

## **ОБРАТНАЯ ЗАДАЧА ИНТЕГРАЛЬНОГО ОЦЕНИВАНИЯ: ВЫЯВЛЕНИЕ КРИТИЧЕСКИХ СОСТАВЛЯЮЩИХ МЕДИКО- ЭКОЛОГИЧЕСКОЙ СИТУАЦИИ**

**М.Ю. Антомонов, С.Л. Пашинская**

*ДУ «Институт гигиены и медицинской экологии им. А.Н. Марзеева НАМН  
Украины»*

В работе предложен алгоритм анализа и оценки первичных показателей с целью оптимизации их перечня для расчета интегральной оценки. Технология позволяет работать с массивами, имеющими аномальные значения и пропущенные данные. Рассмотрено поэтапное конструирование интегральной оценки. Разработан нелинейный алгоритм интегрирования показателей. Предложена методика выявления критических характеристик.

**Ключевые слова:** интегральная оценка, экологические и гигиенические данные, медицинские и экологические исследования.

У роботі запропоновано алгоритм аналізу та оцінки первинних показників з метою оптимізації їх переліку для розрахунку інтегральної оцінки. Технологія дозволяє працювати з масивами, які мають аномальні значення і пропуски в даних. Розглянуто поетапне конструювання інтегральної оцінки. Розроблено нелінійний алгоритм інтегрування показників. Запропоновано методику виявлення критичних характеристик.

**Ключові слова:** інтегральна оцінка, екологічні та гігієнічні дані, медичні та екологічні дослідження.

### **ВВЕДЕНИЕ**

Реформирование системы здравоохранения, ее адаптация к социально-экономическим условиям, сложившимся в Украине, требует широкого внедрения современных медицинских, информационных технологий, новых средств компьютеризации, связи, интегрированных баз данных, широкого применения математических методов анализа и прогнозирования, которые соответствовали бы общей концепции информатизации в государстве и в области здравоохранения. Согласно действующему Постановлению Кабинета Министров Украины от 17.05.2012 № 397 одним из приоритетных направлений инновационной деятельности в Украине является внедрение принципов доказательной медицины в выполнение научных исследований и при проведении анализа базовых показателей здоровья населения. Доказательную медицину определяют как новейшую технологию сбора, анализа, синтеза и применения научной медицинской информации, которая позволяет принимать оптимальные клинические решения. Эколого-гигиенические характеристики объектов окружающей среды, социально-экономическая ситуация регионов, популяционные показатели здоровья населения, уровень индивидуального здоровья необходимо исследовать как единую систему, сформированную на основе интеграции процессов сбора, обработки, анализа и распространения всей статистической информации.

Объем этой информации, доступ к которой возможен с помощью современных интернет-технологий, может быть чрезвычайно велик.

В частности, такие огромные массивы данных содержатся в базах: «Здоровье для всех» (ВООЗ), «Экологический мониторинг», «ДунаMed», «Федеральный информационный фонд данных социально-гигиенического мониторинга» (Российская Федерация) и другие. Количество показателей, которые характеризуют экологические и медицинские системы, может быть от нескольких сотен до нескольких тысяч. При анализе таких объемов информации необходимо применять различные математические приемы свертки и агрегации данных. Одним из эффективных путей снижения размерности массивов данных является использование комплексных показателей (КП) и интегральных оценок (ИО).

**Целью** нашей работы является разработка методики, алгоритмов, расчетных формул и программной реализации для решения обратной задачи интегрального оценивания — выявления дестабилизирующих факторов в уже имеющейся оценке медико-экологической ситуации.

Реализация поставленной цели предполагает постановку и решение поэтапных задач, отражающих логическую структуру и последовательность исследования.

### **ПОСТАНОВКА ЗАДАЧИ**

Для комплексного оценивания качества среды в медико-экологических исследованиях применяются разнообразные математические конструкции, чаще всего представляющие собой средневзвешенные суммы значений выраженности всех регистрируемых вредных факторов. Одним из подходов построения КП является предложенная нами информационная технология интегрального оценивания. Эта технология была апробирована на разнообразных массивах медицинской, экологической и гигиенической информации, подготовлены соответствующие программные модули и выполнена верификация подхода. Таким образом, конечному пользователю в автоматизированном режиме могут быть представлены массивы КП с их интегральной оценкой.

