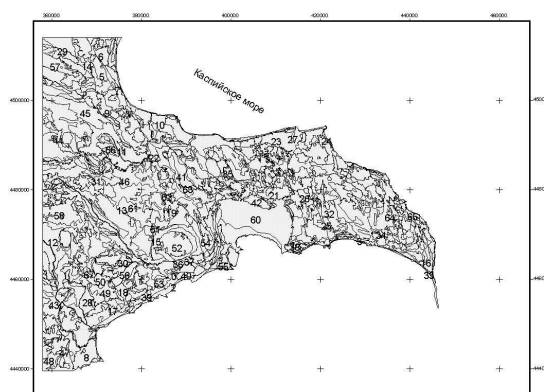


Рис. 2. Выделение локальных ландшафтных структур. Всего 997, составляющих 63 типа ландшафтов. Масштаб 1:100 000.[4]

Таблица 1

<i>Система частных шкал оценки почв по степени проявления отдельных их признаков по определению степени засоления</i>		
Градация по плотному остатку (по В.Р.Волобуеву, 1965)		Оценка пригодности земли
< 0,10	не засолена	лучшие
0,10-0,25	очень слабо засолена	хорошие
0,25-0,50	слабо засолена	средние
0,50-1,00	средне засолена	низкие
1,00-2,00	сильно засолена	условно непригодные
2,00-3,00	очень сильно засолена	условно непригодные
>3,00	солончаки	условно непригодные
по солонцеватости		
Степень солонцеватости (по Антипову-Каратаеву,%)		Оценка пригодности земли
< 5	несолонцеватые	лучшие
5-10	слабосолонцеватые	хорошие
10-15	среднесолонцеватые	средние
15-20	сильносолонцеватые	низкие
20-25	очень сильносолонцеватые	условно непригодные
>25	солонцы	условно непригодные

Таким образом, составной частью информационной среды будут табличные данные, содержащие как оценки экспертов по свойствам и типам почв, так и картографические показатели исследуемой местности. Все таблицы (их 23) собраны в базу данных MS ACCESS. Естественно, что в базе данных таблицы систематизированы, поиск нужных данных осуществляется информационно-поисковой системой, основанной на двух показателях – по пространственному индексу и по метаданным. Существенным является то, что в MS ACCESS существующие взаимосвязи прозрачны и выделяются через схему данных. В нашей ситуации база данных – это набор эталонных таблиц по оценкам экспертов для разрабатываемой нейронной экспертной системы (НЭС).

Этап обработки в ГИС. Осуществлялся сбор атрибутивной информации с оцифрованных тематических карт, приведённых в единую картографическую проекцию. Они составят таблицы картографических показателей по физико-химическим свойствам почв, которые будут классифицированы по принятой в системе LCCS классификаторам природных и антропогенных факторов влияния на качество земель и типу ограничений: климатическое ограничение (K_c), топографическое (K_t), по влажности (K_w), физических и химических свойств почв (K_f), солёность и щёлочность почв (K_n)[2,3].

Таблица 2

ID	Названия локальных ландшафтов	Свойства по K_c	...	Свойства по K_n	Mesto
1	Озеро	r_{11}		r_{1t}	K-39-124B
5	Вулканический конус	r_{51}		r_{2t}	K-39-124C
6	Прибрежная равнина	r_{61}		r_{6t}	K-39125C

Рис. 3. Операционные единицы (а), заданные в виде контуров ЛЛС и таблица(б) связаны по пространственному индексу «Mesto» в ГИС и в базе данных. MS ACCESS

Структура таблицы по картографическим показателям задана в виде таблицы 2, где код ID связан с элементарным участком листа карты, получившим номер ID при визуализации карты классификации в ГИС. Поля «Свойства» – это набор тематических свойств (K_c K_t K_w K_f K_{ch} , K_n), приуроченных к данным локальным ландшафтам при оцифровке соответствующих тематических карт и знаний экспертов. Поле «Mesto» - пространственный индекс, указывающий принадлежность локального ландшафта по местоположению в квадрате разбиения листа карты по номенклатурной разграфке масштаба 1:100 000, это поле является связующим звеном между ГИС и MS ACCESS (рис.3). Пример систематизированных таким образом по характеру и степени засоления показателей по Апшеронскому полуострову приведён в таблице 3. Установлена взаимосвязь таблиц в ГИС и MS ACCESS.

