

## МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ ВЛИЯНИЯ ГЛОБАЛЬНЫХ УГРОЗ НА БЕЗОПАСНОСТЬ СТРАН МИРА

**Ключевые слова:** глобальные угрозы, устойчивое развитие, норма Минковского, корреляционный анализ, кластерный анализ, анализ главных компонент.

### ВВЕДЕНИЕ

Настоящая статья является развитием исследований, представленных в [1, 2], в которых теоретически обосновано влияние системных мировых конфликтов на устойчивое развитие стран и регионов мира в глобальном контексте. На основе анализа данных о мировых конфликтах, которые произошли с 705 г. до нашей эры по настоящее время, в [1] выведена закономерность их течения. Показано, что последовательность жизненных циклов системных мировых конфликтов подчинена закону изменения элементов ряда Фибоначчи, а интенсивность этих конфликтов, которая зависит от уровня технологического развития общества, нарастает по гиперболическому закону. С использованием обнаруженных закономерностей сделана попытка предвидения последующего мирового конфликта, названного конфликтом XXI столетия, и выполнен анализ его природы и главных характеристик: длительности, основных фаз течения и интенсивности.

В работе [2] в количественных величинах представлена совокупность основных десяти глобальных угроз, которые порождают конфликт XXI столетия, и выполнено моделирование их влияния на безопасность стран и регионов мира. Этими угрозами являются *ES* — энергетическая безопасность; *BB* — баланс между совокупным потреблением и биологической продуктивностью Земли; *GINI* — неравенство доходов; *GD* — глобальные болезни; *CM* — детская смертность; *CP* — индекс восприятия коррупции; *WA* — доступ к питьевой воде; *GW* — глобальное потепление; *SF* — степень несостоятельности государства; *ND* — уязвимость к природным катастрофам.

В связи с тем, что оценивать безопасность той или иной страны одновременно в контексте десяти глобальных угроз достаточно сложно, в настоящей работе для более удобного и наглядного исследования этой проблемы воспользуемся методами многомерного статистического анализа (МСА). С помощью этих методов проведем дополнительные исследования корреляционных взаимосвязей между отдельными угрозами, сведем переменные со многими свойствами к нескольким скрытым (латентным) факторам, определяющим эти свойства, и выполним кластеризацию стран по степени влияния глобальных угроз на их безопасность.

### ПОСТАНОВКА ЗАДАЧИ

Воспользуемся данными работы [2, табл. 1] относительно десяти угроз и оценками их влияния на безопасность стран. Представим их в виде матрицы типа объект — свойство

$$X = \langle X^1, X^2, \dots, X^m \rangle, \quad (1)$$

где  $X^i \in \{ES, BB, GINI, GD, CM, CP, WA, GW, SF, ND\}$  — вектор-столбец значений  $i$ -й глобальной угрозы. В работе [1] исходные данные для каждой угрозы нормированы так, чтобы их значения изменялись в диапазоне (0–1). Например, для глобальной угрозы *ES* имеем

$$ES_{\text{norm}} = 1 - \frac{ES - ES_{\min}}{ES_{\max} - ES_{\min}}.$$

После выполнения нормирования для всех глобальных угроз получен нормированный вектор угроз

$$T_{j \text{ norm}} = (ES_{j \text{ norm}}, BB_{j \text{ norm}}, GINI_{j \text{ norm}}, GD_{j \text{ norm}}, CM_{j \text{ norm}}, CP_{j \text{ norm}}, WA_{j \text{ norm}}, GW_{j \text{ norm}}, SF_{j \text{ norm}}, ND_{j \text{ norm}}).$$

© Т.Н. Померанцева, А.А. Болдак, 2010

При этом значению 0 будет отвечать максимальная угроза, значению 1 — минимальная. Иными словами, после выполнения нормирования каждая угроза становится тем более «приближенной» к конкретной стране, чем ее числовое значение ближе к нулю, и тем более «отдаленной» от этой страны, чем ее значение ближе к единице.

Для каждой страны рассчитывается значение индекса безопасности  $I_{sec}$ , которое является нормой Минковского для вектора угроз  $T_{j\text{norm}}$ :

$$I_{sec} = \| T_{j\text{norm}} \| = \sqrt[3]{\sum_{l=1}^{10} (T_{jl\text{norm}})^3}.$$

Таким образом, безопасность той или иной страны будем количественно отображать с помощью индекса безопасности  $I_{sec}$ , представляющего собой интегрированную оценку влияния совокупности глобальных угроз в пространстве Минковского.

Существует, по крайней мере, три категории причин, которые обуславливают применение методов МСА к анализу данных, представленных табл. 1 и соотношением (1).

Во-первых, глобальные угрозы в силу тех или иных причин являются взаимосвязанными. Незнание характера и степени этих связей может существенно исказить представление о реальной безопасности стран. Задача исследования характеристик взаимосвязей необходимо решать с помощью методов корреляционного анализа.

Во-вторых, для удобства и наглядности исследования возникает необходимость свести переменные со многими свойствами к нескольким скрытым (латентным) факторам, определяющим эти свойства. В данном случае безопасность страны может быть упрощенно представлена не всеми десятью глобальными угрозами, а несколькими наиболее существенными факторами. Для этих целей используется метод главных компонент.

В-третьих, необходимо классифицировать страны по степени влияния глобальных угроз на их безопасность, что и обуславливает применение методов кластерного анализа.

#### ИСПОЛЬЗОВАНИЕ МСА ДЛЯ ИЗУЧЕНИЯ БЕЗОПАСНОСТИ СТРАН И РЕГИОНОВ МИРА

**Анализ статистической связи глобальных угроз.** Установление самого факта наличия (или отсутствия) статистически значимой связи между  $X^i$  и  $X^j$ ,  $i, j = \overline{1, m}$ ,  $i \neq j$ , является важной задачей. Это связано с тем, что эффект мультиколлинеарности [3] показателей  $X^i$  создает трудности в содержательной интерпретации параметров модели расчета безопасности стран [1], особенно, когда речь идет о выявлении причинно-следственных связей.

Поскольку в [1, 2] все глобальные угрозы выражены в интервальных шкалах, то степень линейной связи между отдельными угрозами можно оценить с помощью коэффициента корреляции Пирсона [4]

$$r_{k,l} = \frac{\sum_{i=1}^m (x_i^k - \bar{X}^k) \cdot (x_i^l - \bar{X}^l)}{\sqrt{\sum_{i=1}^N (x_i^k - \bar{X}^k)^2} \cdot \sqrt{\sum_{i=1}^N (x_i^l - \bar{X}^l)^2}},$$

где  $\bar{X}^l = \frac{\sum_{i=1}^N x_i^l}{N}$ ;  $\bar{X}^k = \frac{\sum_{i=1}^N x_i^k}{N}$  — средние значения параметров  $X^l$  и  $X^k$ .