Однако, при содержательном анализе уже сформированных комплексных показателей естественным образом возникает обратная задача: за счет каких составляющих (факторов среды) эти показатели приобретают экстремальные (наиболее опасные) значения и как выявить эти переменные по уже имеющимся КП.

### **ФОРМИРОВАНИЕ ИНТЕГРАЛЬНЫХ ОЦЕНОК ПО МЕДИКО-ЭКОЛОГИЧЕСКИМ ДАННЫМ**

Этап начинается с экспресс-обработки массива данных. Обработка массива начинается с дескриптивного анализа (расчет параметрических характеристик). Следующий шаг — логический анализ, позволяющий сделать определенные предположения о данном массиве и, при необходимости, его скорректировать.

Проверка на соответствие экспериментального распределения нормальному распределению позволяет определить методы обработки

(параметрические или непараметрические). Существуют простые методы проверки на принадлежность к нормальному распределению и сложные критерии согласия. Целесообразно для большей достоверности выполнять проверку несколькими методами [1]. Однако, наиболее приемлемым, на наш взгляд, для этих целей является критерий Колмогорова. В отличие от чаще применяемого критерия согласия Пирсона, использовать который рекомендуется на достаточно больших выборках, критерий Колмогорова менее чувствителен к объему выборки. Он не требует использования таблиц, а потому более простой в программной реализации. Дополнительно рекомендуется проводить проверку по числам Вестергарда.

В дальнейшем осуществляется проверка исходного массива на наличие выбросов и их анализ. Если аномальные значения являются грубыми ошибками, их необходимо выбраковать. Если эти значения являются так называемыми «мягкими выбросами», они могут быть информативными и сигнализировать о реальных отклонениях в экологической ситуации конкретных населенных мест. При этом критерием определения типа выбросов является величина заранее выбранного сигмального отклонения. Для выявления аномальных значений можно использовать различные методы в зависимости от объема выборки и вида распределения. Экспериментальная проверка показала, что целесообразнее на малых и средних выборках использовать метод Граббса. На больших объемах рекомендуется применять метод Райта в сочетании с непараметрическим методом по определению межквартильного интервала.

Также данные проверяются на комплектность. При обработке массивов с неполными данными существует два подхода — удаление некомплектных фрагментов и восстановление пропущенных данных [2, 3]. Однако при восстановлении данных уменьшается точность расчетов, при удалении теряется репрезентативность выборки. Удаление целесообразно тогда, когда пустоты в таблице составляют значительный процент (определяется пользователем, в зависимости от требований поставленной задачи) от объема выборки, например, 20%. Самый простой (хотя и не очень точный) метод, который чаще всего используется в статистических пакетах, — заполнение пропусков средними значениями величин, находящихся в данном столбце. В нашей работе использована линейная интерполяция. При этом строились линейные регрессионные модели для смежных столбцов с полностью заполненными данными (аргумент функции), по которым рассчитывались недостающие элементы (значения функции).

Следующий шаг — отбор информативных характеристик. Эта процедура может выполняться как с помощью формальных математических методов (например, в рамках факторного анализа), так и с помощью экспертного оценивания.

На этом этапе также рассчитываются весовые коэффициенты при переменных, которые определяются их значимостью [4, 5].

Одним из средств уменьшения размерности массива является формирование индексов и обобщенных показателей, которые представляют собой объединение по определенным правилам множества исходных переменных в одну характеристику. Информативность, точность и

достоверность комплексного показателя зависит от правильного выбора метода обработки данных в зависимости от общей постановки задачи. Комплексные показатели, в свою очередь, могут быть использованы для интегральной оценки объекта исследования.

Процесс конструирования ИО предполагает выполнение следующих действий: 1) выбор или расчет «нормы» показателей; 2) расчет безразмерных эквивалентов; 3) нормирование показателей или их эквивалентов; 4) формирование интегральных оценок [6].

