

СОХРАНЕНИЕ БИОРАЗНООБРАЗИЯ И МОНИТОРИНГ СОСТОЯНИЯ ВИДОВ

Аналитическая компонента базы данных

1. Введение

Всероссийская база данных содержит информацию о состоянии тех видов, которые являются особо ценными и находятся под угрозой исчезновения. Информационно-аналитическая система (ИАС) задумана как механизм мониторинга состояния видов и посвящена проблемам управления процессами сохранения и устойчивого использования видов. ИАС позволяет анализировать состояние видов на различных уровнях – национальном, региональном и глобальном.

Аналитическая компонента ИАС состоит из следующих блоков:

1. Блок построения **индексов состояния редких видов** на основе экспертных оценок. Результатом работы блока построения индексов является *набор индексов* для каждого вида ИАС, который удовлетворяет условиям устойчивости), есть его значение не изменяется при незначительном изменении экспертных оценок критериев или добавлении новых видов), непротиворечивости (индекс не противоречит интегральным оценкам экспертов, или ранжированию видов), адекватности и согласованности (индекс соответствует модели связи между критериями, одобренной экспертами).

В блок входят: алгоритм сбора экспертных оценок и проверки непротиворечивости экспертных оценок; алгоритм построения индексов состояния видов; алгоритм согласования экспертных оценок.

2. Блок **выявления угрожаемых видов и выявления приоритетов**. Этот блок выполняет мониторинг изменения состояния редких видов (определенного на множестве базовых критериев) и изменения индексов состояния видов. Результатом работы блока является классификация видов по категориям угрожаемости и обоснованное *изменение категории статуса* видов. Для этого должны быть введены эвристические правила классификации видов в зависимости от изменения их состояния. При значительном комплексном изменении состояния вид классифицируется как угрожаемый и заносится в отдельный список. По результатам анализа причин возникновения угрожаемых видов составляется список угроз, который определяет *приоритетные меры* по охране того или иного класса видов.

3. Блок **прогноза состояния видов**. Блок подразумевает детальный анализ поведения заданного вида (его фазовой траектории в пространстве измеряемых показателей). Этот анализ производится на основе привлеченных данных по указателям, которые хранятся в ИАС. На основе имеющихся временных рядов делается прогноз состояния вида и анализ соответствия этого прогноза и изменения приоритета данного вида. Блок состоит из алгоритма предобработки и шкалирования измеряемых данных, алгоритма настройки эвристических функций, отображающих пространство измеряемых показателей в пространство производных описаний и алгоритма настройки регрессионной модели, отображающей пространство производных описаний в пространство прогнозируемых переменных состояния видов.

4. Блок **анализа адекватности мер по охране территорий**. Результатом работы блока является оценка экспертизы хозяйственных проектов в области ресурсопользования с учетом мер по охране редких видов. Исходная информацией алгоритма построения оценки содержится в базе списка угроз редких видов и в базе принятых охранных мер. Необходимая составляющая блока – база моделей принятия адекватных охранных мер по охране природных территорий.

В текущей реализации ИАС рассматривались первый и третий блоки: блок построения индексов состояния редких видов и блок прогноза состояния видов. Ниже эти два блока описываются подробнее.

2. Построение индексов на основе экспертных оценок

Дано множество редких видов, которые требуется сравнить, то есть узнать, какие из объектов лучше, а какие хуже. Результатом сравнения является оценка (индекс) — набор чисел, где каждое указывает на состояние. Для построения индексов собираются данные о каждом из объектов. Для этого определен список критериев, по которым оценивается каждый из объектов.

Полученная в результате сбора информации таблица называется таблицей "объект-признак". По вертикали в этой таблице расположены объекты, по горизонтали — показатели.

При помощи этой таблицы получают индексы. Методы получения индексов делятся на две группы. Первая группа дает точную оценку каждому объекту, а вторая дает порядковый номер объекта, то есть шкалу рейтинга. Эти методы получения индексов называются методами "без учителя", так как результат зависит только от собранных данных и применяемого метода. Известны методы главных компонент, расслоения Парето, экстремальной группировки признаков и многие другие.