Таблица 3. Пример заполнения таблицы картографических показателей

База знаний НЭС				
ID ЛЛС	Степень засоления	Ограничения	Характер засоления	Mesto
9	Незасоленные	< 0,25	SO ₄ -HCO ₃ -Na	2.1
26	Незасоленные	< 0,25	SO ₄ -Cl-Na	2.2
55	Незасоленные	< 0,25	SO ₄ -Na	1.2
1	Слабозасоленные	0,25 - 0,5	SO ₄ -Ca	1.2
5	Слабозасоленные	0,25 - 0,5	SO ₄ -HCO ₃ -Ca	1.2

База знаний НЭС				
ID ЛЛС	Степень засоления	Ограничения	Характер засоления	Mesto
43	Среднезасоленные	0,5 - 1,0	SO ₄ -Cl-Na	1.3
67	Среднезасоленные	0,5 - 1,0	SO ₄ -Na	1.1
48	Сильнозасоленные	1,0-2,0	SO ₄ -Cl-Na	1.1
80	Сильнозасоленные	1,0-2,0	SO ₄ -Na	1.1

Здесь, ID- номер элементарного участка территории (контур), полученный в ГИС; степень засоления – разделена на 5 категорий (в таблице не показана категория солончак), ограничения – пороги ограничений, накладываемые экспертами; характер засоления - химический состав; Mesto – пространственный индекс по местоположению элементарного участка по условию разделения всего участка на квадраты, согласно номенклатурной разграфке в проекции Гаусса-Крюгера (масштаб 1:100 000, рис.3.а), для удобства номенклатурные номера К-39-124В заменены порядковым номером (замена производилась соответственно в ГИС и в MS ACCESS).

Математическая постановка задачи

Пусть в результате классификации изображения по спектральным характеристикам получена векторная модель, определяющая контура локальных ландшафтов исследуемой местности, то есть исследуемая местность разделена на элементарные участки по типам ландшафтов, составляющих множество V элементарных участков территории. Для того, чтобы отобразить в необходимой форме исходную картографическую информацию на данную территорию предполагается, что информация о полезной нагрузке карт, хранимая в памяти компьютера в цифровом виде, будет представлена в виде матрицы R размерности $(M \times N)$, имеющий вид [3]:

$$R = \left\| r_{i,j} \right\|, \text{ где } i = 1, \dots, M; j = 1, \dots, N. \quad (1)$$

Каждый элемент матрицы R является вектором параметров, характеризующих каждый (i,j) -й элементарный участок листа карты по некоторому m -му ($m = 1, \dots, T$) набору тематических свойств:

$$r_{i,j} = (r_{ij}^1, \dots, r_{ij}^T), \quad (2)$$

природа компонентов векторов $r_{i,j}$ в общем случае принципиальной роли не играет.

Тогда, если задаётся набор тематических свойств $\{P_n\}$, ($n = 1, \dots, K$), по которым необходимо районировать исследуемую территорию и пороговые границы ограничений по всему набору тематических свойств, то требуется для каждого P_n по их пороговым ограничениям поставить в соответствие множество $V_i \in V$, ($1 \leq i \leq (M \times N)$) элементарных участков земной поверхности (ЛЛС), которые обладают свойством P_n .

В нашем случае вектор P_n задан следующим образом:

$$P_n = (P_{Kc}, P_{Kt}, P_{Kw}, P_{Kf}, P_{Kch}, P_{Kn}),$$

где каждый P_{Ki} - одно из значений из набора признаков, по которым известно, что каждый признак ранжирован по порогам ограничений (рис.4).

В НЭС каждые из соответствующих наборов признаков представлены эталонными таблицами ранжирования по порогам ограничений, следуя системе частных шкал (таблицы типа1), сгруппированных по типам физико-химических свойств K_c , K_t , K_w , K_f , K_{ch} , K_n . Соответственно, согласно условиям ранжирования экспертами вводится оценка отнесения к одному из классов пригодности земли ($S_1, \dots, S_5$), (S_1 - лучшие; S_2 - хорошие; S_3 - средние; S_4 - низкие; S_5 - условно непригодные).

При такой постановке задачи ниже описывается итеративная процедура картографирования процессов засоления на базе интеграции данных дистанционного зондирования, НЭС и ГИС.