1. В качестве «нормы» можно использовать статистические характеристики имеющихся эмпирических выборок: средние, минимальные и максимальные значения и т. д. При этом следует учитывать относительный характер этих значений, поскольку они изменяются при изменении исходного массива данных. В качестве «нормы» могут использоваться некоторые «внешние» характеристики: фоновые или критические значения показателей, гигиенически допустимые уровни факторов, справочные значения и т. д. [7].

2. Безразмерные эквиваленты получаются путем деления исходных значений массива на соответствующую «норму» или показатели вариабельности, измеренные в тех же единицах, что и сами переменные.

3. Нормирование данных применяется для того, чтобы все значения признаков попадали в удобный для сравнения диапазон. Обычно это интервал  $[0, 1]$ . При этом надо сразу определить смысловую направленность: значение показателей тем «больше», чем они «лучше» или наоборот. Естественно, одной и той же логике должно подчиняться преобразование для всех переменных.

В зависимости от особенностей характера изменения показателя при разных способах нормировки можно получить максимальные значения эквивалентов (например, равные единице) для средних («наиболее нормальных»), минимальных («наиболее плохих») или, наоборот, максимальных («наиболее хороших») значений исходных переменных.

Нами для нормирования была выбрана экспоненциальная функция «полуколокол» [1], которая записывалась в следующем виде:

$$\tilde{x} = \exp(-a(x - x^*)^2), \quad (1)$$

где  $x$  — значение исходной переменной,  $a$  — параметр, определяющий форму «колокола» (его «ширину»),  $x^*$  — стандарт, который мы принимаем как минимальное, максимальное или эталонное значение в зависимости от того, какие характеристики для системы оптимальные.

По этой формуле нормированные эквиваленты «плохих» значений приближаются к нулю, но не достигают его. «Наилучшие» значения равны единице.

4. Конкретный вид функции свертки нормированных показателей состояния исследуемого объекта при расчете интегральной оценки зависит от вида решаемой задачи и требований, предъявляемых к результату. Наиболее удобный и часто употребляемый способ интегрирования показателей — аддитивная линейная свертка с учетом весовых коэффициентов. Иногда при

расчете интегральной оценки обращаются к методу геометрической средней, методу коэффициентов, методу суммы мест, методу расстояний. Их целесообразно применять лишь при относительно малом числе оцениваемых показателей и в случае, если большинство их значений близки к единице.

Линейная свертка, однако, корректна только тогда, когда все критерии попарно независимы по предпочтению, проверка чего на практике почти невозможна при большом наборе критериев [8]. В противном случае в интегральном оценивании возможна компенсация низких значений показателей высокими.

## **ВЫЯВЛЕНИЕ КРИТИЧЕСКИХ СОСТАВЛЯЮЩИХ ПО ИНТЕГРАЛЬНЫМ ОЦЕНКАМ**

Часто в процессе медико-экологического анализа возникает задача определения таких элементов, которые значительно ухудшают состояние объекта, но не выявляются при интегральном оценивании, назовем эти элементы слабым звеном интегральной оценки. Предлагается сформировать корректирующую интегральную оценку, значение которой будет смещаться в сторону слабого звена. Имея рассчитанные интегральные оценки — основную и корректирующую, мы можем решить обратную задачу интегрального оценивания.

Таким образом, под критическими значениями понимаются такие, из-за которых интегральные оценки принимают «наихудшие» значения.

Алгоритм реализуется следующим образом:

1. Расчет средневзвешенных оценок исследуемых объектов по нормированным эквивалентам:

$$I_1 = \frac{\sum_{i=1}^n w_i \tilde{x}_i}{\sum_{i=1}^n w_i}, \quad (2)$$

где  $I_1$  — интегральная средневзвешенная оценка,  $\tilde{x}_i$  — нормированные эквиваленты,  $w_i$  — весовые коэффициенты.

2. Расчет корректирующих интегральных оценок. Этот шаг был реализован тремя вариантами.

В первом варианте рассчитывались «классические» смещенные интегральные оценки по формуле средневзвешенной арифметической:

$$I_{2(1)} = \frac{\sum_{i=1}^n w_i (\tilde{x}_i - \tilde{x}_{\min})}{\sum_{i=1}^n w_i}, \quad (3)$$

где  $I_{2(1)}$  — смещенная интегральная средневзвешенная оценка,  $\tilde{x}_i$  — нормированные эквиваленты,  $\tilde{x}_{\min}$  — минимальное значение эквивалента,  $w_i$  — весовые коэффициенты.