Приведем алгоритм, используемый в данной работе для получения индексов состояния редких видов. Задано множество $Y = \{v_1, \dots, v_m\}$ редких видов (далее-объектов) и множество критерий, описывающих виды (далее-показателей) $\Psi = \{\psi_1, \dots, \psi_n\}$. Объект v_i описан с вектором-строкой $\mathbf{a}_i = \langle a_{i1}, \dots, a_{in} \rangle$, $\mathbf{a}_i \in \mathbf{R}^n$. Множество описаний объектов представлено в виде матрицы исходных данных $A = \{a_{ij}\}_{i,j=1}^{m,n}$. Объект v_i , имеющий максимальный по значению интегральный индикатор (наибольшую экспертную оценку, если она рассматривается в качестве интегрального индикатора), $q_i = \max\{q_1, \dots, q_m\}$, является наилучшим. Объект v_i , имеющий минимальный по значению интегральный индикатор $q_i = \min\{q_1, \dots, q_m\}$, является наихудшим. Показатель ψ_j , имеющий максимальный по значению вес (наибольшую экспертную оценку, если она рассматривается в качестве веса показателя), $w_j = \max\{w_1, \dots, w_n\}$, является наиважнейшим при нахождении интегрального индикатора. *Индексом объекта* v_i с номером I называется скаляр $q_i \in \mathbf{R}^1$ поставленный в соответствие набору \mathbf{a}_i описаний объекта. При рассмотрении множества объектов Y , вектор $\mathbf{q} = \langle q_1, \dots, q_m \rangle^T$, $\mathbf{q} \in \mathbf{R}^m$ считается интегральным индикатором множества объектов, описанных матрицей $A \in \mathbf{R}^{m \times n}$.

Алгоритм расслоения Парето. Описания \mathbf{a}_i множества объектов Y представлены в виде $T = \bigcup_{\zeta=1}^l S_\zeta$ при условии $S_\zeta \cap S_\eta = \emptyset$ и $\zeta \neq \eta$ где множество Парето ζ -го слоя $S_\zeta = \{\mathbf{a}_i : i \in \{1, \dots, m\}\}$ и l — номер слоя, соответствующий множеству $\{\mathbf{a}_i\}$. Вектор $\mathbf{a}_\xi = \langle a_{\xi 1}, \dots, a_{\xi n} \rangle$ называется недоминируемым, если не найдется ни одного вектора \mathbf{a}_i , такого, что $a_{ij} > a_{\xi j}$, $i = 1, \dots, m$ and $j = 1, \dots, n$. Для всех $\zeta = 1, \dots, l$ множество S_ζ определено как набор недоминируемых векторов, не принадлежащих множеству $S_{\zeta-1}$. В соответствие каждому вектору \mathbf{a}_ξ , $\xi = 1, \dots, m$ поставлен индекс ζ множества S_ζ , которому принадлежит вектор $\mathbf{a}_\xi \in S_\zeta$. Полученное множество индексов $\Xi = \{\zeta_\xi\}_{\xi=1}^m$ преобразуется приводится к стандартному виду таким образом, что индекс $\mathbf{q}_3 = \{\max(\Xi) - \zeta_\xi\}_{\xi=1}^m$.

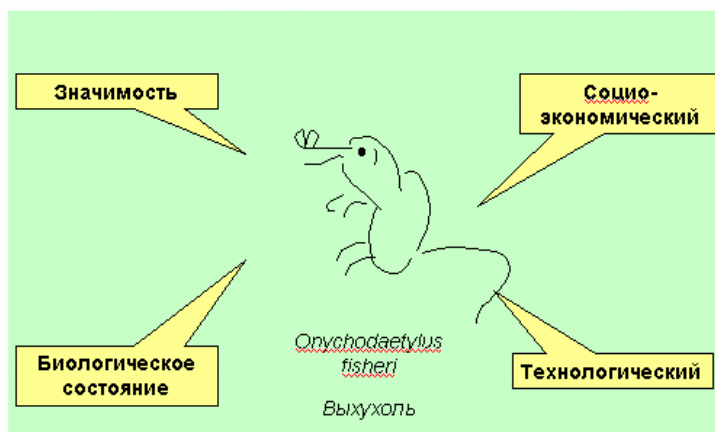
Часто эксперты, хорошо разбирающиеся предметной области, желают назначить оценки объектов самостоятельно. Эти оценки могут быть трех видов. Во-первых, эксперт может сравнить объекты, или назначить какой объект лучше, а какой — хуже. Во-вторых, эксперт может назначить оценки в некоторой удобной ему шкале, то есть точнее, чем в предыдущий раз. Данные оценки называются оценками в ранговых шкалах. И в третьих, эксперт может назначить оценки с большой степенью точности. Эти оценки выставлены в линейных шкалах. Экспертные оценки могут подтверждать или опровергать индексы, полученные с помощью методов "без учителя". Имеет смысл сравнить эти два набора оценок и по возможности, скорректировать или уточнить оценки объектов. Такое уточнение называется согласованием экспертных оценок. Предлагаются специальные методы согласования экспертных оценок, выставленные в ранговых и линейных шкалах.