Результаты вычисления парных коэффициентов корреляции для анализируемых глобальных угроз приведены в табл. 2.

Из [4, 5] следует, что коэффициенты корреляции, которые по модулю больше 0.7, свидетельствуют о сильной связи (коэффициенты детерминации больше 50 %, т.е. один признак определяет другой более чем наполовину). Коэффициенты корреляции, которые по модулю меньше 0.7, но больше 0.5, говорят о связи средней силы (коэффициенты детерминации меньше 50 %, но больше 25 %). Наконец, коэффициенты корреляции, которые по модулю меньше 0.5, говорят о слабой связи (коэффициенты детерминации меньше 25%).

**Таблица 1.** Показатели безопасности жизни в 2009 году

Страна	Рей- тинг	<i>Isec</i>	<i>ES</i>	<i>BB</i>	<i>GINI</i>	<i>GD</i>	<i>CM</i>	<i>CP</i>	<i>WA</i>	<i>GW</i>	<i>SF</i>	<i>ND</i>
1	2	3	4	5	6	7	8	9	10	11	12	13
<b>Очень высокий уровень безопасности</b>												
Канада	1	1.846	0.81	0.871	0.674	0.762	0.955	0.87	1	0.427	0.904	1
Норвегия	2	1.83	0.909	0.401	0.742	0.74	0.973	0.79	1	0.694	0.908	1
Швеция	3	1.828	0.116	0.595	0.75	0.762	0.973	0.93	1	0.792	0.93	1
Австралия	4	1.805	0.8	0.687	0.648	0.775	0.955	0.87	1	0.417	0.896	0.996
Дания	5	1.794	0.107	0.349	0.753	0.699	0.973	0.93	1	0.675	0.974	1
Финляндия	6	1.792	0.091	0.649	0.731	0.73	0.973	0.9	1	0.54	0.962	1
Новая Зеландия	7	1.792	0.146	0.646	0.638	0.742	0.955	0.93	1	0.728	0.908	1
Швейцария	8	1.768	0.068	0.302	0.663	0.78	0.964	0.9	1	0.799	0.842	1
Уругвай	9	1.767	0.137	0.598	0.551	0.721	0.884	0.69	1	0.94	0.97	1
Франция	10	1.753	0.031	0.363	0.673	0.773	0.964	0.69	1	0.782	0.956	1
Испания	11	1.735	0.031	0.278	0.653	0.766	0.964	0.65	1	0.723	0.974	0.999
Чили	12	1.735	0.094	0.465	0.451	0.728	0.929	0.69	0.95	0.869	1	0.999
Люксембург	13	1.727	0.005	0.278	0.74	0.727	0.973	0.83	1	0.139	0.946	1
Коста-Рика	14	1.726	0.187	0.414	0.502	0.722	0.902	0.51	0.97	0.952	0.974	1
Австрия	15	1.726	0.088	0.356	0.709	0.738	0.964	0.81	1	0.69	0.834	1
Нидерланды	16	1.723	0.16	0.329	0.691	0.5	0.964	0.89	1	0.616	0.924	1
Германия	17	1.719	0.028	0.349	0.717	0.721	0.964	0.79	1	0.654	0.856	1
Хорватия	18	1.717	0.052	0.394	0.71	0.603	0.955	0.44	1	0.836	0.958	1
Словения	19	1.715	0.041	0.349	0.716	0.679	0.973	0.67	1	0.728	0.894	1
<b>Высокий уровень</b>												
Италия	20	1.695	0.031	0.309	0.64	0.751	0.973	0.48	1	0.725	0.932	0.978
Ирландия	21	1.694	0.013	0.36	0.657	0.701	0.964	0.77	1	0.642	0.826	1
Португалия	22	1.693	0.051	0.319	0.615	0.715	0.973	0.61	1	0.793	0.834	0.999
Великобритания	23	1.689	0.044	0.302	0.64	0.734	0.955	0.77	1	0.694	0.776	1
Япония	24	1.676	0.012	0.281	0.751	0.815	0.973	0.73	1	0.66	0.5	0.998
Литва	25	1.673	0.041	0.462	0.64	0.551	0.938	0.46	1	0.874	0.826	1
Тунис	26	1.669	0.069	0.407	0.602	0.578	0.831	0.44	0.93	0.921	0.978	0.998
Бельгия	27	1.665	0.011	0.292	0.67	0.735	0.964	0.73	1	0.616	0.72	1
Польша	28	1.663	0.027	0.363	0.655	0.612	0.947	0.46	1	0.735	0.888	1
Эстония	29	1.661	0.047	0.52	0.642	0.53	0.955	0.66	1	0.556	0.886	1
Бразилия	30	1.659	0.17	0.595	0.43	0.569	0.831	0.35	0.9	0.937	0.972	0.999
Словакия	31	1.658	0.017	0.411	0.742	0.59	0.938	0.5	1	0.754	0.798	0.997
США	32	1.649	0.028	0.278	0.592	0.665	0.938	0.73	1	0.339	0.892	0.999
Египет	33	1.647	0.041	0.384	0.656	0.449	0.742	0.28	0.98	0.929	0.966	0.987
Колумбия	34	1.645	0.111	0.499	0.414	0.655	0.848	0.38	0.93	0.954	0.868	0.994
Греция	35	1.643	0.045	0.285	0.657	0.675	0.964	0.47	1	0.703	0.782	0.999
Панама	36	1.643	0.11	0.438	0.439	0.726	0.84	0.34	0.9	0.939	0.888	0.998
Латвия	37	1.642	0.141	0.547	0.623	0.479	0.929	0.5	0.99	0.898	0.662	1
Иордания	38	1.642	0.004	0.38	0.612	0.555	0.813	0.51	0.97	0.88	0.892	1
Израиль	39	1.638	0.018	0.278	0.608	0.767	0.964	0.6	1	0.676	0.65	1
Венгрия	40	1.638	0.268	0.404	0.731	0.556	0.947	0.51	0.99	0.806	0.646	1

