

РЕЦЕНЗИЯ

от проф. дмн. Христо Илиев Семерджиев

(ФМИ при ПУ "Паисий Хилендарски")

върху дисертационния труд "Статистическо моделиране на качеството на въздуха"

с автор Атанас Вълев Иванов

за присъждане на образователната и научна степен „доктор“

в област на висше образование 4. Природни науки, математика и информатика

професионално направление 4.5 Математика

докторска програма Математическо моделиране и приложение на математиката.

Научни ръководители:

1) проф. д-р Снежана Георгиева Гочева-Илиева, ПУ „Паисий Хилендарски“

2) доц. д-р Дойчин Тодоров Бояджиев, ПУ „Паисий Хилендарски“

1. Общо описание на представените материали

Със заповед № Р33-466 от 03.02.2015 г. на Ректора на Пловдивския университет „Паисий Хилендарски“ съм определен за член на жури за осигуряване на процедура за защита на дисертационен труд на тема „Статистическо моделиране на качеството на въздуха“ за придобиване на образователна и научна степен „доктор“ в област 4. Природни науки, математика и информатика, професионално направление 4.5 Математика.

Представеният от Атанас Вълев Иванов комплект материали е в съответствие с Чл.36 (1) от Правилника за развитие на академичния състав на ПУ и включва следните документи:

- Молба до Ректора на ПУ за разкриване на процедурата за защита на дисертационен труд;
- Автобиография в европейски формат;
- Копие от диплома за висше образование ОКС „магистър“;
- Заповед за записване в редовна докторантура;
- Заповед за определяне на темата на дисертационен труд;
- Заповед за провеждане на изпит от индивидуалния план и съответен протокол за издържан изпит по специалността;
- Заповед за смяна името на темата на дисертационен труд;
- Заповед за отчисляване от докторантура с право на защита;
- Протоколи от катедрени съвети, свързани с докладване на готовност за откриване на процедурата и с предварително обсъждане на дисертационния труд;
- Списък на научните публикации по темата на дисертацията;
- Дисертационен труд;
- Автореферат;
- Копия от научните публикации;
- Декларация за оригиналност и достоверност на приложените документи;
- Справка за спазване на специфичните изисквания на съответния факултет;

Докторантът е приложил 4 броя публикации.

2. Кратки биографични данни за докторанта

Атанас Вълев Иванов има бакалаврска и магистърска степен по Математика и Приложна математика, придобити в ПУ „Пайсии Хилендарски“. От 2011 г. той е редовен докторант към катедра Приложна математика и моделиране. По време на докторантурата той задълбочава интереса си в областта на математическата статистика и нейните приложения, резултат на което е представения дисертационен труд. Владее добре писмено и говоримо английски език.

3. Актуалност на тематиката и целесъобразност на поставените задачи

В дисертацията се прилагат съвременни статистически методи за прогнозиране и изследване на връзката на основни замърсители на въздуха в населените места в България за идентифициране комплексни зависимости и получаване на краткосрочни прогнози за превенция и поддържане на качеството на въздуха. Тези задачи са извънредно актуални и важни, тъй като са пряко свързани със здравето на хората.

4. Познаване на проблема

Докторантът показва отлично познаване на състоянието на проблема и творчески подход при избора на статистически методи за решаване на поставените задачи. Ако се съди по получените резултати може да се направи извода, че тези методи съответстват в максимална степен на характера на задачите.

5. Методика на изследването

В изследването са използвани данни от автоматизирани станции. Методите и моделите ARIMA/SARIMA и факторния анализ са добре познати и са реализирани в повечето популярни софтуерни системи и Методът GPSRR (обобщена PathSeeker регуляризираща регресия) на Salford Systems, което е новост в подобни изследвания.

6. Характеристика и оценка на дисертационния труд

Дисертационният труд (ДТ) е структуриран в увод, 4 глави, заключение, 134 печатни страници текст, списък от 115 информационни източника, 61 фигури и 23 таблици.

