

Процедура за анализ и оценка на качеството на водите в река Струма

К. Бошнаков¹, Б. Куманова¹, М. Върбанов², Е. Михайлов¹

Procedure for Analysis and Assessment of Water Quality in the Struma River

K. Boshnakov¹, B. Koumanova¹, M. Varbanov², E. Mihailov¹

¹ University of Chemical Technology and Metallurgy (UCTM), 8 Sv. Kliment Ohridski Blvd., 1756 Sofia, Bulgaria, kb@uctm.edu

² Department Geography, National Institute of Geophysics, Geodesy and Geography, Bulgarian Academy of Sciences, Acad. G. Bonchev St., Bl. 3, 1113 Sofia, Bulgaria, marian_varbanov@geophys.bas.bg

Key Words: Neural classifiers; state of river waters; Struma river; presence of inadmissible pollution.

Abstract. The aim of the present work is to study the possibilities for developing a system for monitoring the state of the water on the Struma River and the presence of unacceptable pollution. The development of the algorithms of the monitoring system is in accordance with the normative documents of the Republic of Bulgaria and is built on the basis of monitoring measurements for the quality of the river waters from six monitoring points for the period 2010-2019. The following variables characterizing water quality were analyzed: ammonium nitrogen (NH₄-N), nitrate nitrogen (NO₃-N), nitrite nitrogen (NO₂-N), pH, biological oxygen demand (BOD), electrical conductivity (EC), total nitrogen (TN), total organic carbon (TOC), total phosphorus (TP), orthophosphate phosphorus (PO₄-P), dissolved oxygen (DO), temperature (T) and chemical oxygen demand (COD). The section of the river Struma from before Blagoevgrad to the border (bridge for the village of Topolnitsa) was chosen for the research. The available measurement data from six monitoring points in the considered section of the Struma River are as synchronized as possible. Correlation analyzes were performed for the possibility of combining the available measurements from the six monitoring points and for the presence of a correlation dependence between the measured variables characterizing the water quality in the Struma River. The presence of correlations between the variables directs the use of the principal components method as a mathematical apparatus for building the monitoring system. Procedures have been developed for classifying the status of river waters depending on current measurements / analyzes. Expressions for the main components for eleven of the measured variables characterizing water quality are derived. Research has been carried out to develop a procedure for classifying these measurements into the terms "excellent", "good" and "moderate". Possibilities for classifications with neural networks have been studied. Measuring some of these variables is a lengthy procedure and can hardly be used for operational actions. For this reason, the possibilities for developing a procedure for monitoring and classifying the status of river waters on the basis of variables, which in principle can be measured automatically (pH, EC, DO and T), have been explored. They also apply the principal components method and classification with neural networks in the same classification categories. The developed approaches for monitoring the quality of the waters of the Struma River can be applied for each measurement in the six monitored monitoring points.

1. Увод

Целта на настоящата работа е разработване на математическо и алгоритмично осигуряване за мониторингова система за качеството на водите на река Струма.

За функциониране на мониторинговата система е необходимо постъпване на измервания/анализи за текущото състояние на нормативно определени физикохимични показатели. Изследвани са два подхода: първият се основава на стойностите на всички физикохимични показатели, характеризиращи качеството на речните води, а втория – само на променливи, които по принцип могат да се измерват в реално време. Информацията за качеството на водите и наличието на аномални явления, рефлектиращи в качеството и, генерирана от мониторинговата система е предназначена за съответната басейнова дирекция [2,3,4].

В редица литературни източници са представени използваните математически методи за оценка на състоянието на речни води в различни точки на света. В [6] са проведени изследвания за класифициране и предсказване на качеството на водите на река Донг Наи във Виетнам за периода от 2012 до 2019 г. За класификация на качеството на водите се използват невронни мрежи с право разпространение, а за прогнозиране – нелинейна авто-регресионна-пълзящо средно невронна мрежа (NAR-MA).