Во втором варианте рассчитывались интегральные оценки по формуле средневзвешенной геометрической:

$$I_{2(2)} = \left( \prod_{i=1}^n \tilde{x}_i^{w_i} \right)^{\frac{1}{\sum_{i=1}^n w_i}}, \quad (4)$$

где  $I_{2(2)}$  — средневзвешенная геометрическая оценка,  $\tilde{x}_i$  — нормированные эквиваленты,  $w_i$  — весовые коэффициенты.

Третий вариант — расчет смещенных интегральных оценок как корня из произведения минимального эквивалента и средневзвешенной суммы:

$$I_{2(3)} = \sqrt{\frac{\sum_{i=1}^n w_i \tilde{x}_i}{\sum_{i=1}^n w_i} \cdot \tilde{x}_{\min}}, \quad (5)$$

где  $I_{2(3)}$  — смещенная интегральная оценка,  $\tilde{x}_{\min}$  — минимальный эквивалент, соответствующий худшему показателю.

3. Сравнение средневзвешенной и смещенной интегральных оценок. Если набор показателей, характеризующих объект, относительно однородный, различие между оценками  $I_1$  и  $I_2$  будет незначительным. Если же различие существенно, то показатель, эквивалент которого минимальный, является «слабым звеном» среди всех характеристик объекта. Если же хотя бы один из эквивалентов частных показателей равен нулю, то, очевидно, будет равной нулю интегральная оценка  $I_2$  во втором и третьем вариантах. Состояние объекта в этом случае будет считаться наихудшим. Аналогично агрегируя показатели по группе исследуемых объектов, получаем общую интегральную оценку. Полученное значение переводим в балльную шкалу и присваиваем соответствующим уровням шкалы вербальную оценку с градациями: «нормальный», «удовлетворительное», «плохой», «очень плохой» и «критический» (соответствует значению «0»). Осуществляя итерацию и изымая каждый раз очередной минимальный эквивалент, обнаружим все дестабилизирующие факторы объекта.

## ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ПРЕДЛОЖЕННОГО АЛГОРИТМА

Программная реализация «Экспресс-анализ» является объединением отдельных модулей, предназначенных для оперативной предварительной обработки таблиц первичных данных с последующим конструированием интегральной оценки, и разработана в среде MS Excel средствами встроенного языка макросов Visual Basic for Application (VBA).

С помощью предложенной программной реализации был обработан массив данных Госсанэпидемслужбы Ровенской области, который содержит медицинские и демографические показатели, а также санитарно-гигиенических показатели состояния систем водоснабжения атмосферного воздуха области. Рассмотрим соотношения построенных нами различных вариантов интегральных оценок (таблица 1).

**Таблица 1.**

*Сравнительная таблица интегральных оценок медико-экологического состояния районов Ровенской области*