1	2	3	4	5	6	7	8	9	10	11	12	13
<b>Средний уровень</b>												
Мексика	41	1.637	0.081	0.37	0.539	0.726	0.742	0.36	0.97	0.866	0.94	0.984
Болгария	42	1.636	0.081	0.431	0.708	0.506	0.911	0.36	0.99	0.797	0.81	1
Албания	43	1.633	0.118	0.394	0.689	0.476	0.866	0.34	0.96	0.945	0.762	1
Доминиканская республика	44	1.633	0.08	0.404	0.484	0.551	0.777	0.3	0.95	0.93	0.948	0.999
Чехия	45	1.631	0.033	0.339	0.746	0.626	0.973	0.520	1	0.591	0.722	0.997
Россия	46	1.63	0.562	0.578	0.601	0.321	0.911	0.21	0.97	0.637	0.938	0.981
Аргентина	47	1.627	0.038	0.622	0.487	0.685	0.875	0.29	0.96	0.958	0.558	1
Вьетнам	48	1.627	0.196	0.411	0.656	0.595	0.866	0.27	0.85	0.963	0.826	0.998
Малайзия	49	1.625	0.161	0.438	0.508	0.627	0.911	0.51	0.99	0.803	0.768	0.998
Намибия	50	1.623	0.659	0.608	0.257	0.236	0.599	0.45	0.87	0.956	0.964	1
Перу	51	1.62	0.127	0.513	0.48	0.647	0.813	0.36	0.83	0.963	0.886	0.997
Марокко	52	1.615	0.016	0.414	0.605	0.596	0.697	0.35	0.81	0.957	0.986	0.987
Ямайка	53	1.615	0.048	0.411	0.545	0.596	0.768	0.31	0.93	0.864	0.944	0.999
Гондурас	54	1.602	0.178	0.285	0.462	0.702	0.795	0.26	0.87	0.964	0.826	1
Гватемала	55	1.596	0.222	0.431	0.449	0.542	0.724	0.31	0.95	0.97	0.824	0.995
Таиланд	56	1.594	0.067	0.387	0.58	0.662	0.938	0.35	0.99	0.873	0.218	0.998
Молдова	57	1.592	0.008	0.428	0.668	0.398	0.875	0.29	0.92	0.857	0.792	0.998
Алжир	58	1.58	0.241	0.404	0.647	0.629	0.706	0.32	0.85	0.906	0.868	0.984
Ботсвана	59	1.58	0.093	0.591	0.395	0.235	0.198	0.58	0.95	0.92	0.946	0.991
Босния и Герцеговина	60	1.576	0.051	0.397	0.738	0.546	0.884	0.32	0.97	0.866	0.17	1
Южная Корея	61	1.575	0.006	0.326	0.684	0.586	0.955	0.56	0.92	0.646	0.634	1
Сальвадор	62	1.568	0.213	0.397	0.476	0.657	0.804	0.39	0.84	0.968	0.702	0.988
Турция	63	1.565	0.083	0.39	0.564	0.52	0.786	0.46	0.96	0.894	0.736	0.955
<b>Низкий уровень</b>												
Филиппины	64	1.562	0.177	0.418	0.555	0.605	0.786	0.23	0.85	0.968	0.692	0.998
Боливия	65	1.551	0.188	0.891	0.399	0.513	0.554	0.3	0.85	0.955	0.528	0.995
Эквадор	66	1.55	0.086	0.424	0.464	0.641	0.813	0.2	0.94	0.927	0.494	0.997
Индонезия	67	1.55	0.142	0.441	0.657	0.55	0.768	0.26	0.77	0.945	0.748	0.998
Индия	68	1.549	0.122	0.411	0.632	0.522	0.492	0.34	0.86	0.964	0.862	0.972
Бенин	69	1.544	0.251	0.445	0.635	0.439	0.216	0.31	0.67	0.99	0.976	0.999
Непал	70	1.539	0.345	0.414	0.528	0.732	0.59	0.27	0.9	0.996	0.454	0.985
Румыния	71	1.534	0.609	0.407	0.69	0.526	0.875	0.38	0.57	0.853	0.682	0.999
Тринидад и Тобаго	72	1.534	0.097	0.424	0.611	0.524	0.706	0.36	0.91	0.449	0.972	1
Сенегал	73	1.532	0.16	0.435	0.587	0.457	0.465	0.34	0.76	0.985	0.878	1
Камерун	74	1.531	0.329	0.489	0.554	0.381	0.225	0.23	0.66	0.987	0.982	0.998
Китай	75	1.531	0.066	0.387	0.531	0.6	0.822	0.36	0.77	0.859	0.788	0.994
Танзания	76	1.529	0.361	0.431	0.654	0.337	0.34	0.3	0.62	0.996	0.946	0.997
Казахстан	77	1.525	0.425	0.458	0.661	0.293	0.768	0.22	0.86	0.606	0.904	1
Украина	78	1.52	0.01	0.418	0.719	0.379	0.822	0.25	0.96	0.768	0.474	0.999
Никарагуа	79	1.506	0.227	0.469	0.569	0.61	0.742	0.25	0.79	0.973	0.508	0.973
Грузия	80	1.505	0.144	0.452	0.596	0.498	0.75	0.39	0.82	0.971	0.404	0.987
Бангладеш	81	1.504	0.136	0.418	0.666	0.513	0.537	0.21	0.74	0.992	0.722	0.992

1	2	3	4	5	6	7	8	9	10	11	12	13	14
Азербайджан	82	1.504	0.651	0.025	0.39	0.635	0.444	0.349	0.19	0.77	0.876	0.936	1
Таджикистан	83	1.501	0.654	0.161	0.424	0.674	0.336	0.501	0.2	0.59	0.971	0.902	0.989
Венесуэла	84	1.494	0.66	0.346	0.438	0.518	0.679	0.84	0.19	0.83	0.822	0.594	0.937
Киргизия	85	1.492	0.662	0.17	0.448	0.697	0.417	0.679	0.18	0.77	0.961	0.54	0.998
Пакистан	86	1.487	0.667	0.158	0.414	0.694	0.529	0.305	0.25	0.91	0.973	0.35	0.991
Замбия	87	1.487	0.668	0.347	0.499	0.492	0.268	0.091	0.28	0.58	0.993	0.922	1
Кения	88	1.485	0.67	0.313	0.431	0.575	0.345	0.296	0.21	0.61	0.99	0.89	0.998
ЮАР	89	1.484	0.67	0.048	0.431	0.422	0.254	0.501	0.49	0.88	0.7	0.948	0.993
Армения	90	1.479	0.676	0.023	0.407	0.662	0.485	0.813	0.29	0.92	0.874	0.91	0.043
Камбоджа	91	1.469	0.686	0.285	0.428	0.583	0.445	0.421	0.18	0.41	0.999	0.864	0.997
<b>Очень низкий уровень</b>													
Нигерия	92	1.438	0.716	0.375	0.435	0.563	0.376	0.118	0.27	0.48	0.979	0.814	1
Зимбабве	93	1.429	0.725	0.267	0.414	0.499	0.076	0.51	0.18	0.81	0.974	0.502	0.994
Узбекистан	94	1.393	0.762	0.032	0.401	0.632	0.429	0.394	0.18	0.82	0.832	0.48	1
Эфиопия	95	1.344	0.811	0.357	0.418	0.7	0.461	0.314	0.26	0.22	0.997	0.918	0.349
Мозамбик	96	1.331	0.824	0.38	0.513	0.527	0.296	0.144	0.26	0.43	0.997	0.956	0.313