За класификация на качеството на водите на река Муда (Малайзия) се използват клъстерен анализ, метод на главните компоненти (PCA) и дискриминантен анализ (DA) [7]. За същата цел в [8] се използват невронни мрежи с право разпространение. За моделиране на качеството на речните води в [9] се използва комбинация от метода на главните компоненти и многомерна линейна регресия, като посредством PCA се намалява размерността на задачата за линейния регресионен анализ. Аналогично в [10] за целите на класификацията се използват методът на главните компоненти и статистически техники.

2. Мониторингови пунктове по течението на река Струма

За провеждане на изследванията в статията са използвани данни от регулярни измервания на състава на водите на река Струма, които се провеждат във всеки мониторингов пункт (МП) четири пъти в годината – по едно за всеки сезон. За провеждане на изследванията е избран участъкът на река Струма от преди Благоевград до границата (мост за с. Тополница).

Съгласно класификацията, направена на типовете повърхностни водни тела в Западнорелски район за басейново управление [2], по поречието на река Струма има един участък – от изворите и до язовир Студена, който е тип R3 (планински тип) и след него има единадесет участъка, които са тип R5 (полупланински тип).

Разглежданият участък от река Струма се отнася към R5 (полупланински тип) и в него са разположени следните пунктове за мониторинг: р. Струма преди гр. Благоевград (МП1); р. Струма след Благоевград (МП2); р. Струма след гр. Симитли, шос. мост на Е79 в м. Орловец (МП3); р. Струма преди гр. Кресна (МП4); р. Струма след вливане на р. Санданска Бистрица, след гр. Сандански (МП5); р. Струма при границата (мост за с. Тополница) (МП6).

При регулярните измервания се измерват/анализират следните физикохимични променливи, характеризиращи качеството на водите: амониен азот (NH₄-N), нитратен азот (NO₃-N), нитритен азот (NO₂-N), рН, биологична потребност от кислород (BOD), електропроводимост (EC), общ азот (TN), общ органичен въглерод (TOC), общ фосфор (TP), ортофосфати (PO₄-P), разтворен кислород (DO), температура (T) и химична потребност от кислород (COD). За провеждане на тези изследвания се използват данни от измервания, проведени в периода 2010-2019 г.

В Наредба № Н-4 за характеризирание на повърхностните води [1] са включени изисквания към следните физикохимични елементи за качество, пояснени по-горе: NH₄-N, NO₃-N, NO₂-N, рН, BOD, EC, TN, TP, PO₄-P и DO. Състоянието на речните води според наредбата е класифицирано в три групи: „отлично“, „добро“ и „умерено“. Изследванията в работата са съобразени изцяло с изискванията в цитираната наредба.

3. Анализ на данните от регулярни измервания на качеството на водите на река Струма

Табл. 1. Извадка от провежданите измервания/анализи

№	NH ₄ -N	NO ₃ -N	NO ₂ -N	рН	BOD
	mg/l	mg/l	mg/l	-	mg/l
1	0.029	0.94	0.024	8.63	2.1
2	0.027	0.7	0.0194	8.28	2.4
3	0.033	0.606	0.014	8.3	2.57
...

№	EC	TN	TOC	TP	
	µS/sm	mg/l	mg/l	mg/l	
1	466.9	1.1	2.4	0.052	
2	485.8	0.98	2.5	0.074	
3	336	1.5	3.5	0.022	
...
№	PO ₄ -P	DO	T	COD	
	mg/l	mg/l	mg/l	mg/l	
1	0.014	9.57	24.3	6	
2	0.029	8.95	18.4	6.8	
3	0.0093	10	10.2	15	
...

3.1. Проверка на съвместимостта на данните от мониторинговите пунктове

В един мониторингов пункт съгласно регламента за измерване максималният брой измервания в периода 2010-2019 г. е 40. За някои пунктове реално проведените измервания са по-малко на брой. Този брой е недостатъчен, за да се направят достоверни статистически изводи.

По тази причина първото изследване, свързано с данните от измерванията, е статистическа проверка на това дали данните от различните мониторингови пунктове са съвместими, което предполага квазистационарност на състава на речните води за времето, през което те преминават от МП1 до МП6. За всеки мониторингов пункт данните са комплектувани по дати и физикохимични параметри. Проведен е корелационен анализ между данните от отделните мониторингови пунктове. Стойностите на всички коефициенти на корелация (r_{ij}) надвишават 0.99 и за всички коефициенти на корелация коефициентите на Пирсън изпълняват условието за значимост $p_j < 0.05$, което позволява наличните измервания от шестте мониторингови пункта да бъдат обединени и използвани за целите на това изследване.