№	Название района	Интегральные оценки									
		$I_1$	$I_{2(1)}$	$I_{2(2)}$	$I_{2(3)}$	$I_1 - I_{2(1)}$	$I_1 / I_{2(1)}$	$I_1 - I_{2(2)}$	$I_1 / I_{2(2)}$	$I_1 - I_{2(3)}$	$I_1 / I_{2(3)}$
1	Березновский	0,82	0,46	0,80	0,54	0,36	1,77	0,02	1,03	0,28	1,52
2	Владимирецкий	0,85	0,59	0,82	0,48	0,27	1,45	0,03	1,03	0,38	1,79
3	Гошанский	0,77	0,57	0,73	0,39	0,20	1,34	0,04	1,06	0,38	1,98
4	Демидовский	0,84	0,46	0,81	0,57	0,38	1,83	0,03	1,04	0,27	1,48
5	Дубенский	0,88	0,38	0,86	0,66	0,50	2,33	0,02	1,02	0,21	1,32
6	Дубровицкий	0,88	0,56	0,85	0,54	0,32	1,58	0,03	1,04	0,35	1,65
7	Заречненский	0,69	0,50	0,63	0,37	0,20	1,39	0,06	1,10	0,33	1,88
8	Здолбуновский	0,71	0,42	0,66	0,45	0,29	1,68	0,04	1,06	0,26	1,57
9	<b>Корецкий</b>	0,68	0,59	0,55	0,26	<b>0,10</b>	<b>1,16</b>	<b>0,13</b>	<b>1,24</b>	<b>0,43</b>	<b>2,67</b>
10	Костопольский	0,76	0,54	0,70	0,41	0,22	1,40	0,06	1,09	0,35	1,86
11	Млиновский	0,83	0,45	0,79	0,55	0,37	1,82	0,03	1,04	0,27	1,49
12	Острожский	0,81	0,58	0,73	0,42	0,22	1,38	0,08	1,11	0,38	1,90
13	<b>Радивилловский</b>	0,80	0,64	0,71	0,36	<b>0,16</b>	<b>1,25</b>	<b>0,09</b>	<b>1,12</b>	<b>0,44</b>	<b>2,23</b>
14	Ровенский	0,92	0,56	0,91	0,58	0,37	1,65	0,02	1,02	0,34	1,59
15	Рокитновский	0,77	0,54	0,70	0,42	0,23	1,43	0,07	1,10	0,35	1,83
16	Сарненский	0,85	0,40	0,84	0,62	0,45	2,14	0,01	1,01	0,23	1,37

По результатам сравнения оценок видно, что наименьшие различия средневзвешенной оценки ( $I_1$ ) и «классической» смещенной ( $I_{2(1)}$ ), и напротив, наибольшие различия средневзвешенной оценки с двумя другими вариантами  $I_{2(2)}$  и  $I_{2(3)}$  указывают на одни и те же районы — Корецкий и Радивилловский. В этих районах предположительно и находятся потенциально опасные факторы, которые могут значительно ухудшать состояние исследуемых объектов, но при этом не выявляться в средневзвешенной оценке.

Недостаток первого варианта — «классическая» смещенная оценка не реагирует на наличие нулевого эквивалента, который указывает на критическое состояние объекта. Средневзвешенная геометрическая оценка более информативна, но создает определенные сложности в вычислениях. Третий вариант, комбинированная свертка, проще в расчетах и содержит в своей конструкции минимальный элемент, по которому мы определяем потенциально опасный фактор, и таким образом решаем обратную задачу.

Рассмотрим подробнее процедуру выявления проблемных районов Ровенской области на примере заболеваемости. Проанализируем соотношение интегральных оценок — средневзвешенной и смещенной комбинированной (таблица 2).

**Таблица 2.**

*Оценки районов Ровенской области по заболеваемости населения*

№	Название района	Статистические характеристики					
		$I_1$	Ранг1	$I_{2(3)}$	Ранг2	Состояние	$I_1 / I_{2(3)}$
1	Березновский	0,57	15	0,26	15	плохое	<b>2,19</b>
2	Владимирецкий	0,79	3	0,62	1	удовлетворительное	1,27
3	Гошанский	0,63	11	0,41	11	плохое	1,54
4	Демидовский	0,68	9	0,49	4	плохое	1,39
5	Дубенский	0,81	1	0,42	9	плохое	1,93
6	Дубровицкий	0,67	10	0,47	5	плохое	1,43
7	Заречненский	0,59	13	0,25	16	плохое	<b>2,36</b>
8	Здолбуновский	0,73	8	0,41	10	плохое	1,78
9	Корецкий	0,54	16	0,28	14	плохое	1,93
10	Костопольский	0,74	6	0,47	6	плохое	1,57
11	Млиновский	0,79	2	0,54	2	удовлетворительное	1,46
12	Острожский	0,74	7	0,46	8	плохое	1,61
13	Радивиловский	0,74	5	0,46	7	плохое	1,61
14	Ровенский	0,61	12	0,40	12	плохое	1,53
15	Рокитновский	0,76	4	0,50	3	удовлетворительное	1,52
16	Сарненский	0,58	14	0,31	13	плохое	1,87

В Березновском и Заречненском районах значение  $ИО_1$  более чем в два раза превышает значение  $ИО_2$ . Это свидетельствует о том, что помимо общего плохого состояния, некоторые заболевания вызывают особое опасение. Какие именно показатели настораживают, можно определить по таблице эквивалентов: в проблемных районах это будут те показатели, которым соответствуют наименьшие эквиваленты.