ДТ е посветен на важния негативен проблем замърсяването на въздуха. Той се дължи на големите производствени мощности и химически производства без обезпечаване на пречистване на отпадъците, засиления трафик на неекологични превозни средства, неконтролируемо отделяне на отпадъци от ежедневната човешка дейност. Замърсяването на околната среда води до сериозни климатични промени, съпроводени от глобални резки температурни скокове. Освен климатичните изменения, замърсяването на въздуха, почвите и водите водят до увеличаването на тежки заболявания. В големите страни като САЩ, ЕС, Япония и др. са установени системи за постоянен мониторинг и контрол на нивата на замърсяване. На проблемите със замърсяване на околната среда, в частност, замърсяване на въздуха са посветени множество научни изследвания. ДТ е посветен на статистическо моделиране на реални данни за замърсяване на въздуха в българските градове Кърджали, Благоевград и Шумен, получени с помощта на съответните автоматизирани станции на Изпълнителната агенция за околната среда.

В глава 1, озаглавена Въведение, е направен литературен обзор на монографии, периодика, публикации на научни конференции, посветени на изследването на чистотата на въздуха, като е обърнато по-голямо внимание на трудовете, свързани с рецензирания ДТ. В цитираната литература често използвани подходи за статистическо

моделиране в науките за околната среда са методите на многомерния линейен и нелинеен анализ – регресия, метод на главните елементи, факторен анализ (ФА), клъстерен анализ. Докторантът е запознат с резултатите от изследванията на данни от замърсявания с прахови частици в множество градове в Южен Онтарио, град Мумбай, Индия, Атина, Бърно, Малайзия, Китай, Полша за различни периоди от време. В някои от посочените изследвания са направени предсказания за няколко дни напред. Данните за чистотата на въздуха са представени във вид на времеви редове. Тъй като в самите наблюдения за даден замърсител, освен времето, се съдържа информация за породилите го условия в околната среда, то характерът на данните се определя като стохастичен. В ДТ се използва методологията на Бокс-Дженкинс, която е известна като ARIMA (Auto Regressive Integrated Moving Average). Тази методология е доразвита и за случая на отчитане на сезонност (Seasonal ARIMA) или SARIMA. Известно е мнението на Milionis и Davies, че при анализиране на регресивните и стохастични модели за замърсяване на въздуха стохастичните модели са за предпочитане пред регресивните. От многобройните специализирани софтуерни пакети по математика и по статистика, които могат да се използват за анализ на времеви редов в ДТ се използват главно SPSS IBM и Wolfram Mathematica. В §1.1.3 се засягат проблемите на математическото моделиране за замърсяването на въздуха и на екологични процеси. От една страна, в България има много населени места със силно влошено качество на въздуха. От друга страна, мониторинговите автоматизирани станции са натрупали огромна база данни за голям брой замърсители. Докторантът счита, че основната цел на ДТ е създаване, изследване и приложение на стохастични модели на основни замърсители на въздуха в населените места в България за идентифициране на комплексни замърсявания от общи източници, откриване на зависимости, получаване на краткосрочни прогнози за превенция и поддържане на качеството на въздуха. За целта дисертантът разглежда замърсявания с 6 основни замърсители: прахови частици (PM10), серен диоксид (SO2), азотен диоксид (NO2), азотен оксид (NO), озон (O3) и азотни оксиди (NOx). Възникват следните задачи:

1. С помощта на стохастичен сезонен SARIMA метод да се построят и изследват модели на замърсяване на въздуха с PM10 и SO2 на град Кърджали.
2. За всичките 6 основни замърсителя на град Шумен да се проведе времеви анализ чрез построяване и приложение на стохастични ARIMA модели като за всичките тези замърсители да се намери групиране чрез многомерен ФА.
3. Същото да се направи за шестте замърсителя на въздуха в град Благоевград. Тези модели да се приложат за краткосрочни прогнози.