3.2. Статистически анализ на физикохимичните показатели, характеризиращи качеството на речните води

Проведен е статистически анализ на 124 комплекта данни от измервания в шестте мониторингови пункта. Резултатите за оценките на математическите очаквания (\bar{x}) и средноквадратичните отклонения (S_x) са представени в табл. 2.

Табл. 2. Оценки на математическите очаквания (\bar{x}) и средноквадратичните отклонения (S_x)

	NH ₄ -N	NO ₃ -N	NO ₂ -N	рН	BOD
	mg/l	mg/l	mg/l	-	mg/l
\bar{x}	0.105	0.802	0.020	8.176	2.343
S_x	0.126	0.318	0.015	0.384	0.845
	EC	TN	TOC	TP	
	µS/sm	mg/l	mg/l	mg/l	
\bar{x}	387.873	1.213	3.096	0.110	

S_x	74.137	0.522	0.860	0.116
	PO4-P	DO	T	COD
	mg/l	mg/l	°C	mg/l
\bar{x}	0.080	10.020	12.915	7.810
S_x	0.058	1.766	7.043	4.858

За проверка на наличието на статистическа зависимост между отделните физикохимични параметри, характеризиращи качеството на водата в река Струма, е проведен корелационен анализ, включващ по двойки всички физикохимични показатели. От изчислените 78 корелационни коефициента 23 са значими, което показва, че част от променливите, характеризиращи качеството на водата в река Струма, са взаимнокорелирани. Изследвания по тези въпроси са публикувани в [5].

В такива случаи е подходящо използването на метода на главните компоненти (Principal Component Analysis (PCA) за линейна трансформация на данните от измерванията в нов комплект от некорелирани данни със значително по-малък брой променливи, с минимална загуба на информация и позволяващи анализиране на тенденциите в подлежащия на мониторинг процес [11,12].

Наличието на статистическа зависимост между отделни показатели е определящо за избора на математическите методи, които ще бъдат използвани за разработване на мониторинговата система за качеството на водите на река Струма.

3.3. Анализ на измерванията от гледна точка на изискванията на Наредба 4

Подробният анализ на данните за река Струма в периода 2010-2019 г. по отношение на изискванията на Наредба 4 [1] за характеризирани на повърхностните води за полупланински типове реки (R4 и R5) (табл. 3) показва следното. Различните анализи във времето попадат в различни групи, постулирани в наредбата: „отлично“, „добро“ и „умерено“. Наблюдават се някои особености в качеството на речните води на река Струма в изследваното течение. Концентрацията на разтворения кислород DO попада преимуществено в първите две категории „отлично“ и „добро“. Електропроводността EC попада изцяло в категорията „отлично“. В наредбата за рН са зададени гранични стойности само за категорията „добро“. Направена е експертна категоризация на разполагаемите 70 комплекта измервания/анализи, съответстваща на изискванията на Наредба 4. В трите категории попадат, както следва – в „отлично“ – 15 комплекта данни, в „добро“ – 26 данни, и в „умерено“ – 29 данни.

Табл. 3. Категоризация на физикохимичните показатели съгласно Наредба 4

	DO	pH	EC	NH4-N
	mg/l	-	µS/sm	mg/l
Отлично	10.5-8.0	-	700	<0.04

Добро	8.0-6.0	6.5-8.5	750	0.04-0.4
Умерено	<6.0	-	>750	>0.4
	NO3-N	NO2-N	TN	PO4-P
	mg/l	mg/l	mg/l	mg/l
Отлично	<0.5	<0.01	<0.5	<0.02
Добро	0.5-1.5	0.01-0.03	0.5-1.5	0.02-0.04
Умерено	>1.5	>0.03	>1.5	>0.04
	TP	BOD₅		
	mg/l	mg/l		
Отлично	<0.025	<1.2		
Добро	0.025-0.075	1.2-3		
Умерено	>0.075	>3		