Так же, как для индексов, эксперт может назначить веса критериев. Пользуясь этими весами, мы можем вычислить оценку объектов. Пусть эксперт назначает оценку объектов. Тогда мы можем по ней вычислить вес критериев. Пусть он делает и то, и другое. Тогда мы можем согласовать оценки объектов, веса критериев и данные. Полученные результаты предлагаются экспертам на обсуждение в следующем виде:

Начальные значения	(веса)	w_0^T
(индексы)	Согласованные значения	w_α^T
q_0	q_α	A

В результате получаем, во-первых, точные оценки объектов, во-вторых, обоснованные оценки объектов, то есть, мы будем знать, какой критерий какое влияние оказал на индекс — так мы объясним оценку по весам экспертов. И, в-третьих, получаем веса критериев, которые могут быть использованы для дальнейшего вычисления индекса.

Проиллюстрируем вышесказанное на примерах. Каждый вид, занесенный в базу данных ИАС, описан несколькими наборами показателей-критериев. В частности приведем четыре из них: значимость, биологическое состояние, социо-экономический набор, технологический набор.



Каждый вид описан с помощью четырех наборов критериев

На основе каждого из наборов должен быть построен индекс. Для этого используются вышеперечисленные методы и алгоритмы, удовлетворяющие определенным требованиям.

1	Индекс I это величина, наиболее информативно отражающая состояние вида.
2	Состояние вида S описывается набором критериев: $S = (k_1, k_2, \dots, k_n)$
3	Индекс вида зависит от состояния вида: $I = f(S)$.
4	Основой алгоритма f , с помощью которого вычисляется индекс I , являются: - метод расслоения Парето, - метод согласования экспертных оценок.

Индекс построен на основе набора критериев

Исходными данными для алгоритма являются наборы критериев, но сами критерии также можно рассматривать как индексы. Таким образом, строится трехуровневая иерархическая система индексов: индексы базового уровня (критерии), индексы второго уровня (по наборам критериев) и индекс (или интегральный индикатор) третьего, верхнего уровня – число описывающее состояние данного вида.

Индекс (интегральный индикатор) вида			
индексы			
Значимость	Биологическое состояние	Социо-экономический	Технологический
Возможные потери Доля ареала Роль в биоценозах	Численность Темп прироста Встречаемость/плотность Площадь ареала Структура ареала Популяционная структура Генетическое разнообразие Половозр.соцо. структура Физиологическое состояние Состояние местообитаний	Ресурсное значение Научное значение Значение в качестве индикатора состояния ПК	Доступность знаний Мониторинг Иск. воспроизводство Реинтродукция Сохранение Ex-situ
Состояние Тренд			
			критерии

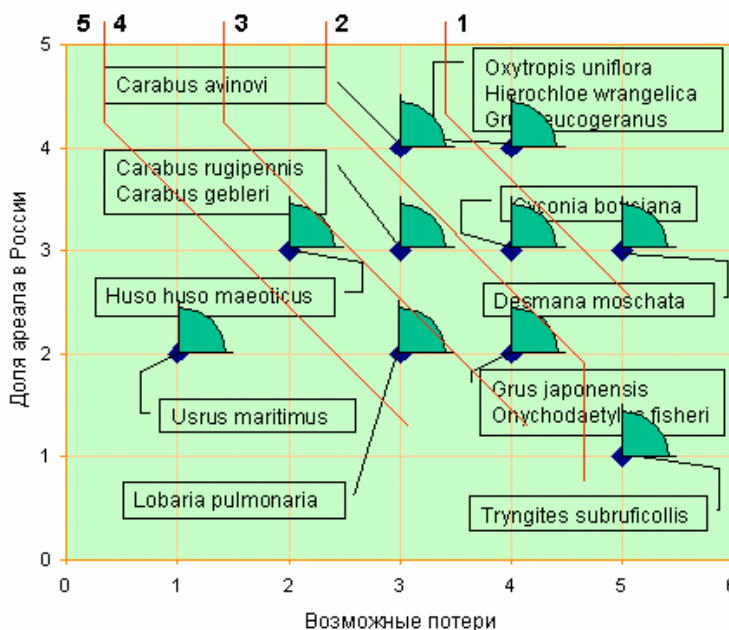
Трехуровневая иерархия индексов

Для данного вида эксперт выставляет оценки в лингвистических шкалах по каждому из критериев. Результатом оценивания видов является нижеприведенная таблица. Таким образом, все виды, занесенные в базу данных, являются сравнимыми в рамках определенных шкал. В первой колонке приведены критерии двух наборов, а в последующих колонках – экспертные оценки, выставленные для четырнадцати видов.