На первом этапе НЭС по системе эталонных таблиц обучается выносить решение по каждому из шести подклассов задания тематических свойств данного региона (рис.5). Решение выносится о принадлежности к одному из классов пригодности ($S_1, \dots, S_5$) по следующей схеме:

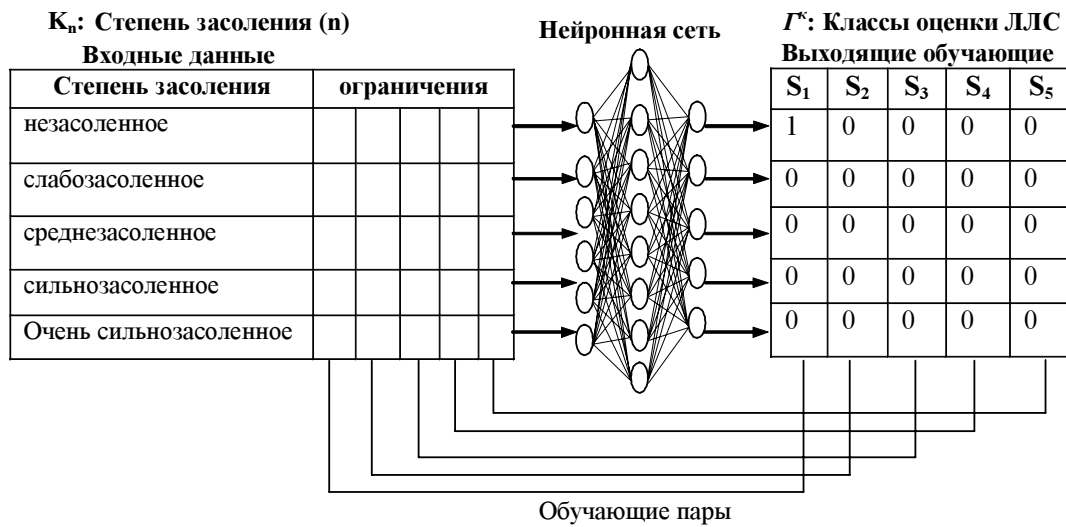


Рис. 4. Схема обучения нейросетевого алгоритма по эталонной таблице (K_n)

Тогда, показывая НЭС всевозможные варианты значений признаков в рамках порогов ограничений по каждой k -ой эталонной таблице на выходе имеем матрицы оценок Γ^{kl} (рис.4):

$$\Gamma^{kl} = \|q^{kl}_{u,v}\|, \quad q^{kl}_{u,v} \in [0,1], \quad (3)$$

где ответ 1-означает, что по набору признаков P_{Kc} и по остальным соответственно их порогам ограничений выносится решение о принадлежности к одному из классов пригодности и 0 – в противном случае, l - количество вариантов.

В процессе обучения для фиксированной паре uv выносится решение отнесения к классу пригодности $S_v \in (S_1, \dots, S_5)$ по функции принадлежности $\mu_k(uv)$ если:

$$\mu_k(uv) = \max_{i,j} \mu_{i,j}(\Gamma^k), \quad (4)$$

где $\mu_{i,j}(\Gamma^k) = \frac{1}{l} \sum_{i=1}^l w_{i,j}^l q_{i,j}^{kl}$, $w_{i,j}^l$ - коэффициенты относительной важности по l -му варианту

отнесения к классам ($S_1, \dots, S_5$), вычисляемые как веса нейронной сети, которые хранятся в базе знаний [3].

Следовательно, если НЭС выносит решение по паре uv , то это означает, что по набору признаков P_{Kc} по пересечению u – ой строки и v -го столбца имеется максимальное количество голосов. Подобным образом, НЭС обучается по всем эталонным таблицам.

Все шесть сетей имеют пять узлов вывода и пять возможных выводных обучающих образцов. Во время обучения, сети "учились" генерировать один из образцов вывода по схеме на рис. 5. Полные классы пригодности выносятся по схеме $S_3 K_c$, если данный регион по климатическим условиям классифицируется как средне пригодный со слабозасоленными почвами.