Таблица 2. Корреляционная матрица глобальных угроз

	<i>ES</i>	<i>BB</i>	<i>GINI</i>	<i>GD</i>	<i>CM</i>	<i>CP</i>	<i>WA</i>	<i>GW</i>	<i>SF</i>	<i>ND</i>
<i>ES</i>	1.00	0.46	-0.12	-0.20	-0.23	-0.06	-0.34	0.03	0.10	-0.06
<i>BB</i>	0.46	1.00	-0.33	-0.18	-0.17	-0.01	-0.09	0.11	0.07	-0.02
<i>GINI</i>	-0.12	-0.33	1.00	0.24	0.33	0.34	0.16	-0.44	-0.07	-0.03
<i>GD</i>	-0.20	-0.18	0.24	1.00	<b>0.70</b>	0.60	0.54	-0.38	0.00	0.17
<i>CM</i>	-0.23	-0.17	0.33	<b>0.70</b>	1.00	0.58	<b>0.78</b>	-0.55	-0.04	0.22
<i>CP</i>	-0.06	-0.01	0.34	0.60	0.58	1.00	0.56	- <b>0.67</b>	0.28	0.16
<i>WA</i>	-0.34	-0.09	0.16	0.54	<b>0.78</b>	0.56	1.00	-0.51	-0.05	0.36
<i>GW</i>	0.03	0.11	-0.44	-0.38	-0.55	- <b>0.67</b>	-0.51	1.00	-0.16	-0.14
<i>SF</i>	0.10	0.07	-0.07	0.00	-0.04	0.28	-0.05	-0.16	1.00	-0.11
<i>ND</i>	-0.06	-0.02	-0.03	0.17	0.22	0.16	0.36	-0.14	-0.11	1.00

Так, согласно данным, приведенным в табл. 2, можно выделить две группы глобальных угроз:  $G_1 = \{GD, CM, CP, WA, GW\}$ , когда между угрозами наблюдается сильная связь, и  $G_2 = \{ES, BB, SF, ND, GINI\}$ , когда угрозы слабо зависимы.

**Применение метода главных компонент для анализа глобальных угроз.** Использование метода главных компонент позволит выявить минимальное число скрытых факторов (главных компонент), отображающих основные свойства исходных десяти угроз, и одновременно уменьшить степень зависимости этих факторов от собственных остаточных случайных компонент.

Выполним поиск нового ортонормированного базиса векторов-компонент [6, 7], разложив матрицу  $X$  в виде

$$X = T \cdot P^T + E, \quad (2)$$

где  $T$  — матрица счетов размерности  $n \times m'$  ( $m' \leq m$ ) ( $m'$  — размерность пространства глобальных угроз;  $m$  — количество главных компонент, выбранных для проецирования);  $P$  — матрица нагрузок размерности  $m' \times m$ ;  $E$  — матрица остатков.

Матрица счетов задает множество векторов  $T_i = \langle t_i^j \rangle$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, m'}$ , определяющих проекции векторов  $X_i$ ,  $i = \overline{1, n}$ , в пространстве главных компонент (число

компонент равно  $m' \leq m$ ). Нахождение главных компонент связано с вычислением собственных векторов ковариационной матрицы [3, 7], определяемой как

$$C = (c_{ij}, c_{ij} = \text{cov}(X^i, X^j)), i, j = \overline{1, m}, \quad (3)$$

где  $\text{cov}(X^i, X^j) = \frac{\sum_{k=1}^n (x_k^i - \bar{X}^i) \cdot (x_k^j - \bar{X}^j)}{n-1}$  — ковариация признаков показателей  $X^i$  и  $X^j$ .

Матрица нагрузок задает отображение базиса исходного пространства в пространство главных компонент, в котором можно оценить значимость каждой глобальной угрозы и зависимость между ними.

**Таблица 3.** Факторные нагрузки глобальных угроз

Глобальная угроза	Нагрузки на факторы				
	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$
<i>ES</i>	-0.332	<b>0.701</b>	-0.002	-0.445	-0.063
<i>BB</i>	-0.280	<b>0.709</b>	0.372	-0.156	0.239
<i>GINI</i>	0.480	-0.192	<b>-0.550</b>	-0.499	-0.199
<i>GD</i>	<b>0.767</b>	0.009	0.088	0.055	0.321
<i>CM</i>	<b>0.875</b>	-0.011	0.139	-0.042	0.221
<i>CP</i>	<b>0.796</b>	0.389	-0.140	0.059	-0.010
<i>WA</i>	<b>0.821</b>	-0.016	0.350	0.105	0.089
<i>GW</i>	<b>-0.740</b>	-0.281	0.244	0.165	0.192
<i>SF</i>	0.055	0.533	-0.418	<b>0.658</b>	-0.227
<i>ND</i>	0.315	-0.024	0.652	-0.016	<b>-0.662</b>

В табл. 3 представлены факторные нагрузки в пространстве десяти глобальных угроз и пяти главных факторов  $F_1, \dots, F_5$ . Как видим, глобальные угрозы  $G_1 = \{GD, CM, CP, WA, GW\}$  имеют наибольшие нагрузки на первый фактор. Группа  $G_2$  разбивается на четыре подгруппы:  $G_{22} = \{ES, BB\}$ ,  $G_{23} = \{GINI\}$ ,  $G_{24} = \{SF\}$ ,  $G_{25} = \{ND\}$  (второй индекс соответствует номеру фактора с максимальной нагрузкой). При этом угрозы подгруппы  $G_{22}$  имеют практически одинаковые нагрузки на второй фактор, но статистически слабо связаны между собой (см. также табл. 2).

Анализ глобальных угроз группы  $G_1$  в пространстве главных компонент говорит о том, что эта группа может быть разделена на две подгруппы:  $G_{11} = \{GD, CM, WA\}$  и  $G_{12} = \{CP, GW\}$ . Для угроз группы  $G_{11}$  характерны высокий уровень положительной корреляции и направленность соответствующих векторов вдоль оси первой компоненты. Угрозы группы  $G_{12}$  отличаются высоким уровнем отрицательной корреляции, что также подтверждается данными, представленными в табл. 2.