Виды	<i>Carabus rugipennis</i>	<i>Carabus avinovi</i>	<i>Carabus gebleri</i>	<i>Lobaria pulmonaria</i>	<i>Oxytropis uniflora</i>	<i>Hierochloa wrangelica</i>	<i>Huso huso maeoticus</i>	<i>Onychodaetylus fisheri</i>	<i>Cyconia boissiana</i>	<i>Grus leucogeranus</i>	<i>Grus japonensis</i>	<i>Tryngites subruficollis</i>	<i>Desmana moschata</i>	<i>Urus maritimus</i>
Значимость для сохранения биоразнообразия														
✓ Всп. потери	3	3	3	3	4	4	2	4	4	4	4	5	5	1
Доля ареала	3	4	3	2	4	4	3	2	3	4	2	1	3	2
Роль в биоценозе	1	1	1	1	1	2	1	1	1	1	1	1	1	1
Биологическое состояние														
Размер популяции	2	2	1	3	2	1	1	2	1	1	1	2	1	3
то же, тренд	2	2	2	2	3	3	2	3	1	3	3	4	2	4
Темп прироста	1	1	1	1	2	2	2	1	1	1	1	2	1	2
то же, тренд	1	1	1	2	2	2	1	2	1	2	2	2	1	3
Встречаемость/плот	2	2	1	2	3	3	1	1	2	2	2	2	2	3
то же, тренд	2	2	2	2	3	2	1	3	2	3	3	3	2	4
Площадь ареала	2	2	1	3	1	1	1	2	1	1	1	1	2	3
то же, тренд	2	2	2	2	3	2	1	2	2	3	3	3	2	3
Структура ареала	5	5	1	3	1	1	5	2	2	4	3	2	2	5
то же, тренд	3	3	2	1	2	2	2	2	1	3	2	2	1	3
Популяц. структура	1	1	1	2	1	1	2	2	1	2	2	2	1	1
то же, тренд	2	2	2	1	2	2	1	2	2	2	2	2	2	2
Физиол. состояние	3	3	3	3	3	3	3	3	3	3	2	3	3	3
то же, тренд	2	2	2	2	2	2	1	2	2	2	2	2	2	2
Сост. местообитаний	2	2	1	2	3	3	1	3	3	2	2	3	3	3

Результаты экспертных оценок видов по базовым критериям

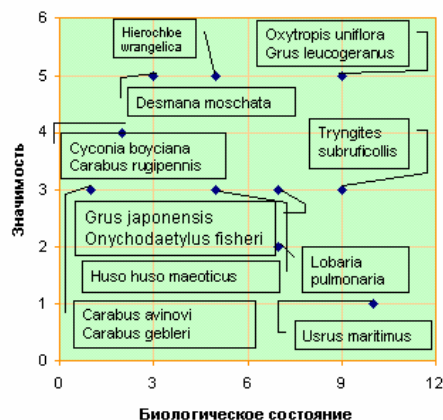
Схематически проиллюстрируем работу алгоритма расслоения Парето. Каждый вид представлен вектором-точкой на плоскости «Возможные потери»—«Доля ареала в России». Каждому слою поставлены в соответствие недоминируемые векторы, не принадлежащие предыдущему слою. На рисунке показаны пять Парето-слоев, в которые вошли четырнадцать видов.



Вычисление индексов с помощью расслоения Парето

Результатом работы алгоритма являются таблица индексов. В первой колонке таблицы приведены вышеупомянутые четырнадцать видов, а во второй и третьей колонке — значения соответствующих индексов второго уровня.