В таблице 3 показан фрагмент массива показателей по следующим заболеваниям: инфекционные и паразитарные болезни (ИП), новообразования (Н), болезни крови и кроветворных органов (ККО), болезни эндокринной системы (ЭС), расстройства психики и поведения (РПП), болезни нервной системы (НС), болезни глаза и придаточного аппарата (ГПА), болезни уха и сосцевидного отростка (УСО), болезни системы кровообращения (СК), болезни органов дыхания (ОД), болезни органов пищеварения (ОП), болезни кожи и подкожной клетчатки (КПК).

Анализ нормированных эквивалентов показателей заболеваемости (таблица 3) и определение их минимальных значений на проблемных территориях указывает на то, что особое беспокойство вызывают новообразования (Н) в Дубенском районе, болезни крови и кроветворных органов (ККО) в Березновском и эндокринной системы (ЭС) в Заречненском районах.



Таблица 3.

Выявление опасных уровней заболеваемости населения районов Ровенской области

№	Название района	Показатели заболеваемости											
		ИП	Н	ККО	ЭС	РПП	НС	ГПА	УСО	СК	ОД	ОП	КПК
1	Березновский	0,69	0,86	<b>0,12</b>	0,36	0,88	0,92	0,47	0,76	0,51	0,88	0,16	0,56
2	Владимирецкий	0,91	0,73	0,68	0,75	1,00	0,91	0,62	0,84	1,00	1,00	0,63	0,75
3	Гошанский	0,45	0,40	0,83	0,77	0,52	0,43	0,96	0,80	0,33	0,35	0,32	0,78
4	Демидовский	0,96	0,68	0,76	0,59	0,40	0,83	0,36	0,88	0,96	0,70	0,37	0,99
5	Дубенский	0,51	<b>0,22</b>	1,00	1,00	0,49	0,99	1,00	1,00	0,90	1,00	1,00	1,00
6	Дубровицкий	0,82	0,42	0,87	0,60	0,88	1,00	0,42	0,60	0,43	0,86	0,73	0,99
7	Заречненский	0,13	0,95	0,55	<b>0,11</b>	0,26	0,32	0,68	1,00	0,30	0,99	0,60	0,44
8	Здолбуновский	0,74	0,24	0,34	0,77	0,34	0,97	0,97	0,95	0,55	0,94	0,93	0,88
9	Корецкий	0,80	0,66	0,69	0,23	0,23	0,84	0,39	0,21	0,65	0,59	0,52	0,95
10	Костопольский	0,94	0,81	0,95	0,75	0,50	1,00	0,45	0,55	0,56	1,00	0,68	0,95
11	Млиновский	0,38	0,87	0,97	0,91	0,72	0,99	0,81	0,89	0,96	0,61	0,73	0,72
12	Острожский	0,96	0,73	0,81	0,84	0,49	0,91	0,29	0,92	0,66	0,61	0,52	0,61
13	Радивилловский	0,58	0,76	0,96	0,75	0,30	0,88	0,42	0,42	0,97	0,91	0,73	0,66
14	Ровенский	1,00	0,98	0,61	0,59	0,61	0,56	0,55	0,33	0,74	0,27	0,31	0,46
15	Рокитновский	0,82	1,00	0,84	0,50	0,38	0,37	0,96	0,82	1,00	0,64	0,67	1,00
16	Сарненский	0,69	0,88	0,78	0,72	0,17	0,55	0,43	0,63	0,68	0,77	0,20	0,24

Аналогично проведем анализ уровня загрязненности атмосферного воздуха в Ровенской области (таблица 4). Из анализа исключены Березновский, Демидовский, Острожский и Радивилловский районы из-за недостатка данных.

Таблица 4.