Из работы [7] следует, что сумма дисперсий глобальных угроз равна сумме дисперсий всех главных компонент, которые упорядочены в соответствии с долями их дисперсий. Поэтому анализируя изменение относительной доли дисперсии, вносимой первыми  $m'$  главными компонентами, можно определить число компонент, которое целесообразно использовать для последующего рассмотрения. При этом для выбора достаточного числа  $m' \leq m$  главных компонент часто применяют кумулятивную дисперсию [7]

**Таблица 4.** Характеристики факторов

Номер фактора	Собственное значение	Доля общей дисперсии, %	Кумулятивная дисперсия, %
$F_1$	<b>3.73</b>	37.31	37.31
$F_2$	<b>1.55</b>	15.47	52.78
$F_3$	<b>1.27</b>	12.69	65.47
$F_4$	0.95	9.50	74.98
$F_5$	0.79	7.88	82.86
$F_6$	0.57	5.68	88.54
$F_7$	0.43	4.29	92.83

$$D_i = \frac{\sum_{j=1}^i \lambda_j}{m}, i = \overline{1, m},$$

где  $\lambda_j, j = \overline{1, m}$ , — собственные числа ковариационной матрицы  $C$ .

На основе данных, приведенных в табл. 4, и в соответствии с критериями Кайзера [8] и Кэттеля [9] для последующего исследования достаточно выде-

лить три первых главных фактора. Это позволит представить около 65 % данных об исходных глобальных угрозах.

Как видим, глобальные угрозы групп  $G_1$  и  $G_{22}$  являются наиболее информативными относительно компаративного анализа стран по защищенности от влияния глобальных угроз.

**Определение значимости глобальных угроз.** Пусть матрица факторных нагрузок  $P$  определяется соотношением (2) при  $m = m'$ . Тогда дисперсию глобальной угрозы  $X^i$  можно определить как  $D(X^i) = \sum_{j=1}^m p_{ij}^2$ . В то же время дисперсия, вносимая фактором  $F_j$ , определяется по формуле

$$D(F_j) = \sum_{i=1}^m p_{ij}^2 = \lambda_j.$$

Предположим, что сумма нормированных нагрузок глобальных угроз на главные компоненты равна доли общей дисперсии, определяемой этой главной компонентой,

$$D_i = \frac{\sum_{k=1}^m p_{ik}^2}{\sum_{j=1}^m \lambda_j}, \quad i = \overline{1, m}, \quad p_{ik} \in P. \quad (4)$$

Тогда с учетом того, что  $\sum_{i=1}^m D_i = 1$ ,

получим весовые коэффициенты, определяющие степень значимости глобальных угроз

$$X^i \rightarrow K^i = \sum_{j=1}^m D_j p_{ij}^2,$$

где  $D_j$  вычисляются с использованием соотношения (4).

На рис. 1 представлена диаграмма значимости глобальных угроз. В этом случае они могут быть разделены на следующие категории:  $G_1 = \{GD, CM, CP, WA, GW\}$  — высокая степень значимости (более 0.10);  $G_2 = \{ES, BB, SF, ND, GINI\}$  — низкая степень значимости (менее чем 0.10).

Таким образом, для рассматриваемых стран наиболее значимыми угрозами являются  $G_1 = \{GD, CM, CP, WA, GW\}$ . При этом наблюдается высокая степень их парной корреляции.

На рис. 2 представлены результаты кластеризации стран, полученные с помощью метода  $k$ -средних [3]. Кластеризация выполнена по уровню безопасности в пространстве главных компонент. Выделим четыре группы стран.

Первая группа (кластер  $C_4$ ): Австралия, Норвегия, Канада. Для этих стран характерны высокие значения как первой  $F_1$ , так и второй  $F_2$  главных компонент.

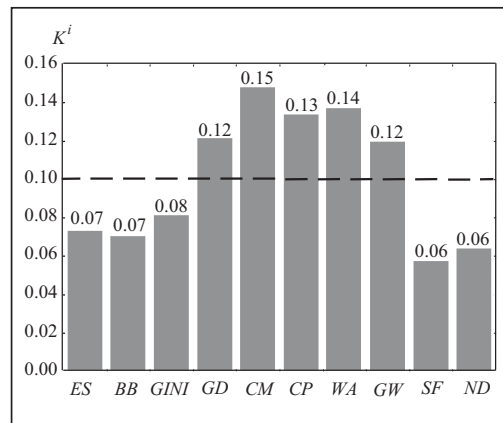


Рис. 1. Диаграмма значимости глобальных угроз

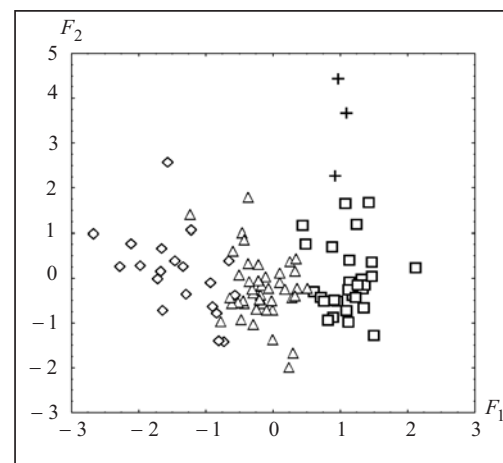


Рис. 2. Результаты кластеризации стран по уровню защищенности от глобальных угроз:  $\diamond$  — кластер  $C_4$ ,  $\Delta$  — кластер  $C_3$ ,  $\square$  — кластер  $C_2$ ,  $+$  — кластер  $C_1$

Вторая группа (кластер  $C_3$ ): Южная Корея, Чехия, Израиль, Греция, США, Словакия, Эстония, Польша, Бельгия, Япония, Великобритания, Португалия, Ирландия, Италия, Словения, Хорватия, Германия, Нидерланды, Австрия, Люксембург, Чили, Испания, Франция, Уругвай, Швейцария, Новая Зеландия, Финляндия, Дания, Швеция. Здесь характерны высокие значения компоненты  $F_1$  и стандартные отклонения компоненты  $F_2$ .

