Available online at www.sciencedirect.com
SCIENCE @ DIRECT®
ELSEVIER
Atmospheric Environment 39 (2005) 2145–2155
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

Generalised additive modelling of air pollution, traffic volume and meteorology

Magne Aldrin*, Ingrid Hobæk Haff

Norwegian Computing Center, P.O. Box 114 Blindern, N-0314, Oslo, Norway

Received 22 September 2004; received in revised form 2 December 2004; accepted 17 December 2004

Abstract

We present a general model where the logarithm of hourly concentration of an air pollutant is modelled as a sum of non-linear functions of traffic volume and several meteorological variables. The model can be estimated within the framework of generalised additive models.

Although the model is non-linear, it is simple and easy to interpret. It quantifies how meteorological conditions and traffic volume influence the level of air pollution. A measure of relative importance of each predictor variable is presented.

Separate models are estimated for the concentration of PM_{10}, $PM_{2.5}$, the difference PM_{10}–$PM_{2.5}$, NO_2 and NO_x at four different locations in Oslo, based on hourly data in the period 2001–2003. We obtain a reasonably good fit, in particular for the largest particles, PM_{10} and PM_{10}–$PM_{2.5}$, and for NO_x. The most important predictor variables are related to traffic volume and wind. Further, relative humidity has a clear effect on the PM variables, but not on the NO variables. Other predictor variables, such as temperature, precipitation and snow cover on the ground are of some importance for one or more of the pollutants, but their effects are less pronounced.

Keywords: Air quality modelling; Urban air quality; Particulate matter; Nitrogen oxides; Forward validation

1. Introduction

The growing health problems caused by traffic-related air pollution has resulted in an increased interest in analysis and prediction of the air quality. Several methodologies, both deterministic and statistical, have been proposed. These are often based on linear or non-linear regression models where the concentration of an air pollutant at a specific site is related to traffic volume and meteorological variables.

Levy et al. (2003) relates the concentrations of $PM_{2.5}$, ultra-fine particles and polycyclic aromatic hydrocarbons to traffic volume, wind direction and distance from the road, using linear mixed effects regression models. Chaloulakou et al. (2003) use linear regression to relate PM_{10} and $PM_{2.5}$ concentrations to predictor variables as temperature, wind speed, wind direction, time of year and day of week. They recognise that the meteorological variables are non-linearly related to the concentrations of PM_{10} and $PM_{2.5}$. To handle this, they convert the meteorological predictor variables into binary variables which are used as predictor variables in a modified linear model. Several authors use non-linear methods. Gardner and Dorling (1999), Kukkonen et al. (2003) and Schlink et al. (2003) all conclude that neural networks (see for instance Ripley (1996) for a general reference) are superior to linear techniques in predicting PM_{10}, NO_2,

*Corresponding author. Tel.: +47 22 85 25 00.
E-mail address: magne.aldrin@nr.no (M. Aldrin).

1352-2310/$ - see front matter
doi:10.1016/j.atmosenv.2004.12.020