Анализ состояния атмосферного воздуха

№	Название района	Статистические характеристики					
		$I_1$	Ранг1	$I_{2(3)}$	Ранг2	Состояние	$I_1/I_{2(3)}$
1	Березновский	–	–	–	–	–	–
2	Владимирецкий	0,77	9	0,52	9	удовлетворительное	1,48
3	Гошанский	0,90	6	0,69	6	удовлетворительное	1,30
4	Демидовский	–	–	–	–	–	–
5	Дубенский	0,80	8	0,52	8	удовлетворительное	1,54
6	Дубровицкий	0,92	5	0,81	5	нормальное	1,14
7	Заречненский	0,94	4	0,85	4	нормальное	1,11
8	Здолбуновский	0,70	10	0,47	10	плохое	1,49
9	Корецкий	0,95	3	0,85	3	нормальное	1,12
10	Костопольский	0,59	12	0,00	12	<b>критическое!</b>	–
11	Млиновский	0,60	11	0,00	12	<b>критическое!</b>	–
12	Острожский	–	–	–	–	–	–
13	Радивилловский	–	–	–	–	–	–
14	Ровенский	0,98	1	0,92	1	нормальное	1,07
15	Рокитновский	0,96	2	0,86	2	нормальное	1,12
16	Сарненский	0,87	7	0,63	7	удовлетворительное	1,38

Из таблицы видно, что ранги по двум интегральным оценкам одинаковы и значения  $ИО_1$  и  $ИО_2$  существенно не отличаются, за исключением Костопольского и Млиновского районов, состояние которых определено как критическое. Критическое состояние Костопольского района вызвано содержанием в воздухе диоксида азота и фенола в концентрациях,

превышающих предельную допустимую. В Млыновском районе недопустимая концентрация формальдегида и сажи.

Результат обработки позволил ранжировать районы Ровенской области, исходя из значений интегральных оценок, определить районы с наиболее неблагоприятной экологической обстановкой и уровнем заболеваемости, выявить дестабилизирующие факторы в данных районах.

## **Выводы**

Предложенный алгоритм позволяет оперативно осуществить обработку массива данных, привести его к формату, удобному для дальнейшего более детального анализа и формирования интегральной оценки.

По результатам выполнения первой задачи (определение интегральных оценок) можно провести ранжирование эколого-гигиенических объектов, исходя из интегральных оценок; определить объекты с наиболее неблагоприятной экологической обстановкой и уровнем заболеваемости.

Решение второй задачи (расчет смещенных интегральных показателей) позволяет: выявить критические элементы в системе показателей; сдвинуть средневзвешенную оценку в сторону слабого звена системы; трансформировать результаты в бальную шкалу и сформировать вербальную оценку.

1. Лемешко Б.Ю. Исследование особенностей и мощности некоторых критериев нормальности / Б.Ю. Лемешко, А.П. Рогожников // *Метрология*. — 2009. — № 4. — С. 3–24.
2. Васильев В.И. Восстановление пропущенных данных в эмпирических таблицах / В.И. Васильев, А.И. Шевченко // *Искусственный интеллект*. — 2003. — № 3. — С. 317–324.
3. Злоба Е. Статистические методы восстановления пропущенных данных / Е. Злоба, И. Яцкив // *Computer Modeling & New Technologies*. — 2002. — Vol. 6. — № 1. — С. 51–61.
4. Бакуменко Л.П. Интегральная оценка качества и степени экологической устойчивости окружающей среды региона (на примере Республики Марий Эл) / Л.П. Бакуменко, П.А. Коротков // *Прикладная эконометрика*. — 2008. — № 1(9). — С. 73–92.
5. Павлов С.Б. Экологічний ризик для здоров'я населення / С.Б. Павлов // *Медицинские исследования*. — 2001. — Т. 1, вып. 1. — С. 16–19.
6. Антомонов М.Ю. Математическая обработка и анализ медико-биологических данных / М.Ю. Антомонов. — Киев, 2006. — 558 с.
7. Шуйский В.Ф. Количественная оценка и нормирование сложных антропогенных воздействий на макрозообентос / В.Ф. Шуйский, Т.П. Занцинская, Д.С. Петров // *Сб. науч. тр. ГосНИОРХ*. — Вып. 326. — 2000. — С. 137–144.
8. Файнзильберг Л.С. Правдоподобные, но неверные решения при построении диагностических правил / Л.С. Файнзильберг // *Материалы восьмой дистанционной научно-практической конференции с международным участием «Системы поддержки принятия решений. Теория и практика. СППР'2012»*. — Киев: ИПММС НАН Украины, 2012. — С. 31–34.