Третья группа (кластер  $C_2$ ): Армения, Киргизия, Венесуэла, Грузия, Никарагуа, Украина, Казахстан, Китай, Тринидад и Тобаго, Румыния, Непал, Индонезия, Эквадор, Боливия, Филиппины, Турция, Сальвадор, Босния и Герцеговина, Алжир, Молдова, Таиланд, Гватемала, Гондурас, Ямайка, Марокко, Перу, Малайзия, Вьетнам, Аргентина, Россия, Доминиканская Республика, Албания, Болгария, Мексика, Венгрия, Иордания, Латвия, Панама, Колумбия, Египет, Бразилия, Тунис, Литва, Коста-Рика. Здесь характерны стандартные отклонения первой и второй компонент.

Четвертая группа (кластер  $C_1$ ): Мозамбик, Эфиопия, Узбекистан, Зимбабве, Нигерия, Камбоджа, ЮАР, Кения, Замбия, Пакистан, Таджикистан, Азербайджан, Бангладеш, Танзания, Камерун, Сенегал, Бенин, Индия, Ботсвана, Намибия. Характерны очень низкие значения первой главной компоненты.

Для каждого кластера была проведена оценка значимости глобальных угроз, результаты которой представлены на рис. 3. Для кластера  $C_1$  наиболее значимыми являются угрозы  $\{BB, GINI, CP, WA\}$ ; для кластера  $C_2$  — угрозы  $\{GINI, GD, WA, GW\}$ ; для кластера  $C_3$  — угрозы  $\{ES, BB, GINI, CM, CP\}$ ; для кластера  $C_4$  — угрозы  $\{ES, GINI, GD, CM, CP, GW\}$ .

Таким образом, глобальные угрозы можно разбить на три категории:  $L_1 = \{GINI\}$  — наиболее значимый показатель;  $L_2 = \{CP, WA\}$  — группа показателей со средней степенью значимости;  $L_3 = \{ES, BB, GD, CM, GW\}$  — низкая степень значимости.

Как видим, страны, близкие по уровню защищенности от глобальных угроз, отличаются, в первую очередь, по уровню показателя  $L_1 = \{GINI\}$ , а также по уровню показателей  $L_2 = \{CP, WA\}$ .

**Исследование зависимости между уровнем безопасности стран (норма Минковского) и глобальными угрозами.** Для детального анализа глобальных угроз, с которыми сталкиваются страны, необходимо локализовать выборку, по которой оценивается корреляция между угрозами. Естественно предположить, что в такую выборку должны попадать «похожие» между собой страны, степень схожести которых может оцениваться, например, как евклидово расстояние в пространстве угроз. Кроме того, степень влияния угроз на коэффициент корреляции должна быть тем больше, чем ближе страна находится к точке, в которой анализируется корреляция.

В соответствии с этими предположениями определим взвешенное среднее [10] в виде

$$m(X, W) = \frac{\sum_i w_i x_i}{\sum_i w_i}, \quad (5)$$

где  $X$  — выборка данных,  $W$  — весовая функция.

Определим  $W$  как функцию, зависящую от расстояния, в виде

$$W(x, t) = e^{-\lambda d(x, t)},$$

где  $d(x, t)$  — расстояние между точками  $x, t \in R^n$ ;  $\lambda$  — параметр распределения. Подставив  $W$  в (5), получим соотношение для вычисления взвешенного локализованного среднего в точке  $t$  по выборке  $X$

$$m(X, t) = \frac{\sum_i e^{-\lambda d(t, x_i)} \cdot x_i}{\sum_i e^{-\lambda d(t, x_i)}}, \quad x_i \in X.$$



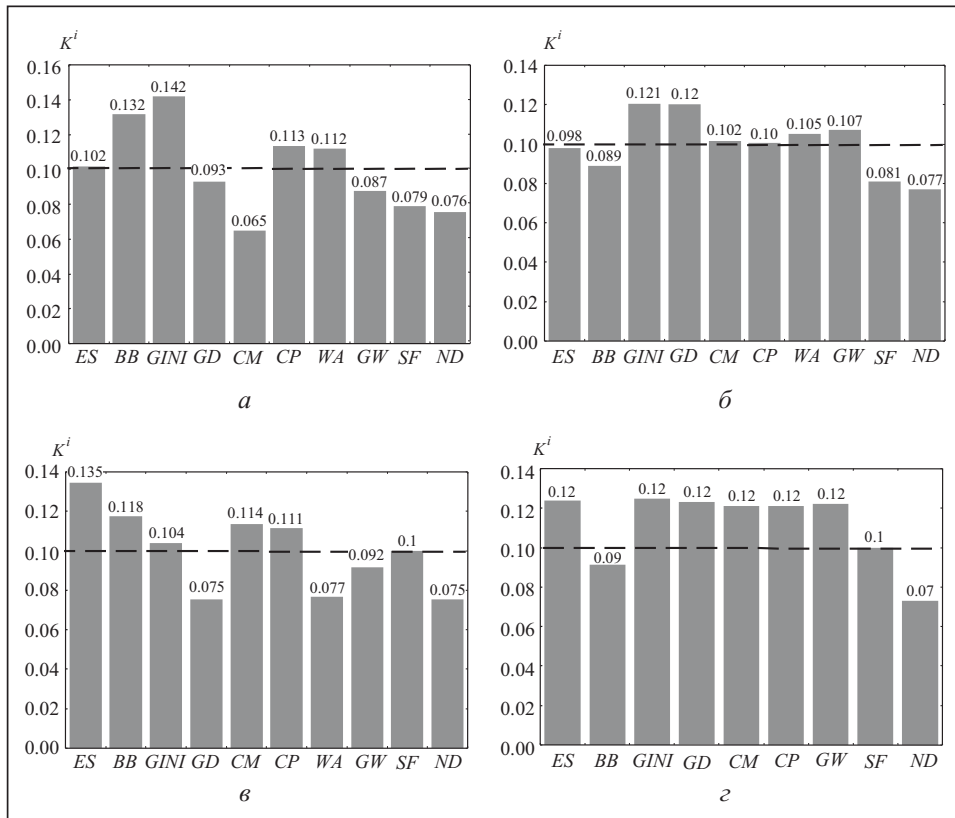


Рис. 3. Диаграмма значимости глобальных угроз для кластера  $C_1$  (а); кластера  $C_2$  (б); кластера  $C_3$  (в); кластера  $C_4$  (г)

Аналогично можно определить взвешенную локализованную ковариацию

$$\text{cov}(X, Y, t) = \frac{\sum_i e^{-\lambda d(t, x_i)} (x_i - m(X, t))(y_i - m(Y, t))}{\sum_i e^{-\lambda d(t, x_i)}}$$

и взвешенную локализованную корреляцию (Weighted Localized Correlation (WLC)):

$$\text{corr}(X, Y, t) = \frac{\text{cov}(X, Y, t)}{\sqrt{\text{cov}(X, X, t) \cdot \text{cov}(Y, Y, t)}}$$

Выбор параметра распределения весов  $\lambda$  можно осуществить таким образом, чтобы ограничить область влияния точек, расположенных на больших расстояниях. Например, точки, равноудаленные от точки  $t$ , в которой вычисляется WLC, имеют одинаковый вес, равный  $s$  (масштаб распределения):

$$e^{-\lambda(t)m(d_t)} = s; \quad \lambda(t) = \frac{\ln(s)}{m(d_t)}, \quad (6)$$

где  $m(d_t)$  — среднее расстояние от точек выборки до точки  $t$ .