Следующий этап реализации многоступенчатого итеративного процесса - выделение и классификация объектов на изображении на основе методов, закладываемых пользователем в базу знаний. Каждый полигон ранее (ЛЛС) отнесён к классам объектов соответственно легенде землепользования с указанием типов почв. Таким образом, закончен первый вариант дешифрирования изображения. Теперь, на основе

проведенной классификации и сравнительного анализа с возможными предыдущими классификациями или картографическими данными возможны операции разбиения, слияния и изменения классовой принадлежности сегментов ЛЛС на основе их пространственных отношений, динамики и методов, закладываемых пользователем в базу знаний.

Эта процедура может применяться многократно, и на каждом шаге правила, заложенные в базу знаний, могут корректировать форму и принадлежность объектов. Полученные в итоге объекты могут быть анализированы с помощью связи ГИС и НЭС, при этом в ГИС они могут сопровождаться атрибутами, описывающими свойства объектов, которые могут использоваться для пространственного анализа объектов.

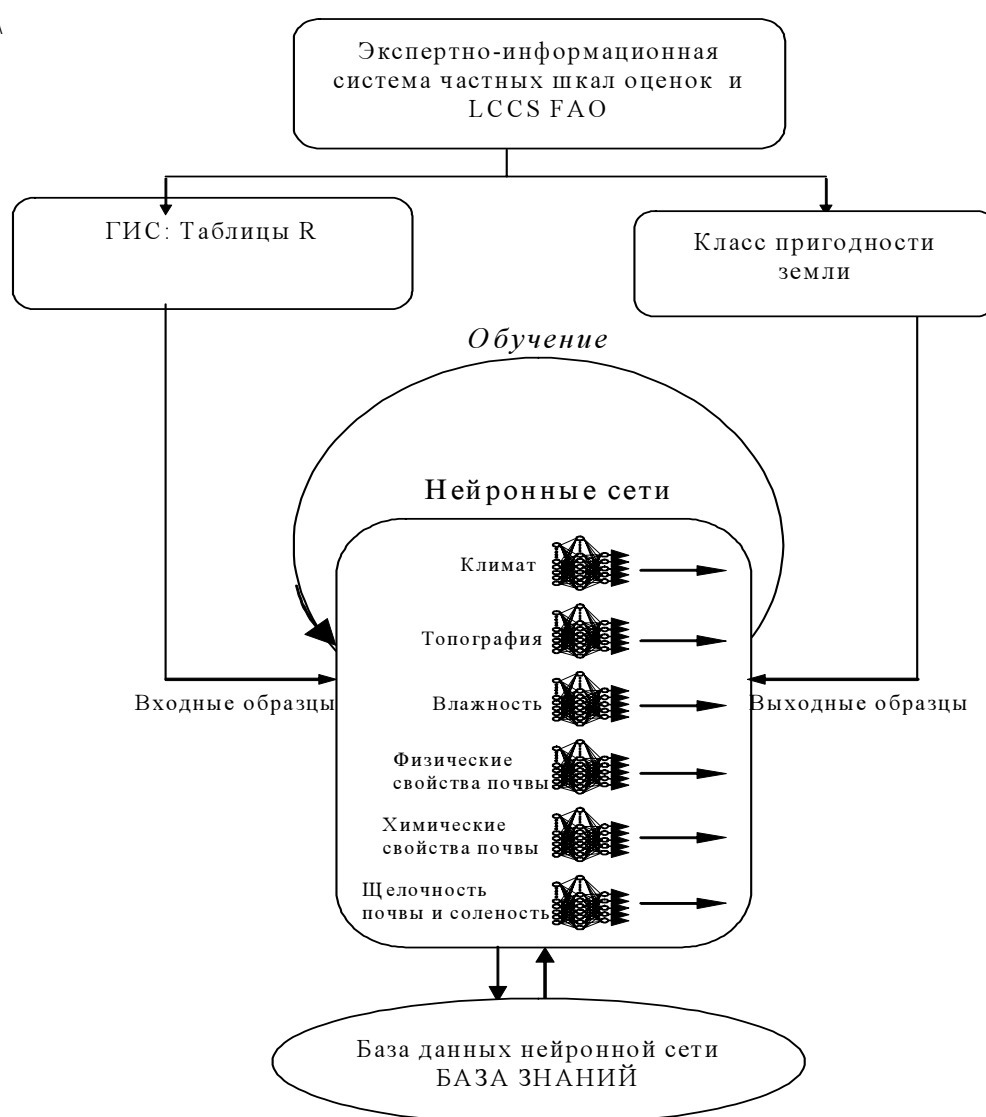


Рис. 5. Обобщённая схема обучения искусственных нейронных сетей

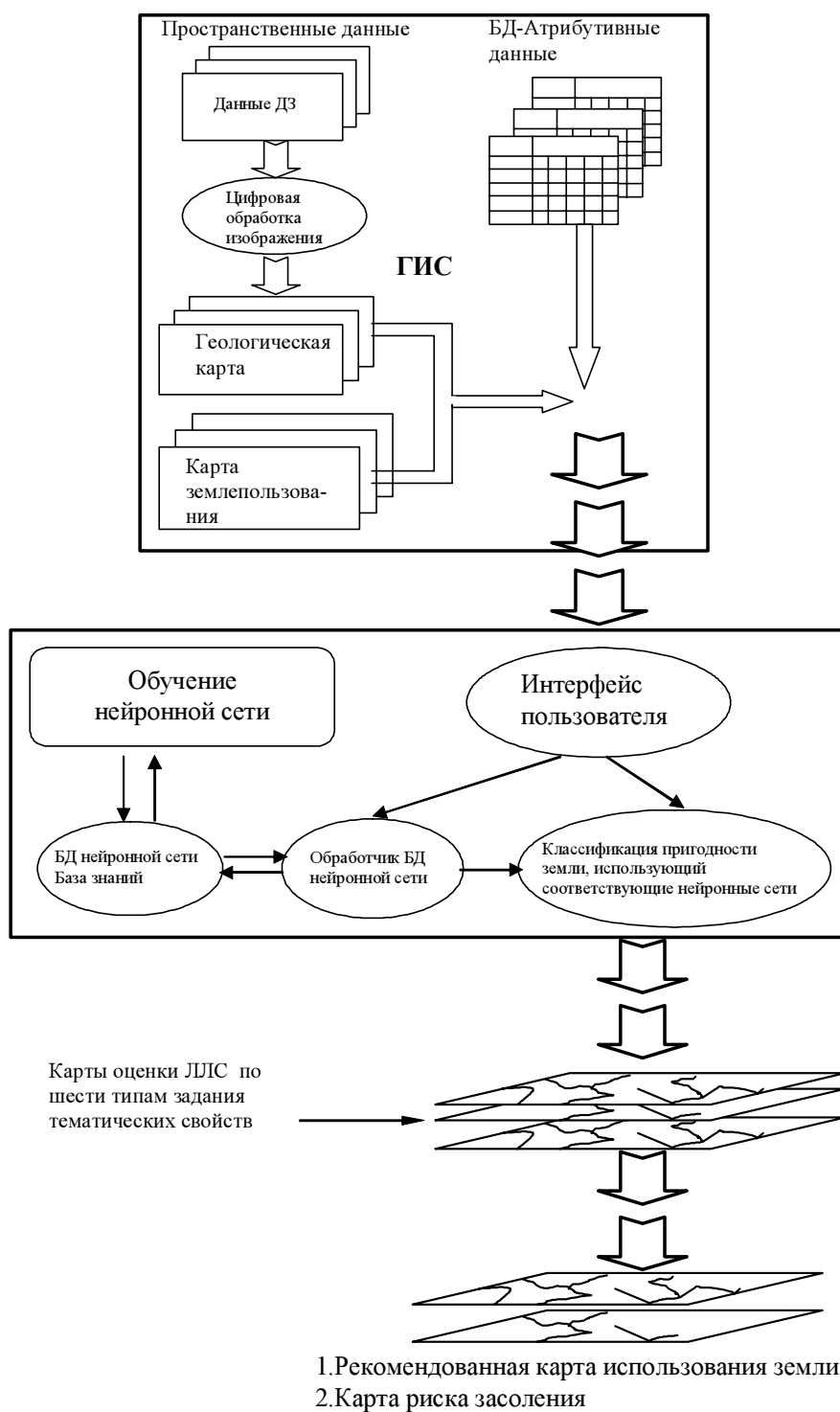


Рис. 6. Автоматизированная система оценки земли, использующая обученную нейронную сеть

Этап цифрового картографирования

К данному этапу мы имеем обученную по эталонным таблицам НЭС, полную картографическую информацию в ГИС (литология, формы рельефа, типы почв и растительности, землепользования). Применяя метод наложения тематических слоёв, мы можем моделировать, получая новые картографические модели и тем самым, формируя достаточно полную информационную модель, чтобы принимать объективные решения по конкретной задаче (рис.6).

