Atmospheric Environment 67 (2013) 304–312

Contents lists available at SciVerse ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

A sequential factorial analysis approach to characterize the effects of uncertainties for supporting air quality management

S. Wang, G.H. Huang*, A. Veawab

Faculty of Engineering, University of Regina, Regina, Saskatchewan, Canada S4S 0A2

HIGHLIGHTS

► A sequential factorial analysis was proposed to examine the effects of uncertainty.
► Influence of a number of factors of interest was studied in a systematic manner.
► A factor-screening strategy was used to reduce the computational effort.
► Interactive effects of parameters and constraints on model outputs were analyzed.

ARTICLE INFO

Article history:
Received 12 June 2012
Received in revised form
29 October 2012
Accepted 31 October 2012

Keywords:
Air quality management
Factorial analysis
Modeling
Uncertainty

ABSTRACT

This study proposes a sequential factorial analysis (SFA) approach for supporting regional air quality management under uncertainty. SFA is capable not only of examining the interactive effects of input parameters, but also of analyzing the effects of constraints. When there are too many factors involved in practical applications, SFA has the advantage of conducting a sequence of factorial analyses for characterizing the effects of factors in a systematic manner. The factor-screening strategy employed in SFA is effective in greatly reducing the computational effort. The proposed SFA approach is applied to a regional air quality management problem for demonstrating its applicability. The results indicate that the effects of factors are evaluated quantitatively, which can help decision makers identify the key factors that have significant influence on system performance and explore the valuable information that may be veiled beneath their interrelationships.

1. Introduction

Regional air pollution has been a major environmental concern, since it poses a serious threat to human health (Liu et al., 2003). On a global basis, World Health Organization (WHO) estimates that almost 800,000 people die prematurely every year due to air pollution (WHO, 2002). Air pollution is not only a problem of public health; the harmful effects of pollution in materials and environment are also considerable. Therefore, control measures need to be taken to mitigate air pollution. However, an air pollution control system involves many components, such as pollutant emission standards and pollutant treatment efficiencies; these components may be associated with uncertainties. Advanced optimization methods are thus desired for supporting air quality management and pollution control planning under uncertainty.

Over the past decades, a number of inexact optimization methods were developed for dealing with uncertainties in air pollution control systems (Ellis et al., 1985, 1986; Li et al., 2006; An and Eheart, 2007; Lu et al., 2010; Qin et al., 2010). These methods can hardly reveal the interactive effects of uncertainties on system performance, even though they are capable of tackling uncertainties that exist in various forms (e.g., interval numbers, fuzzy sets, and probability distributions). In fact, uncertain parameters are not independent of each other; they interact in different ways. The potential interactions among parameters may greatly influence the performance of air pollution control systems. Therefore, the interactive effects of parameters should not be neglected or underestimated in practical applications.

Previously, factorial analysis was widely used to reflect the potential interrelationships among uncertain parameters and their impacts on system performance. Maqsood et al. (2003) conducted a set of 2^4 factorial experiments for quantitatively analyzing the combined effects of four uncertain input parameters on modeling outputs; for each factorial experiment, a 100-run Monte Carlo simulation was undertaken through a multiphase compositional simulator. Lin et al. (2008) proposed a simulation-aided 2-level factorial analysis approach for characterizing the interactive

* Corresponding author. Tel.: +1 306 585 4095; fax: +1 306 585 4855.
E-mail addresses: gordonhuangevse@gmail.com, huang@iseis.org (G.H. Huang).

1352-2310/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.atmosenv.2012.10.066

effects of composting factors on composting processes. Qin et al. (2008) developed a factorial-design-based stochastic modeling system (FSMS) to investigate the impacts of uncertainties associated with hydrocarbon-contamination transport in subsurface. Mabilia et al. (2010) performed an experimental test according to a fractional factorial design with replicated central point, which identified the best configuration for a formaldehyde passive sampler by statistically evaluating the significance of effects of factors and their interactions. Onsekizoglu et al. (2010) used a two-level factorial experimental design for evaluating the effects of temperature difference between the feed and permeate side of the membrane, concentration of the osmotic agent and flow rate. Zhou and Huang (2011) proposed a factorial two-stage stochastic programming (FTSP) approach for supporting water resources management under uncertainty. Previous studies focus on the investigation of the effects of parameters. In optimization problems, however, constraints may also have a considerable influence on model outputs. Their effects are as important as those of input parameters, and should be analyzed thoroughly. On the other hand, it is unwise to carry out a single, large, and comprehensive experimental design when there are too many factors involved in practical applications. One single design is unlikely to answer all the questions adequately. One potential solution is to conduct a sequence of factorial analyses for examining the effects of factors in a systematic manner.