При масштабе распределения, равном единице, WLC совпадает с коэффициентом корреляции Пирсона. Как видно из (6), параметр распределения весов вычисляется для каждой точки  $t$ , в качестве которых выступают точки выборки. Для каждой новой точки заново вычисляется значение среднего расстояния  $m(d_t)$ . Следовательно, предложенный метод оценки локальной зависимости угроз является адаптивным.

Интерпретация значений взвешенной локализованной корреляции WLC сведена в табл. 5.

**Таблица 5.** Интерпретация значений WLC

Значение WLC	Характеристика связи	Поведение глобальных угроз
$[-1.0, -0.5)$	Высокая степень отрицательной корреляции (более 25%). Рост одной угрозы приводит к снижению другой.	Исследуемая угроза имеет невысокий (по сравнению с остальными) вклад в общую удаленность от глобальных угроз.
$[-0.5, -0.3)$	Средняя степень отрицательной корреляции (9–25%). Рост одной угрозы приводит к снижению другой.	При уменьшении конкретной угрозы общая удаленность от совокупности глобальных угроз уменьшается в средней степени.
$[-0.3, 0.3]$	Низкая степень корреляции (менее 9%)	Общая удаленность от глобальных угроз слабо зависит от исследуемой угрозы.
$(0.3, 0.5]$	Средняя степень положительной корреляции (9–25%). Рост одной угрозы приводит к росту другой.	Общая удаленность от глобальных угроз возрастает в средней степени при уменьшении конкретной угрозы.
$(0.5, 1.0]$	Высокая степень положительной корреляции. Рост одной угрозы приводит к росту другой (более чем на 25%).	Исследуемая угроза существенно влияет на общую удаленность от совокупности глобальных угроз.

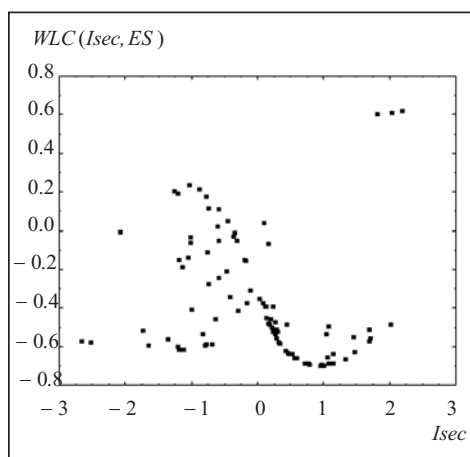


Рис. 4. WLC между *Isec* и *ES*

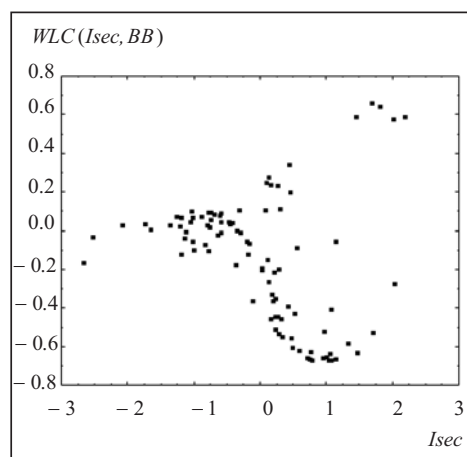


Рис. 5. WLC между *Isec* и *BB*

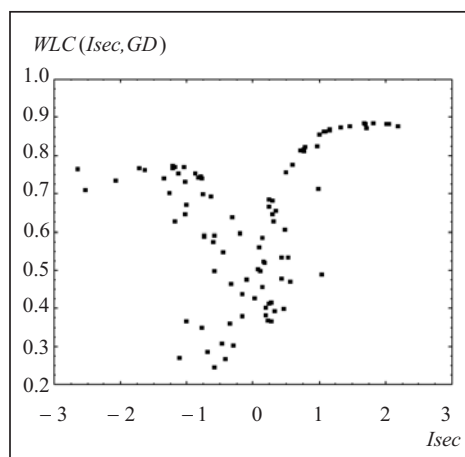


Рис. 6. WLC между *Isec* и *GD*

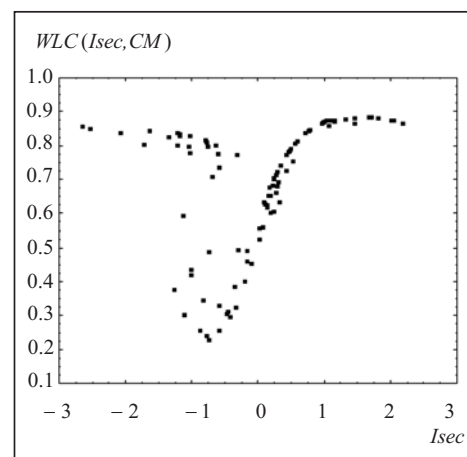


Рис. 7. WLC между *Isec* и *CM*

На рис. 4–8 приведены графики значений WLC между нормой Минковского ( $I_{sec}$ ) и глобальными угрозами  $L_3 = \{ES, FB, GD, CM, GW\}$ .

Анализируя связь между нормой Минковского и уровнем энергетической безопасности (рис. 4), можно сказать, что для стран с высокими значениями нормы Минковского (Австралия, Норвегия, Канада) обнаруживается положительная связь средней степени, что свидетельствует о значимом вкладе уровня энергетической безопасности в общую оценку безопасности страны. С другой стороны, для стран со значениями нормы Минковского, близкими к средним, наблюдается отрицательная корреляция среднего и высокого уровней, что свидетельствует о компенсации низкого уровня энергетической безопасности за счет других составляющих.