Таким образом, по информации из ГИС формируется вектор физико-химических параметров почв данного региона для каждой ЛЛС

$$P_n = (P_{Kc}, P_{Kt}, P_{Kw}, P_{Kf}, P_{Kch}, P_{Kn}) ,$$

который вводится в НЭС для сравнения с её эталонной базой знаний, чтобы принять решения о степени засоления и роли различных факторов в этом процессе. Решение НЭС заносится в базу данных соответствующей тематическому слою карты оценок почв ЛЛС в ГИС по шести типам задания их свойств (рис.6). В результате, используя базу знаний и выполняя этапы описанной выше итеративной процедуры получена карта процессов засоления почв Апшеронского полуострова.

Заключение

Таким образом, база знаний для Апшеронского полуострова была сгенерирована в форме обучаемых нейронных сетей на базе системы частных шкал оценки почв по степени проявления отдельных их признаков при определении степени засоления. По многозональному снимку Ландсат 1998 г. получена карта землепользования (ЛЛС). Каждая единица ЛЛС в ГИС имеет атрибутивную базу картографической легенды по физико-химическим параметрам, типам почв и растительности, литологии данного региона. На этапах реализации многоступенчатого итеративного процесса, описанного в статье, была получена карта по процессам засоления исследуемого региона, фрагмент которой приводится на рис.7.

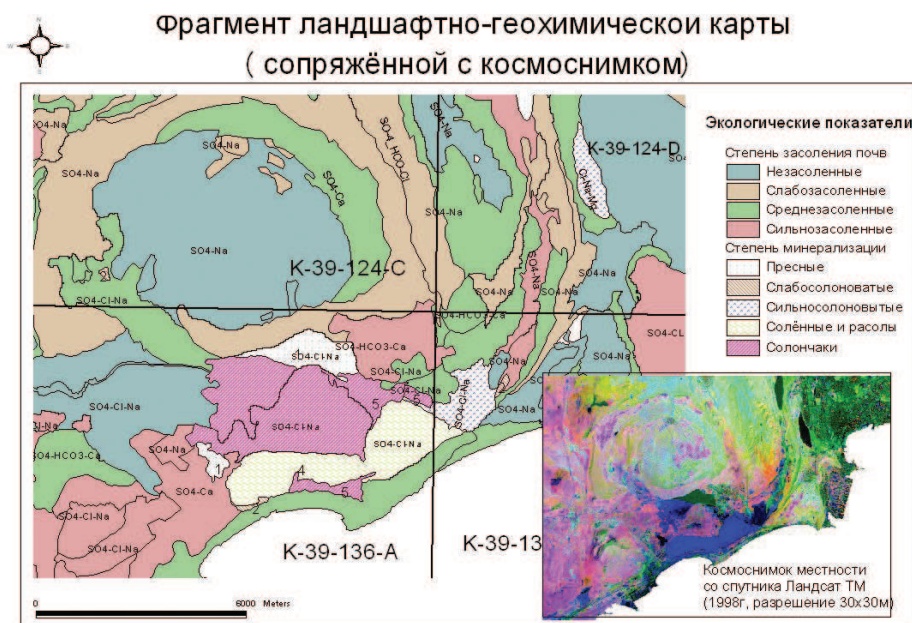


Рис. 7. Фрагмент комплексной карты процессов засоления почв Апшерона

Литература

1. Antonio Di Gregorio. The Land Cover Classification System (LCCS) // Classification Concepts and User: FAO UN, Rome, 1998. 95 с.
2. Мамедов Г.Ш. Карта экологической оценки почв Azerbaijan и её значения // Обзорная информация АЗНИИТИ, 1992. Серия «Сельское хозяйство». С.1-25.
3. Исмамова Х.Р., Талыбова С.С., Абдуллаева С.Н. Один способ интеграции экспертной базы знаний в систему тематического дешифрирования изображений // Сб. научных трудов Национальной Авиационной Академии, посвященный 70-летию ак. А.М-Дж. Пашаеву. Баку, 2004. «Cild 6». №1. С.138-145.