В ДТ се използват следните методи за моделиране:

1. Трансформация на Йео-Джонсън

$$\psi_{YJ}(\lambda, x) = \begin{cases} \{(x+1)^\lambda - 1\} / \lambda & x \geq 0, \lambda \neq 0 \\ \log(x+1) & x \geq 0, \lambda = 0 \\ -\{(-x+1)^{2-\lambda} - 1\} / (2-\lambda) & x < 0, \lambda \neq 2 \\ -\log(-x+1) & x < 0, \lambda = 2 \end{cases}, \quad \lambda \in [-2, 2]$$

За определяне на параметъра λ дисертантът е разработил съответен код на системата Wolfram Mathematica за променливата x . При подбор на оптималната стойност на λ се сравнява с оптималната стойност на показателя на Колмогоров-Смирнов за нормалното разпределение на данните

2. ФА, който е широко известна статистическа техника и позволява редуцирането на броя на независимите изходни променливи чрез групиране на силно корелиращи се помежду си в общ фактор.
3. Стохастични ARIMA и SARIMA методи. Едномерния ARIMA (авторегресионен анализ и плаващи средни) е название на обобщен клас от методи, въведени от Бокс и Дженкинс. ARIMA широко се прилага в екологичната статистика за оценка и предсказване поведението на времеви редове.

Глава 1 няма приносен характер.

Глава 2 е озаглавена Приложение на стохастичен времеви анализ и факторен анализ. В тази глава се съдържат две емпирични изследвания на чистотата на въздуха. Проведен е анализ на замърсявания с прахови частици (PM10) и серен диоксид (SO₂) в зависимост от времето чрез SARIMA модели. По-нататък са построени, изследвани и приложени за предсказване модели за 6-те въздушни замърсители с ФА и ARIMA с използване на метеорологични параметри.

§2.1. Тук е направен времеви статистически анализ на замърсяването на въздуха за град Кърджали като се акцентира на основните замърсители, влияещи на човешкото здраве - прахови частици (PM10) и серен диоксид (SO₂). За целта са използвани резултати на почасови данни от мониторинга за 4 последователни години от 2008 до 2011г.. Установено е, че не съществува корелационна връзка между двата замърсителя. Изследването е проведено със статистическия софтуер SPSS. Числовите резултати са табулирани.

§2.2. Проведен е анализ на времеви редове и предсказване на замърсяването на въздуха за същите 6 замърсителя с ФА и ARIMA с използване на метеорологичните параметри – атмосферно налягане, посока на вятъра, приземна температура на въздуха, слънчева радиация, относителна влажност и скорост на вятъра. Тук е представено статистическо изследване на замърсяването на въздуха в средно големия град в северна България Шумен. Времевите редове се базират на почасови данни за 2011г. и първото тримесечие на 2012г.. Построените модели са приложени за предсказване на измерените стойности и за краткосрочно прогнозиране. ФА е приложен за всичките 6 изследвани замърсителя. Процедурата се състои от последователно прилагане на следващите стъпки: изчисляване на корелационната матрица, тест за адекватност на ФА, извличане на фактори, въртене на факторите и изчисляване на стойностите на факторните променливи. За съжаление, рецензентът не откри в ДТ корелационната матрица и таблица 2.10.

В **глава 3**, озаглавена Времеви анализ на ежедневните замърсители на въздуха в град Благоевград, се изследват времеви редове за замърсяването на въздуха на 6-те въздушни замърсителя в град Благоевград с приложение на факторен анализ и SARIMA метод.

В **глава 4**, озаглавена Приложение на SARIMA и GPS регресия за изследване замърсяване на въздуха с PM10 на град Шумен, се изследват данни за концентрациите на фини прахови частици PM10 за град Шумен. Концентрациите превишават многократно допустимите норми. Целта е да се построят съответни емпирични статистически модели с отчитане на метеорологичните параметри на базата на два различни подхода - стохастичен SARIMA и General Pathseeker регресия, да се сравни способността им да предсказват реалните нива на замърсяване с PM10, както и приложимостта им за краткосрочно прогнозиране.