От анализите на проведените измервания/анализи може да бъде направен изводът, че при попадане в категорията „умерено“ е необходимо да бъде проведено детайлно изследване на физикохимичните параметри от конкретното измерване

4. Разработване на невронни класификатори

Първоначално е разработен класификатор за категоризиране на постъпващи измервания/анализи с участие на следните физикохимичните показатели: NH4-N, NO3-N, NO2-N, BOD, TN, TP, PO4-P, за които при различните измервания се наблюдава принадлежност и към трите категории. Разработен е класификатор, построен на базата на невронни мрежи с право разпространение на сигнала. За целите на класификатора на категорията „отлично“ е присвоена числена стойност 1, на категорията „добро“ – 2, и за „умерено“ – 3. Обучението се извършва на базата на 70 комплекта данни. Получените резултати от обучението на невронните мрежи са показани в табл. 4.

Табл. 4. Резултати от обучението на невронните мрежи

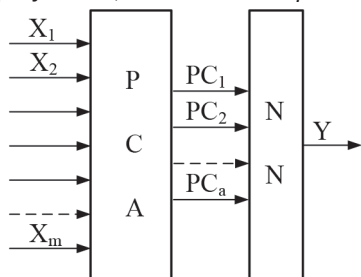
№	Структура на невронната мрежа	k	R
1	7-7-1	64	0.970
2	7-5-1	46	0.962
3	7-3-1	28	0.957

В табл. 4 първата цифра в структурата на невронната мрежа е броят на входните неврони, втората – броят на невроните в скрития слой, и третата е броят на изходните неврони. С k е означен броят на параметрите в невронната мрежа. Коефициентът на множествена корелация (R) се отнася за всички данни, в които са включени данните, въз основа на които е направено обучението, данните за верификация и тестовите данни

От таблицата се вижда, че директното включване на изследваните променливи като входове на невронната мрежа води до нарастване на броя на оценяваните параметри в мрежата, което от своя страна изисква значително по-голям брой данни от измервания/анализи.

5. Приложение на метода на главните компоненти за разработване на невронни класификатори

За преобразуване на корелираните променливи в независими или несвързани променливи и редуциране на размерността на задачата, в резултат на което да се включи по-малък брой променливи, следващите изследвания включват прилагане на метода на главните компоненти (Principal Component Analysis (PCA)). Схемата, по която се прави преобразуването, е показана на *фиг. 1*.



Фиг. 1. Схема с включване на метода на главните компоненти

На *фиг. 1* с X_1, X_2, \dots, X_m са означени физикохимичните променливи, а с PC_1, PC_2, \dots, PC_a – главните компоненти. С PCA е означен методът на главните компоненти, с NN – невронна мрежа, и с Y – изходната променлива на невронната мрежа, която в случая е категорията, към която се отнася конкретното измерване.

5.1. Разработване на невронни класификатори за широк спектър от измервания

Анализът на главните компоненти е направен за всички физикохимични показатели за качеството на речните води, като в допълнение е включена и температурата (T), която е влияещ фактор главно върху концентрацията на разтворения кислород (DO). По този начин предмет на разглеждане са следните променливи NH₄-N, NO₃-N, NO₂-N, pH, BOD, EC, TN, TP, PO₄-P, DO, T. Стойностите на тези променливи се разполагат по редове на матрицата X, (в термините на PCA), която е с размерност (n,m), като n съответства на броя на измерванията, а m – на броя на променливите. За конкретния случай n=70, а m=11.

Получени са следните собствени стойности на ковариационната матрица:

$\lambda_1=4.322, \lambda_2=1.565, \lambda_3=1.315, \lambda_4=0.9599, \lambda_5=0.7732, \lambda_6=0.6921, \lambda_7=0.5453, \lambda_8=0.4751, \lambda_9=0.1908, \lambda_{10}=0.1229$

Получените три главни компонента са показани в *табл. 5*.

