

МОДЕЛИРАНЕ НА ЗАМЪРСЯВАНЕТО С ФПЧ10 НА АТМОСФЕРНИЯ ВЪЗДУХ В ГРАД ПЛОВДИВ ЧРЕЗ РЕГРЕСИОННИ ДЪРВЕТА

Десислава Войникова

Резюме. Чистотата на атмосферния въздух е един от факторите за човешкото здраве, а замърсяването му с фини прахови частици с диаметър под 10 микрона (ФПЧ10) е проблем на съвремето. В това изследване е приложен методът на Класификационни и регресионните дървета (CART) и са построени модели за среднодневните данни за замърсяването на въздуха с PM10 в град Пловдив за период над 12 години от 01.01.2011 г. до 31.01.2023 г. Най-добрият получен модел, описващ 90% от данните, е използван за предсказване на концентрациите един ден напред, което е приложено 7 пъти. Разгледани са различни модели като са използвани в комбинации метеорологични данни, времеви данни, лагирани променливи, трансформирани данни и кросвалидация. Построени са и контингентни таблици, които показват, че CART моделите предсказват правилно до 94% от измерените стойности на PM10 по отношение на европейската среднодневна допустима прагова стойност за защита на човешкото здраве (Директива 2008/50/ЕО на Европейския парламент) от 50 микрограма на кубичен метър.

Ключови думи: Класификационни и регресионни дървета, фини прахови частици с диаметър до 10 микрона (ФПЧ10), замърсяване на атмосферния въздух, предсказване на замърсяване на атмосферния въздух.

Десислава Войникова

Пловдивски университет „Паисий Хилендарски“

Факултет по математика и информатика,

ул. „Цар Асен“ 24, Пловдив, България

Автор за кореспонденция: desi_voynikova@uni-plovdiv.bg

MODELING OF PM10 AIR POLLUTION IN PLOVDIV USING REGRESSION TREES

Desislava Voynikova

Abstract. The purity of the atmospheric air is one of the factors for human health and its pollution with particulate matter of less than 10 microns of

diameter (PM10) is a contemporary problem. In this study, the method of Classification and Regression Trees (CART) was applied and models were built for the PM10 daily mean value in the city of Plovdiv for a period of more than 12 years from 01 January 2011 to 31 January 2023. The best model obtained, describing 90% of the data, was used to forecast concentrations one day ahead, and it was applied 7 times. Different models were created employing combinations of meteorological data, time-counting variables, lagged variables, transformed data, and cross-validation. Contingency tables were also constructed, showing that the CART models correctly predict up to 94% of the measured values with respect to the limit value for PM10 for the protection of human health (Directive 2008/50/EC of the European Parliament) of 50 micrograms per cubic meter.

Key words: Classification and regression trees (CART), Particulate matter PM10, Air pollution modeling, Pollution forecast.