7. Приноси и значимост на разработката за науката и практиката

Дисертацията има научно-приложен характер. В нея е предложен нов подход към анализа на данни и моделирането на зависимости. При този подход са използвани най-новите достижения на статистическата наука, като идеите са реализирани прецизно и

умело. От друга страна, получените модели могат да имат важни приложения в екологията.

8. Преценка на публикациите по дисертационния труд

По представения дисертационен труд са публикувани 4 статии. Една от тях е в рефирирано списание с висок импакт-фактор (Stochastic Environmental Research and Risk Assessment), две са в сборниците на Американския институт по физика, а една в трудове на IX Национална научно-техническа конференция с международно участие "ЕКОЛОГИЯ И ЗДРАВЕ 2012", три от тях са докладвани на конференции. Всичките са написани в съавторство като приемам, че приносът на авторите е равностоен.

9. Лично участие на докторанта

След прочитането на дисертацията и проведената предварителна защита считам, че основната работа по построяването и анализа на моделите е извършена самостоятелно.

10. Автореферат

Авторефератът представя адекватно на ДТ основните резултати и достижения. Написан е сбито, точно и ясно в пълно съответствие с формалните изисквания.

11. Критични забележки и препоръки

В дисертацията е използван методът GPSRR (обобщена PathSeeker регуляризираща регресия) на Salford Systems, който не се е наложил още в тази област. Въпреки това намирам, че идеята е оригинална и показва отлични резултати. Тази моя забележка може да бъде разглеждана като препоръчка за бъдещи работи и налагане методът в областта.

12. Лични впечатления

Докторантът Атанас Иванов защити блестящо дипломна работа за получаване на бакалавърска степен под ръководството на рецензента върху трудната тема от аналитичната механика третираща механични системи с променлива маса, но тази работа не намери приложение. Рецензираният ДТ има хубави научно-приложни приноси. Самият асистент Атанас Иванов е находка за катедрата ПММ на ПУ и за неговите научни ръководители проф. д-р Снежана Гочева-Илиева и доц. д-р Дойчин Бояджиев, на които аз като бивш ръководител на катедрата ПММ, изказвам дълбоката си благодарност. Разбира се, докторантът е изпитал силно благотворно научно влияние (това е истинският фактор на влияние) на научните си ръководители. Той е имал и удачата да се снабди с реални данни от автоматизираните станции на Кърджали и Шумен.

13. Препоръки за бъдещо използване на дисертационните приноси и резултати

Препоръката е да се продължи използването на методите за тяхното подобряване и приложение в областта на екологията. В хода на рецензирането бяха открити някои печатни грешки, които са съобщени на докторанта.

ЗАКЛЮЧЕНИЕ

След всичко казано дотук следва, че дисертационният труд е с ясно очертана научна тематика. Получени са достатъчно на брой важни научно-приложни приноси. Дисертационният труд отговаря на изискванията на ЗРАСРБ и на ППЗРАСРБ за придобиване на образователна и научна степен „доктор“ в област 4. Природни науки, математика и информатика, професионално направление 4.5 Математика, докторска програма Математическо моделиране и приложение на математиката, отговаря и на специфичните изисквания от Правилника на ПУ. Поради това моята оценка на дисертационния труд на ас. Атанас Вълев Иванов **“СТАТИСТИЧЕСКО МОДЕЛИРАНЕ НА КАЧЕСТВОТО НА ВЪЗДУХА”** е положителна. Предлагам на уважаемите членове на научното жури да гласуват за присъждането на образователна и научна степен „доктор“ на Атанас Вълев Иванов.

21.02.2015 г.

Репензент

гр. Пловдив

проф. дмн Христо Семерджиев