## **INVERSE PROBLEMS OF INTEGRATED EVALUATION: IDENTIFICATION OF CRITICAL COMPONENTS OF HEALTH AND ECOLOGICAL SITUATION**

**Pashinskaia S.L., Antomonov M.Yu.**

*State Institution «O.M.Marzeyev Institute for Hygiene and Medical Ecology of National Academy of Medical Sciences of Ukraine»*

**Introduction.** Integrated assessment of environmental quality in medical ecological research used different mathematical structures that are often weighted sum of expression of all reported hazards. The problem of optimal formation of integrated indicators is the direct problem of integral evaluation

**The purpose** of work is to develop methods, algorithms, computational formulas and software implementation for solving the inverse problem of the integral evaluation — identifying destabilizing factors in the assessment of ecological and hygienic objects.

**Results.** Realization of this aim requires the formulation and solution of problems phased: express processing of data array; the selection of informative features; construction of an integrated evaluation.

Algorithm of construction integrated evaluation is implemented as follows: the calculation of the normalized equivalents of selected indicators; calculation of weighted average grade of the objects on the normalized equivalents; calculation shifted integrated evaluation, as the square root of the product of the minimum and weighted average; comparison of average and shifted integrated evaluation; identification of critical elements.

The paper presents an automated technology analysis and evaluation of the primary indicators in order to optimize their list to calculate the integral evaluation. The technology allows to work with arrays having outliers and missing data. We consider a phased construction of an integrated assessment. The nonlinear algorithm of integration indicator formation and the method for identifying critical elements were developed.

**Conclusions.** The proposed technology allows to quickly implement the processing of the data array, to bring it to a format suitable for further, more detailed analysis and to form an integrated assessment. The results of processing may be performed within the environmental and medical objects; identify objects with the most adverse environmental conditions and disease. Calculation of displaced integral indicators shows destabilizing elements in the system of indicators.

**Keywords:** integrated evaluation, ecological and hygienic objects, medical and ecological research.

1. Lemeshko B.Yu., Rogozhnikov A.P. *Investigation of the features and power of some of the criteria of normality*. Metrologiya, 2009, no 4, pp. 3–24 (in Russian).
2. Vasil'yev V.I., Shevchenko A.I. *Recovery of missing data in empirical tables*. Iskusstvennyy intellect, 2003, no 3, pp. 317–324 (in Russian).

3. Zloba Ye., Yatskiv I. *Statistical methods for recovery of missing data*. Computer Modeling & New Technologies, 2002. Vol. 6. no 1. pp. 51–61 (in Russian).
4. Bakumenko L.P., Korotkov P.A. Integral assessment of the quality and environmental sustainability of the region (on the example of the Republic of Mari El). *Prikladnaya ekonometrika*. 2008, no 1(9), pp. 73–92 (in Russian).
5. Pavlov S. B. *Ekologichny rizik for Health Protection of the population*. Meditsinskiye issledovaniya, 2001. T. 1, vyp. 1, pp. 16–19 (in Ukrainian).
6. Antomonov M.Yu. *Mathematical processing and analysis of medical and biological data*, Kiyev, 2006. 558 p (in Russian).
7. Shuyskiy V.F., Zantsinskaya T.P., Petrov D.S. *Quantification and valuation of complex anthropogenic impact on macrozoobenthos*. Sb. nauch. tr. GosNIORKH. Vyp. 326. 2000, pp. 137–144 (in Russian).
8. Faynzilberg L.S. *Plausible, but wrong decisions in the construction of diagnostic rules*. Materialy vosmoy distantsionnoy nauchno-prakticheskoy konferentsii s mezhdunarodnym uchastiyem. «Sistemy podderzhki prinyatiya resheniy. Teoriya i praktika. SPPR '2012». Kiyev: IPMMS NAN Ukrainy, 2012, pp. 31–34 (in Russian).

Получено 07.11.2014