Виды	Значимость	Биологич. состояние
Carabus rugipennis	4	2
Carabus avinovi	3	1
Carabus gebleri	3	1
Lobaria pulmonaria	2	7
Oxytropis uniflora	5	9
Hierochloe wrangelica	5	5
Huso huso maeticus	3	5
Onychodaetylus fisheri	3	7
Cyconia boyciana	4	2
Grus leucogeranus	5	9
Grus japonensis	3	7
Tryngites subruficollis	3	9
Desmana moschata	5	3
Urus maritimus	1	10



Два индекса: Значимость и Биологическое состояние

График, расположенный в левой части рисунка показывает иерархическое соотношение двух полученных индексов, позволяющее построить общий индекс, называемый также интегральным индикатором.

3. Прогноз состояния вида

Состояния вида прогнозируется по результатам анализа временных рядов, описывающих поведение вида в некотором пространстве показателей. Для каждого вида, руководствуясь указателями, хранимыми в базе данных ИАС, составляется собственная система временных рядов-описаний поведения вида. Эта система анализируется на предмет зависимости прогнозируемой переменной от базисных временных рядов. Для этого определяется зависимость переменных состояния вида (далее-объекта) от *экзогенных переменных*. Эта зависимость описывается *математической моделью* объекта. Для определения связи между переменными модели выполняется тест причинно-следственной связи Грэнжера, который заключается в следующем.

Каждая переменная, включаемая в модель управления либо сама влияет на состояние объекта, либо изменяется под влиянием других переменных. Если изменения переменной *a* предшествуют изменениям переменной *b* (при наличии статистической связи между ними), но не наоборот, то переменная *b* зависит от переменной *a*. При выборе переменных для модели управления исключаются безразличные переменные – те, от которых не зависит состояние объекта, и которые не изменяются под влиянием других переменных. Таким образом множество переменных разбивается на подмножества управляемых, неуправляемых и зависимых переменных.

Векторные авторегрессионные модели и модели на основе одновременных уравнений являются эффективными инструментами прогноза состояния вида. В частности, авторегрессионная модель позволяет найти оптимальные управляющие воздействия и спрогнозировать состояние объекта управления при оптимальном управлении.

Прямая задача нахождения состояния $Y = Y_t$ объекта по экзогенным переменным $X = X_t$, согласно векторной авторегрессионной модели, имеет вид

$$Y_t = \sum_{\tau=0}^r (A_{\tau} Y_{t-\tau} + B_{\tau} X_{t-\tau}) + M + \varepsilon_t, \text{ где } Y_t \text{ – состояние объекта управления в момент}$$

времени t , описываемое набором эндогенных переменных; X_t – управляемые и неуправляемые внешние воздействия на объект управления в момент времени t , описываемые набором экзогенных переменных; t – дискретное время $t = 1, \dots, t_0$ и $\tau = 0, \dots, r < t_0$ – временной лаг. В уравнение включены вектор M – регрессионное среднее и вектор ε_t – регрессионный остаток, в общем различный в каждый момент времени.

Пусть состояние объекта управления описано p действительными переменными, а внешние воздействия описаны q действительными переменными. Тогда треугольная матрица $A \in R^{p \times p}$, матрица $B \in R^{p \times q}$ и векторы $Y, M, \varepsilon \in R^p, X \in R^q$.

Для прогноза состояния объекта при различных управляющих воздействиях, экзогенные переменные, элементы вектора X , разделяются на управляемые переменные U и переменные внешнего воздействия Z . Соответственно матрица регрессионных коэффициентов B разделяется на присоединенные матрицы $C | F$.

Полученная приведенная форма уравнения векторной авторегрессии

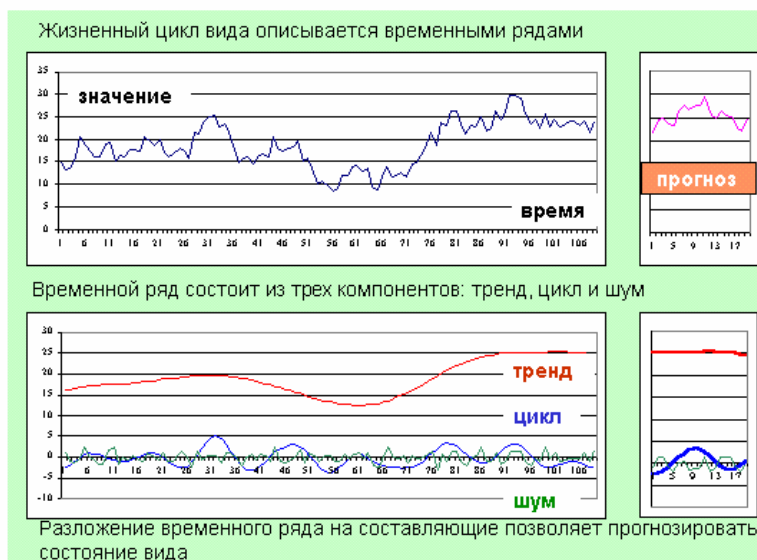
$$Y_t = (I - A_0)^{-1} \left(C_0 U_t + F_0 Z_t + \sum_{\tau=1}^r (A_\tau Y_{t-\tau} + C_\tau U_{t-\tau} + F_\tau Z_{t-\tau}) + M + \varepsilon_t \right) \text{ для получения}$$