Для стран с очень высоким уровнем общей безопасности (Финляндия, Новая Зеландия, Австралия, Канада, Швеция) обнаруживается положительная корреляционная связь средней степени между нормой Минковского и уровнем биологического баланса (рис. 5). В то же время для группы стран с высоким уровнем безопасности (Израиль, Египет, Хорватия, Дания, Греция, США, Польша, Франция, Бельгия, Япония, Ирландия, Швейцария, Австрия, Германия, Словения, Великобритания, Испания, Португалия, Люксембург, Италия, Нидерланды) имеет место высокая отрицательная корреляция.

На рис. 6 прослеживается тенденция к возрастанию корреляции между нормой Минковского и уязвимостью к глобальным болезням для минимальных и максимальных уровней безопасности. Для WLC между нормой Минковского и уровнем детской смертности наблюдается аналогичная тенденция (рис. 7). Между угрозой глобального потепления и нормой Минковского для большинства стран наблюдается отрицательная корреляция средней и высокой степени (рис. 8).

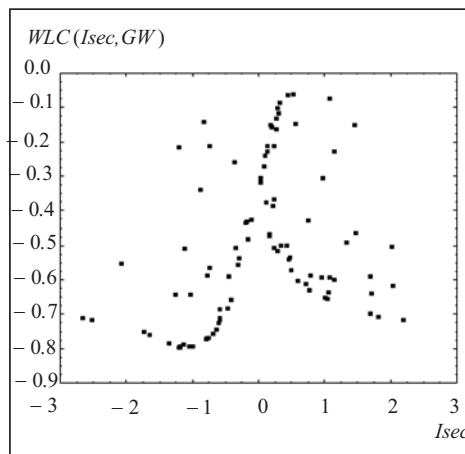


Рис. 8. WLC между  $I_{sec}$  и  $GW$

## ЗАКЛЮЧЕНИЕ

Применение методов многомерного статистического анализа для исследования влияния совокупности глобальных угроз на безопасность стран мира позволило получить следующие результаты.

Выполненный анализ статистической взаимосвязи глобальных угроз позволил рассчитать для них коэффициенты парной корреляции Пирсона. Согласно полученным результатам выделены две группы глобальных угроз:  $G_1 = \{GD, CM, CP, WA, GW\}$ , для которой между угрозами наблюдается сильная связь, и  $G_2 = \{ES, BB, SF, ND, GINI\}$ , содержащей слабо зависимые угрозы. При этом упрощается содержательная интерпретация параметров модели расчета уровня безопасности стран, а именно идентификация причинно-следственных связей за счет снижения эффекта мультиколлинеарности.

На основе применения метода главных компонент и в соответствии с критериями Кайзера и Кэттеля безопасность стран представлена не всеми десятью глобальными угрозами, а тремя наиболее существенными латентными факторами. Это позволило представить около 65 % данных об исходных глобальных угрозах и существенно упростить последующие исследования. Показано, что глобальные угрозы групп  $G_1 = \{GD, CM, CP, WA, GW\}$  и  $G_2 = \{ES, BB\}$  являются наиболее информативными с точки зрения компаративного анализа безопасности стран.

Установлено, что для угроз группы  $G_1$  характерны высокая степень значимости и высокая степень их парной корреляции, а для угроз группы  $G_2$  — низкая степень значимости.

С помощью метода  $k$ -средних выполнена кластеризация стран по уровню безопасности в пространстве главных компонент. При этом выделены четыре груп-

пы стран. Первая группа наиболее безопасных стран характеризуется высокими значениями как первого, так и второго фактора. Для второй группы стран с несколько меньшим уровнем безопасности получены высокие значения первого фактора и стандартные отклонения второго фактора. Третья группа стран с низким уровнем безопасности характеризуется стандартными отклонениями первого и второго факторов. Четвертая группа наиболее уязвимых к воздействию глобальных угроз стран характеризуется очень низкими значениями первого фактора. Для каждой группы стран была проведена оценка значимости глобальных угроз.

Выполнена модификация метода взвешенной локальной корреляции, позволяющая провести исследование зависимости уровня безопасности различных стран мира (нормы Минковского) от отдельных глобальных угроз. С использованием этого метода детально проанализированы зависимости между нормой Минковского и наиболее существенными угрозами  $\{EC, BB, GD, CM, GW\}$ . Анализируя связь между нормой Минковского и уровнем энергетической безопасности, установлено, что для стран с высоким уровнем общей безопасности (Австралия, Норвегия, Канада, Финляндия, Новая Зеландия, Австралия, Швеция) обнаруживается положительная связь средней степени, что свидетельствует о значимом вкладе уровня энергетической безопасности в общую оценку безопасности страны. Для этих же стран обнаруживается положительная корреляционная связь средней степени между нормой Минковского и уровнем биологического баланса. Прослеживается также тенденция к возрастанию корреляции между нормой Минковского и уязвимостью к глобальным болезням для минимальных и максимальных уровней безопасности. Показано, что между нормой Минковского и уровнем детской смертности наблюдается аналогичная тенденция, а между угрозой глобального потепления и нормой Минковского для большинства стран мира наблюдается отрицательная корреляция средней и высокой степени.

#### СПИСОК ЛИТЕРАТУРЫ

1. Згуровский М.З. Закономерность течения системных мировых конфликтов и глобальные угрозы XXI столетия // Кибернетика и системный анализ. — 2007. — № 5. — С. 87–99.
2. Аналіз сталого розвитку. Ч. 1. Глобальне моделювання процесів сталого розвитку в контексті якості і безпеки життя людей, 2009. — [www.wdc.org.ua](http://www.wdc.org.ua).
3. Прикладная статистика: Классификация и снижение размерности. Справочное изд. / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин / Под ред. С.А. Айвазяна. — М.: Финансы и статистика, 1989. — 607 с.
4. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное изд. — М.: Финансы и статистика, 1983. — 471 с.
5. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Исследование зависимостей. Справочное изд. / Под ред. С.А. Айвазяна. — М.: Финансы и статистика, 1985. — 487 с.
6. Jolliffe I.T. Principal component analysis: Springer series in statistics, 2nd ed. — NY: Springer, 2002, XXIX. — 487 p.
7. Smith L.I. A tutorial on principal components analysis, 2002. — [http://www.cs.otago.ac.nz/cosc453/student\\_tutorials/principal\\_components.pdf](http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf).
8. Kaiser H.F. The application of electronic computers to factor analysis // Educational and Psychological Measurement. — 1960. — 20. — P. 141–151, .
9. Cattell R.B. The scree test for the number of factors // Multivariate Behavioral Research. — 1966. — N 1. — P. 245–276.
10. A MATLAB Toolbox for computing weighted correlation coefficients, 2008. — <http://www.mathworks.com/matlabcentral/fileexchange/20846>.

Поступила 05.11.2009