Табл. 5. Стойности на главните компоненти PC1, PC2 и PC3

Променлива	PC1	PC2	PC3
NH ₄ -N	0.228	0.025	0.484
NO ₃ -N	0.393	-0.130	-0.151
NO ₂ -N	0.171	0.352	0.275
pH	0.045	0.165	-0.764

BOD	0.254	-0.373	-0.035
EC	0.253	0.312	0.046
TN	0.372	-0.328	0.009
TP	0.373	0.389	-0.001
PO ₄ -P	0.391	0.319	0.002
DO	0.335	-0.024	-0.274
T	-0.302	0.483	-0.071

Изразът за изчисляване на един от главните компоненти е следният:

$$(1) PC1 = 0.228 * NH_4 - N + 0.393 * NO_3 - N + 0.171 * NO_2 - N + 0.045 * pH + 0.254 * BOD + 0.253 * EC + 0.373 * TP + 0.391 * PO_4 - P + 0.335 * DO - 0.302 * T$$

Разработен е класификатор с невронна мрежа с право разпространение, в която PC1, PC2 и PC3 са независими променливи, а Y съдържа числено съответствие на категориите „отлично“, „добро“ и „умерено“.

Табл. 6. Невронен класификатор

No	Структура на невронната мрежа	k	R
1	3-5-1	26	0.879
2	3-3-1	16	0.845

Разработените до момента класификатори с невронни мрежи включват променливи, които по принцип могат да бъдат измервани непрекъснато (pH, EC, DO и T), и такива, които се измерват в резултат на лабораторни анализи, някои от тях отнемат време, измерващо се в дни. Както беше споменато по-горе пълният комплект измервания/анализи се извършва по един път на сезон за всеки един от мониторинговите пунктове (МП).

За да се разполага с актуална информация за състоянието на речните води, е необходимо да се осъществяват непрекъснати измервания.

Измерванията на стойностите на променливите, характеризиращи състоянието на водата, като пример *табл. 4*, се извършва с различни инструменти и методи и са с различна продължителност, като има някои, които са особено бавни като анализа на BOD, което се прави в продължение на пет дни (BOD₅). По тази причина е изследвана възможността да се организира процедура въз основа на променливите, които в световен мащаб се измерват непрекъснато с помощта на сензори, монтирани на място. Такива променливи са pH, EC, DO и T.

5.2. Разработване на невронни класификатори за променливи, подлежащи на непрекъснато измерване

За тази цел от разполагаемите 70 данни, които са класифицирани в категориите „отлично“, „добро“ и „умерено“ и съответстващата числова информация за категория 1, 2, и 3 се създава нов комплект данни, който включва pH, EC, DO, T и Y. И при това изследване се прилага подходът за преобразуване на стойностите на физическите промен-

ливи в главни компоненти. Стойностите на първите три главни компонента са представени в табл. 5.

Табл. 7. Главни компоненти за измерими променливи

№	PC1	PC2	PC3
pH	0.131	0.882	0.257
EC	0.436	0.044	-0.830
DO	0.691	0.150	0.137
T	-0.539	0.444	-0.475

Табл. 8. Структура на невронните мрежи

№	Структура на невронната мрежа	k	R
1	3-7-1	36	0.842
2	3-5-1	26	0.851

Резултатите от табл. 8 показват, че избраният подход за класификация на базата на потенциално непрекъснато измервани променливи (pH, EC, DO, T) е приложим.

6. Изводи и заключение

В работата са представени изследвания, необходим етап от разработване на мониторингова система за качеството на речните води на река Струма. Изследванията са подчинени на ограничените възможности за измерване/анализ на характеризиращите променливи за качеството на водите. Направен е анализ на проведените измервания в шест мониторингови станции по течението на реката в периода 2010-2019 г. Извършена е проверка на съвместимостта на данните от отделните мониторингови пунктове. Проведените изследвания са съобразени с Наредба 4 за характеризирание на повърхностните води. Разработени са и са изследвани невронни класификатори за седем и единадесет измервани променливи. Специално внимание е отделено на възможностите за характеризирание на качеството на водите посредством възможни за непрекъснато измервани променливи.

