

СТАНОВИЩЕ

върху дисертационен труд за придобиване на образователна и научна степен „Доктор“

Автор на дисертационния труд: **маг. инж. Николина Йорданова Иванова**

Тема на дисертационния труд: **ИЗСЛЕДВАНЕ НА ПОДХОДИ ЗА УНИФИЦИРАНЕ НА ГОЛЕМИ ДАННИ ОТ РАЗЛИЧНИ ЕКОЛОГИЧНИ ИЗТОЧНИЦИ С ЦЕЛ УСЪВЪРШЕНСТВАНЕ НА ТЕХНИЯ АНАЛИЗ**

Член на научното жури: **доц. д-р инж. Йорданка Найденова Анастасова**

1. Актуалност на разработвания в дисертационния труд проблем в научно и научно приложно-отношение.

Унифицирането и обработката на големи обеми от хетерогенни данни, постъпващи от различни източници, е особено актуално както в научната, така и в научно-приложната тематика. Интеграцията и стандартизацията на данните осигуряват възможност за прецизни и надеждни анализи, особено когато това се отнася за реални индустриални данни.

Целта на дисертационния труд е разработване и прилагане на методологичен подход за предварителна обработка, унифициране и анализ на големи обеми хетерогенни данни от различни източници, с акцент върху оперативни данни от пречиствателни станции за отпадни води и метеорологични наблюдения. За реализация на целта са дефинирани и изпълнени пет основни задачи, за чието решаване са използвани съвременни подходи и технологии.

Направеното изследване, представено в дисертационния труд, цели разкриването на статистически зависимости между технологични и климатични фактори, които ще подпомогнат управлението на пречиствателни процеси.

2. Степен на познаване състоянието на проблема и творческа интерпретация на литературния материал.

В дисертационния труд авторът показва много добро познаване на тематиката и адекватна интерпретация на литературния материал. Това личи ясно от направеният в първа глава на дисертацията анализ на основните предизвикателства при обработката и анализа на Big Data. Представени са съвременните аналитични подходи, техники и методи, приложими при разработка и анализ на разнородни и комплексни масиви от данни. В резултат ясно е дефинирана целта и конкретизирани задачите в дисертационния труд.

Дисертационният труд е в обем от 145 страници и включва 26 фигури в основния текст, още 18 в двете приложения и 28 таблици. Цитирани са 97 библиографски източника, от които само 1 на български и 96 на английски, които в преобладаващата си част са достъпни и онлайн. Преобладаваща част от източниците са публикувани в последните години, което показва задълбочено познаване на съвременните тенденции в областта.

3. Съответствие на избраната методика на изследване и поставената цел и задачи на дисертационния труд с постигнатите приноси.

Преди избора на методика авторът подробно описва източниците на данни, като е анализирана структурата им, честота, характеристики на наблюдавани показатели и проблемите при дигитализацията и предварителната им обработка. Сравнени са различията в структурата на данните от пречиствателна и метеорологична станции и е осигурена съвместимостта им.

Освен това авторът разглежда подробно методите за запълване на липсващи данни, като са описани механизмите, водещи до появата им и е направена класификация на подходите за справянето с проблема, а именно изтриване, импутация, моделиране и Extract Transform Load процеси. Избрани са и приложени три основни метода за запълване на липсващи стойности - PCHIP, SPLINE и AR и е направена количествена оценка на ефективността им, съобразена с характеристиките на данните и природата им.

Като най-подходящ е избран AR (Autoregressive Model), поради факта, че улавя тенденциите и гарантира висока точност при данни с висока променливост и сезонни колебания, каквито са тези от ПСОВ и метеорологичната станция.

Избраната методика на изследване е приложена за анализ на взаимодействията между хидрохимични и метеорологични параметри, като са обединени и синхронизирани реални данни от ПСОВ и метеорологична станция гр. Кърджали. Приложени са корелационен, крос-корелационен и факторен анализ, които оценяват зависимостите, показващи влиянието на атмосферните условия върху качеството на отпадните води, като са установени три основни фактора – Температурно-химичен (Ф1), Атмосферно-метеорологичен (Ф2) и Хидравличен (Ф3).

Формулирани и проверени са 4 хипотези, като:

Хипотеза 1 потвърждава доминиращата роля на Ф1 при замърсяване и ефективност при пречистването;

Хипотеза 2 потвърждава че Ф2 влияе на сезонната чувствителност на системата;

Хипотеза 3 потвърждава, че Ф3 влияе напълно върху концентрацията на азот и условно за фосфор в отпадните води;

Хипотеза 4 потвърждава, че при внимателна и обоснована селекция на индикатори (намаляването им) се постига опростяване на процеса, като дори се повишава интерпретируемостта на структурата.

Проведеното и представено експериментално изследване ползва интегриран подход, включващ унифициране, факторен анализ и проверка на хипотези, като за целта са ползвани над 268 000 данни и получените резултати са визуализирани и графично.

Считам, че избраната методика на изследване и реализация напълно съответстват на поставената в дисертацията цел и формулираните задачи, като предлагат основа за екологични анализи и разработка на прогнозни модели за управление на водни и климатични ресурси и процеси.

4. Научни и/или научно-приложни приноси на дисертационния труд.