The objective of this study is to propose a sequential factorial analysis (SFA) approach for supporting air quality management and pollution control planning under uncertainty. SFA can not only investigate the interactive effects of input parameters, but also analyze the effects of constraints. Moreover, SFA has the advantage of conducting a series of factorial analyses to examine the effects of parameters and constraints in a systematic manner. The proposed SFA approach will be applied to a regional air quality management problem for demonstrating its applicability. The findings will help decision makers identify the key parameters, constraints, and their interactions that have a significant influence on system performance, which plays an important role in the decision-making process.

2. Methodology

Firstly, consider a linear programming (LP) problem as follows:

$$\text{Min} f = CX \tag{1a}$$

subject to:

$$AX \le B \tag{1b}$$

$$X \ge 0 \tag{1c}$$

where $C = (c_1, c_2, \ldots, c_n)$, $A = (a_{ij})_{m \times n}$, and $B = (b_1, b_2, \ldots, b_m)^T$ represent the parameters in the objective function and constraints; $X = (x_1, x_2, \ldots, x_n)^T$ is the vector of decision variables. Constraints (1b) can be converted from inequalities to equalities through introducing non-negative slack variables to the left-hand side of the constraints. Model (1) can then be rewritten into the following form:

$$\text{Min} f = CX \tag{2a}$$

subject to:

$$AX + S = B \tag{2b}$$

$$X \ge 0 \tag{2c}$$

$$S \ge 0 \tag{2d}$$

where $S = (s_1, s_2, \ldots, s_m)^T$ represents the vector of the newly introduced slack variables that measure the unused amounts of constrained resources. If a slack variable equals 0, it implies that all resources will be used in order to achieve the minimized objective. The larger a slack variable, the higher the amount of slack resources. A variety of uncertainties are inherent in future-oriented planning efforts, which have different impacts on model response. Thus it is necessary to conduct a systematic analysis of uncertainties. Sensitivity analysis (also called post-optimally analysis) is used extensively in practice; it is performed by varying one parameter over its range with the other parameters held constant. Such a one-parameter-at-a-time approach can only reflect the single effects of parameters, while it fails to consider any possible interaction among them. In fact, joint effects exist among many parameters, which may greatly influence system performance. Therefore, they should not be neglected or underestimated.

Factorial analysis, a multivariate inference method, can thus be introduced to reveal the potential interrelationships among factors and their impacts on system performance (Box et al., 1978). The most important case of factorial analysis is 2^k factorial design which consists of k factors with each at two levels. The statistical model for a complete 2^k design would include 2^k-1 effects that comprise k main effects, $\binom{k}{2}$ two-factor interactions, $\binom{k}{3}$ three-factor interactions, …, and one k-factor interaction (Montgomery, 2001). To estimate an effect or to compute the sum of squares for an effect, the contrast associated with that effect needs to be determined by expanding the right-hand side

$$\text{Contrast}_{AB\cdots K} = (a \pm 1)(b \pm 1)\cdots(k \pm 1) \tag{3}$$

where the sign in each set of parentheses is negative if the factor is included in the effect and positive if the factor is not included; "1" needs to be replaced by (1) that denotes all factors at their low levels when expanding Eq. (3); the high level of a factor is represented by its lowercase letter and the low level of a factor is denoted by the absence of its letter in the treatment combination. Once the contrasts for the effects are determined, the effects and the sum of squares can be estimated according to:

$$\ell_{AB\cdots K} = \frac{2(\text{Contrast}_{AB\cdots K})}{n2^k} \tag{4}$$

and

$$\text{SS}_{AB\cdots K} = \frac{(\text{Contrast}_{AB\cdots K})^2}{n2^k} \tag{5}$$

where $\ell_{AB\cdots K}$ is the single or joint effects of factors; $\text{SS}_{AB\cdots K}$ is the sum of squares for the effects; n denotes the number of replicates.

A 2^k factorial design requires $2 \times 2 \times \cdots \times 2 = 2^k$ runs. As the number of factors in a 2^k factorial design increases, the number of runs required for a complete design increases exponentially, resulting in a great computational burden. In fact, most systems are dominated by main effects and low-order interactions, and most high-order interactions are negligible (Montgomery, 2001). Therefore, fractional factorial design can be introduced to expose the information on main effects and low-order interactions by running only a fraction of runs of the full factorial design. A 2^k fractional factorial design containing 2^{k-p} runs is called a 2^{k-p} fractional factorial design, which requires the selection of p independent generators based on a criterion that the best possible alias relationships can be obtained (Montgomery, 2001). In other words, care should be taken to ensure that the effects of potential interest are not aliased with each other when choosing the generators. The