Благодарност

Тези изследвания са направени в рамките на Национална научна програма „Опазване на околната среда и намаляване на рисковете от неблагоприятни събития и природни бедствия“, одобрена с решение на Министерския съвет № 577/17.08.2018 г. и подкрепена от Министерството на образованието и науката (МОН) на България (Споразумение № Д01-363/17.12.2020 г.).

Литература

1. Наредба № Н-4 от 14.09.2012 г. за характеризирание на повърхностните води, издадена от министъра на околната среда и водите. Обн. ДВ, бр. 22 от 5.03.2013 г., в сила от 5.03.2013 г., изм. и доп., бр. 79 от 23.09.2014 г., в сила от 23.09.2014 г.
2. Западнореломорски район за басейново управление.
<https://wabd.bg/content/%D0%BF%D1%83%D1%80%D0%B1/%D0%BF%D1%83%D1%80%D0%B1-2010-2015-%D0%B3/>.

3. Мониторинг на водите в басейна на река Струма.
https://wabd.bg/docs/plans/ST/RBMP_STIV.pdf.
4. План за управление на речните басейни на западнореломорски район (2016-2021 г.), приет с решение № 1108/29.12.2016 г. на Министерския съвет.
<https://wabd.bg/content/%D0%BF%D1%83%D1%80%D0%B1/%D0%BF%D1%83%D1%80%D0%B1-2016-2021/>.
5. Boshnakov, K., B. Koumanova, M. Varbanov, E. Mihailov, S. Lavrova-Popova. Evaluation of Key Indicators Based on Automatically Measured Variables and Forecasting the Spread of Pollutants. – *Science, Engineering & Education*, 6, (1) 2021, 1-10 (under print).
6. Than, N. H., C. D. Ly, P. V. Tat. The Performance of Classification and Forecasting Dong Nai River Water Quality for Sustainable Water Resources Management Using Neural Network Techniques. – *Journal of Hydrology*, 596, 2021, 1-19.
7. Azhar, S. C., A. Z. Aris, M. K. Yusoff, M. F. Ramli, H. Juahir. Classification of River Water Quality Using Multivariate Analysis. *Procedia Environmental Sciences*, 30, 2015, 79-84.
8. Wechmongkhonkon, S, N. Poomtong, S. Areerachakul. Application of Artificial Neural Network to Classification Surface Water Quality. – *International Journal of Environmental and Ecological Engineering*, 6, 2012, No. 9, 574-578.
9. Nasir, M. F., M. S. Samsudin, I. Mohamad, M. R. A. Awaluddin, M. A. Mansor, H. Juahir, N. Ramli. River Water Quality Modeling Using Combined Principle Component Analysis (PCA) and Multiple Linear Regressions (MLR): A Case Study at Klang River, Malaysia. *World Applied Sciences Journal*, 14 (Exploring Pathways to Sustainable Living in Malaysia: Solving the Current Environmental Issues), 2011, 73-82.
10. Isiyaka, H. A, A. Mustapha, H. Juahir & P. Phil-Eze. Water Quality Modelling Using Artificial Neural Network and Multivariate Statistical Techniques. *Model. Earth Syst. Environ.*, 2019, 5:583-593.
11. Бошнаков, К., Л. Фан. Мониторинг на процесите при пречистване на отпадъчни води. Международна конференция „Автоматика и информатика“ 2010, София, България, 3-7 октомври 2010, III-455-459.
12. Jolliffe, I. T. *Principal Component Analysis*, Second Edition. http://www.ru.ac.bd/stat/wp-content/uploads/sites/25/2019/03/503_08_Jolliffe_Principal-Component-Analysis.pdf.

За контакти:

Проф. д-р **Коста Бошнаков**
Катедра „Автоматизация на производството“ – ХТМУ
e-mail: kb@uctm.edu

Проф. д-р **Богдана Куманова**
Катедра „Инженерна химия“ – ХТМУ

Доц. д-р **Мариан Върбанов**
Департамент „География“

Национален институт по геофизика,
геодезия и география – БАН

e-mail: marian_varbanov@geophys.bas.bg

Проф. д-р **Емил Михайлов**
Катедра „Физична металургия и топлинни агрегати“ – ХТМУ