краткосрочного прогноза состояния по заданному управлению редуцируется до выражения $Y_t = G_r U_t + H_{t,r}$.

Для решения задачи оптимального прогноза редуцированное выражение обращается $U_t = G_r^+ (Y_t - H_{t,r})$, где псевдообратная матрица G_r^+ получена с помощью сингулярного разложения.

Для анализа одномерных временных рядов используется метод "Гусеница" – SSA (singular structure analysis – анализ сингулярных структур). Базовый метод SSA состоит в преобразовании исходного ряда в многомерный, сингулярном разложении получившейся траекторной матрицы, группировке членов разложения и последующем восстановлении. При этом часто оказывается возможным выделить отдельные аддитивные составляющие исходного ряда, такие как тренд (гладкая и медленно меняющаяся часть ряда), различные колебательные и циклические компоненты, а также шумовую компоненту.

Проиллюстрируем алгоритм прогноза состояния вида следующим рисунком. Поведение, или жизненный цикл вида описывается одним или несколькими временными рядами, на рисунке в качестве такого исходного ряда показан верхний левый, где по оси абсцисс отложено время, в течение которого производились измерения, а по оси ординат – значение измеряемой переменной – состояния вида.



Прогноз состояния вида посредством анализа временных рядов

Для прогноза состояния исходный ряд разлагается на три аддитивных компоненты: тренд, циклическую составляющую и шумовую составляющую (см. нижний график). Для каждой из составляющих строится эвристическая модель поведения, согласованная с экспертами. Посредством этих моделей строится прогноз отдельных компонент (правый

нижний график) и затем спрогнозированный временной ряд восстанавливается согласно принятой ранее модели разложения (правый верхний график).

Заключение

Предлагаемые алгоритмы позволят рассчитать степень угроз для каждого вида, текущее состояние вида и обоснованно, с применением количественных оценок, подойти к процессу отнесения вида к той или иной категории редкости красной книги.

Система индексов разработана на основе экспертных балльных оценок биологического состояния видов, значимости вида для сохранения биоразнообразия. Социально-экономической значимости вида, кумулятивного фактора угроз, необходимого и достаточного условий для восстановления вида.

Использование компьютерных технологий обеспечит динамику процесса мониторинга и управления сохранением редких видов за счет повышения скорости передачи, обработки, представления и анализа информации.

В перспективе ИАС позволит создать открытую, модульную, расширяемую систему поддержки принятия решений в указанной предметной области.

Литература

1. Компьютерные технологии для решения проблем в области сохранения биоразнообразия, - Тезисы 10-й международной конференции "МАТЕМАТИКА, КОМПЬЮТЕР, ОБРАЗОВАНИЕ", Пущино, 20-25 января 2003, С. 416.
2. Кревер О.Н., Стишов М.С., Стрижов В.В., Шакин В.В. Компьютерные технологии для решения проблем, сопровождающих управление процессами сохранения биоразнообразия и мониторинг состояния видов. - Там же, с.276.
3. Смирнов В.Э. Пространственный анализ ландшафтов (На примере Приокско-Террасного заповедника). - Там же, с.305.
4. Михайлов А.В., Комаров А.С., Чертов О.Г. Моделирование стока углерода лесными экосистемами при различных сценариях лесопользования. - Там же, с.294.
5. Шакин В. В. Парето-классификация конечных выборок. Применение многомерного статистического анализа в экономике и оценке качества продукции. V-я научная конференция стран СНГ. Тезисы докладов. М.: Центральный экономико-математический институт РАН, 1993. С. 96-97.
6. Стрижов В.В., Шакин В.В. Прогноз и управление в авторегрессионных моделях. Математические методы распознавания образов. Доклады ММРО-11. М.: 2003, С.178-181.