В дисертационния труд авторът е формулирал следните приноси, които приемам и в резюме са както следва:

Научни:

1. Предложени са допълнителни критерии за усъвършенстване на интегрираната методология за унифициране на хетерогенни екологични и метеорологични данни, включваща синхронизация на измерванията по дати, стандартизация на променливите и запълване на липсващите стойности чрез специализирани импутационни методи, като е приложена успешно с реални оперативна данни от ПСОВ и метеорологична станция.
2. Адаптирани са статистически подходи за извършване на факторен анализ в контекста на екологични системи чрез използване метода на главните компоненти (PCA) и Varimax ротация. Извлечени са температурно-химичен, атмосферно-метеорологичен и хидравличен фактори, които са ясно интерпретируеми и разкриват дълбочинно залегалата структура в сложни екологични данни.
3. Разработена е експериментална процедура за сравнителна оценка на методи за запълване на липсващи стойности, основана на зануляване на избрани интервали с известни стойности от данните и последващо възстановяване чрез PCHIP, SPLINE и авторегресивен модел. Точността на методите е оценена с метриците MAE, RMSE и MAPE, което осигурява обективен критерий за избор на подход при импутация.

Научно-приложни:

1. Валидирана е хипотеза за съществуването на отложени (лагови) зависимости между атмосферни и водни показатели, чрез прилагане на крос-корелационен анализ, като са установени значими зависимости със закъснение от няколко дни, което обогатява

разбирането на динамиката в пречиствателните системи и има потенциал за разработване на прогнозни модели.

2. Създадена е аналитична база от зависимости между климатични и технологични променливи, която позволява по-задълбочено моделиране поведението на пречиствателната система при различни атмосферни условия и оперативни натоварвания.
3. Предложен е подход за оптимизация на мониторинговата мрежа в ПСОВ, чрез обобщение на факторни товари и анализ на обяснена дисперсия. Извлечената редукция е съставена от 15 променливи, които запазват 77,36% от общата обяснена дисперсия, което позволява значително намаляване на броя на наблюдаваните показатели без съществена загуба на информация и създава предпоставки за оптимизация на измервателните мрежи в ПСОВ.
4. Поставена е основа за прилагане на прогнозни модели, базирани на факторни оценки и атмосферни данни, което би подпомогнало оперативното управление и би повишило ефективността на пречиствателните процеси чрез по-информирано вземане на решения.
5. Предложени са оригинални методи за обработка и анализ на Big Data в контекста на ПСОВ, чрез съчетаване на техники от статистическия анализ и машинното обучение, което разширява съществуващите изследвания и дава основа за разработване на практически приложими алгоритми за анализ на реални данни от непрекъснат мониторинг.
6. Систематизирана и апробирана е методология за интегриран анализ на екологични и технологични данни, която обединява процесите на унифициране, редукция на размерността чрез факторен анализ и проверка на хипотези. Тази рамка е приложима за широк спектър от изследвания и показва как съчетаването на класически статистически техники с Big Data подходи може да доведе до значими интердисциплинарни резултати.

5. Преценка на публикациите по дисертационния труд: брой, характер на изданията, в които са отпечатани.

Резултатите от изследванията на инж. Николина Иванова са публикувани в четири статии. Всички статии са представени на международни конференции, които са в България. Публикациите са направени в периода 2021-2024 година и покриват тематиката на представената дисертационна работа, като отразяват постигнатите резултати и приноси.

Една от статиите е самостоятелна, три са в съавторство, като инж. Иванова е първи автор в две от тях. Всички статии са индексирани в Google Scholar и CrossRef и са публикувани в периодични издания, а една от тях е индексирана в Scopus. Към настоящия момент са открити 3 цитирания на две от представените статии. Публикационната дейност на дисертанта отговаря на изискванията за придобиване на ОНС „Ооктор“ и считам, че публикациите са придобили достатъчно публичност в научната общност.

Отчитайки броя и качеството на публикациите считам, че получените резултати са изключително дело на автора и са постигнати под ръководството на научния му ръководител, като не съм установила недостоварност на данните или плагиатство.

6. Мнения, препоръки и бележки.

Дисертационният труд е изготвен в стройна логическа последователност. Коректно са цитирани използваните източници. Описани са подробно използваните съкращения в дисертационния труд. Фигурите и таблиците са цитирани акуратно и допълват много уместно текстовото изложение.

Препоръчвам на докторантката за в бъдеще да публикува получените резултати от работата си и последващите ѝ изследванията в издания, които се индексират в Scopus и WoS.

7. Заключение с ясна положителна или отрицателна оценка на дисертационния труд.

Оценявам **положително** дисертационния труд „**Изследване на подходи за унифициране на големи данни от различни екологични източници с цел усъвършенстване на техния анализ**“ с автор маг. инж. **Николина Йорданова Иванова** и считам, че е изготвен прецизно, акуратно и на

високо професионално ниво, като авторът е постигнал поставените цели и задачи в дисертационния труд. Резултатите и приносите на дисертационния труд са отразени в научни публикации, които са представени на научната общност на международни конференции и публикувани и индексирани. Научните и научно-приложните приноси са значими и съответстват на изискванията за придобиване на научно-образователната степен „Доктор“ съгласно Закона за развитие на академичния състав в Република България, Правилника за приложението му и Правила и процедури за приемане и обучение на докторанти и придобиване на ОНС „Доктор“ и НС „Доктор на науките“ на Минно-геоложки университет „Св. Иван Рилски“, София.

Предлагам на уважаемото Научно жури да присъди образователна и научна степен „Доктор“ на маг. инж. Николина Йорданова Иванова в професионално направление 5.13. Общо инженерство, докторска програма „Компютърни технологии в инженерната дейност“.

Дата: 07.08.2025 г.

ЧЛЕН НА ЖУРИТО:

(Доц. д-р инж. Йорданка Анастасова)