

**БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра теории вероятностей и математической статистики**

Аннотация к магистерской диссертации

**ПРОЕКТИРОВАНИЕ СИСТЕМЫ СБОРА И АНАЛИЗА ДАННЫХ,
ПОЛУЧЕННЫХ С ПОМОЩЬЮ ЭЛЕКТРОННЫХ СРЕДСТВ
ИЗМЕРЕНИЯ**

Герасименя Марина Владимировна

Научный руководитель – кандидат технических наук,
доцент Ж.В. Василенко

Минск, 2017

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Перечень ключевых слов: ВРЕМЕННОЙ РЯД, ПРОГНОЗИРОВАНИЕ, ЯЗЫК R, ОПИСАТЕЛЬНЫЕ СТАТИСТИКИ, КОРРЕЛЯЦИОННЫЙ АНАЛИЗ, РЕГРЕССИОННЫЙ АНАЛИЗ, МНОЖЕСТВЕННАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ, МЕТОД ГЛАВНЫХ КОМПОНЕНТ, НЕЙРОННЫЕ СЕТИ.

Цель работы: с помощью современного языка программирования R осуществить анализ, обработку и прогнозирование реальных временных рядов. Необходимость подобного исследования обусловлена тем, что в настоящее время объем накопленных наблюдений за экологическим состоянием окружающей среды резко увеличился. Современный математический аппарат и его прикладные части позволяют не только проанализировать сложившуюся ситуацию, но и дать прогноз по состоянию объекта в будущем.

Объектом исследования в работе являются еженедельные наблюдения за загрязнителями воздуха в разных частях города и за его пределами в период с января 2012 года по декабрь 2016 года. В процессе работы полученные данные были проанализированы с помощью описательных статистик и графического анализа, был произведен корреляционный и регрессионный анализы, вычисление прогнозных значений временного ряда с помощью множественной линейной регрессии и нейронных сетей.

Полученные результаты могут быть использованы для дальнейшего исследования организациями по мониторингу окружающей среды.

Структура магистерской диссертации:

В главе 1 приводятся основные понятия, связанные с загрязнением воздуха, описаны основные вредные вещества и их влияние на человека и окружающую среду. В главе 2 исследуются основные виды мониторинга за окружающей средой и их общие свойства, приведена общая схема работы системы, разработан первоначальный дизайн для веб-системы. В главе 3 проведен последовательный анализ данных по загрязнению и рассмотрены на практике методы регрессионного анализа и нейронных сетей.

Магистерская диссертация состоит из 52 страниц, 18 рисунков, 9 таблиц, 28 источников, 2 приложения.

АГУЛЬНАЯ ХАРАКТЕРЫСТИКА РАБОТЫ

Спіс ключавых слоў: ЧАСОВЫ ШЭРАГ, ПРАГНАЗАВАННЕ, МОВА R, АПИСАЛЬНАЯ СТАТЫСТЫКІ, КАРЭЛЯЦЫЙНЫЯ АНАЛІЗ, РЭГРЭСІЙНЫ АНАЛІЗ, МНОЖНАЯ ЛІНЕЙНАЯ РЭГРЭСІЯ, МЕТАД ГАЛОЎНЫХ КАМΠΑНАЕНТ, НЕЙРОНАВЫЯ СЕТКІ.

Мэта дысэртацыі: з дапамогай сучаснай мовы праграмавання R ажыццявіць аналіз, апрацоўку і прагназаванне рэальных часовых шэрагаў. Неабходнасць падобнага даследавання абумоўлена тым, што ў цяперашні час аб'ём назапашаных назіранняў за экалагічным станам навакольнага асяроддзя рэзка павялічыўся. Сучасны матэматычны апарат і яго прыкладныя часткі дазваляюць не толькі прааналізаваць сітуацыю, якая склалася, але і даць прагноз па стане аб'екта ў будучыні.

Аб'ектам даследавання ў працы з'яўляюцца штотыднёвыя назірання за забруджвальнікамі паветра ў розных частках горада і за яго межамі ў перыяд са студзеня 2012 года па снежань 2016 года. У працэсе работы атрыманыя дадзеныя былі прааналізаваны з дапамогай апісальных статыстык і графічнага аналізу, быў выраблены карэляцыйныя і рэгрэсійныя аналізы, вылічэнне прагнозных значэнняў часовага шэрагу з дапамогай множнай лінейнай рэгрэсіі і нейронавых сетак.

Атрыманыя вынікі могуць быць выкарыстаны для далейшага даследавання арганізацыямі па маніторынгу навакольнага асяроддзя.

Структура магістарскай дысэртацыі:

У раздзеле 1 прыводзяцца асноўныя паняцці, звязаныя з забруджваннем паветра, апісаны асноўныя шкодныя рэчывы і іх уплыў на чалавека і навакольнае асяроддзе. У раздзеле 2 даследуюцца асноўныя віды маніторынгу за навакольным асяроддзем і іх агульныя ўласцівасці, прыведзена агульная схема працы сістэмы, распрацаваны першапачатковы дызайн для вэб-сістэмы. У раздзеле 3 праведзены паслядоўны аналіз дадзеных па забруджванні і разгледжаны на практыцы метады рэгрэсійнага аналізу і нейронавых сетак.

Магістарская дысэртацыя мае 52 старонкі, 18 малюнкаў, 9 табліц, 28 крыніц, 2 прыкладанні.

BRIEF DESCRIPTION OF THE THESIS

The list of keywords: TIME SERIES, FORECASTING, LANGUAGE R, DESCRIPTIVE STATISTICS, CORRELATION ANALYSIS, REGRESSION ANALYSIS, MULTIPLE LINEAR REGRESSION, PRINCIPAL COMPONENT ANALYSIS, NEURAL NETWORKS.

The purpose of the work: to analyze, process and predict real time series using modern programming language R. Such a study is of current interest due to the fact that the volume of accumulated observations of the ecological environmental state has increased dramatically in the latest days. The modern mathematical apparatus and its applied parts not only allow to analyze the current situation, but also to forecast the state of the object in the future.

The object of research is weekly monitoring of air pollution in different parts of the city and outside of it starting from January 2012 till December 2016. In the process of work, the data were analyzed using descriptive statistics and graphical analysis, correlation and regression analysis were performed, predictive values of the time series were computed using multiple linear regression and neural networks.

The results obtained can be used for further research by environmental monitoring organizations.

Structure of the master's thesis:

Chapter 1 introduces the basic concepts associated with air pollution, describes the main harmful substances and their effects on humans and the environment. In Chapter 2, the main types of environmental monitoring and their general properties are explored, a general scheme for the operation of the system is provided and initial design for the web system is created. In Chapter 3, a sequential analysis of contamination data is carried out and methods of regression analysis and neural networks are considered in practice.

Master's thesis consists of 52 pages, 18 figures, 9 tables, 28 sources, 2 appendices.