Building and Environment 125 (2017) 414–422

Contents lists available at ScienceDirect

Building and Environment

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

Turbulent Schmidt number for source term estimation using Bayesian inference

Fei Xue [a], Xiaofeng Li [a, b, *], Ryozo Ooka [c], Hideki Kikumoto [c], Weirong Zhang [d]

[a] Department of Building Science, School of Architecture, Tsinghua University, Beijing, China
[b] Beijing Key Lab of Indoor Air Quality Evaluation and Control, Beijing 100084, China
[c] Institute of Industrial Science, University of Tokyo, Tokyo, Japan
[d] Department of Architecture, Tokyo Polytechnic University, Japan

ARTICLE INFO

Article history:
Received 25 June 2017
Received in revised form
30 August 2017
Accepted 9 September 2017
Available online 12 September 2017

Keywords:
Source term estimation
Bayesian inference
Turbulent Schmidt number
Urban dispersion
CFD

ABSTRACT

Source term estimation (STE) addresses the retrieval of emission source information, including location and strength, based on available information. STE can be viewed as an assimilation process of the observed concentration data measured by a sensor network and the predicted concentration data provided by a dispersion model. When considering emissions in complex urban areas, computational fluid dynamics (CFD) approaches are generally used to provide building-resolving results; however, the value of a key parameter, the turbulent Schmidt number Sc_t, has remained an arbitrary choice. Therefore, it is important to investigate the role of Sc_t in STE problems and determine its optimum value for the purpose of obtaining better estimation results. In this paper, the impact of Sc_t on STE problems is examined, and Bayesian inference is used to improve estimation accuracy by treating Sc_t as an extra unknown parameter. A wind tunnel experiment with a constant point tracer source in an urban-like geometry is used for demonstration. The results show that Sc_t has a major impact on estimation. Larger Sc_t values shift the estimated location towards the upwind direction and decrease the estimated strength. Compared with a conventional estimation method performed by using a pre-assigned value of $Sc_t = 0.7$, treating Sc_t as an unknown improves point estimates, while the uncertainty increases since the proposed method introduces an extra unknown parameter. For source strength, more improvement in point estimates and a larger increase in uncertainty are shown due to its greater sensitivity of Sc_t compared with the source location.

1. Introduction

Either intentional or accidental releases of hazardous materials in the atmosphere can pose great danger to public safety. In such events, the fast and accurate retrieval of source information, e.g., location and strength, is a crucial technique to allow urban environmental management efforts to make appropriate responses and to reduce further damage. This identification of unknown sources is a typical source term estimation (STE) problem, which has received increased attention over the past decade [1,2].

The STE problem is viewed as an ill-posed inverse problem that is characterized by its non-uniqueness and unstable solutions. To make the STE problem more tractable, multiple methods have been developed using various approaches, as described in the state-of-the-art reviews of Hutchinson et al. [1] and Singh et al. [2]. One intuitive solution is to directly inverse the transport process. In this context, Kato, et al. [3] traced the origin of the continental-scale transport of airborne contaminants using backward trajectory analysis. Another more general alternative is the optimization method of minimizing a cost function, which quantifies the discrepancy between observed and predicted concentrations. Based on the features of different STE problems, a series of different cost functions were constructed, from the simplest form of the classical least squares [4–6] to more robust ones using regularization [7] or renormalization [8]. This optimization method produces point estimates of unknown parameters without rigorous quantification of their uncertainties.

The third approach, Bayesian inference, as used in this study, considers the problem of STE in a probabilistic logical manner. All

* Corresponding author. Department of Building Science, School of Architecture, Tsinghua University, Beijing, China.
E-mail address: xfli@mail.tsinghua.edu.cn (X. Li).

http://dx.doi.org/10.1016/j.buildenv.2017.09.012

parameters are regarded as random variables, rather than as constants, with certain probability distributions. Accordingly, Bayesian inference provides not only the point estimations of unknown parameters but also their probability distributions, thus providing a more natural method of uncertainty estimation. However, to date, the calculation of this probability distribution has been a difficult task. Xue et al. [9] proposed a two-step deterministic method to deconstruct the calculation of the joint posterior distribution of the source parameters into calculations of two separate distributions. More general solutions are the stochastic sampling approaches, e.g., the Markov chain Monte Carlo (MCMC) method, which has been widely applied to the STE problems of whether dealing with single sources [10] or multiple sources [11,12], on flat terrain [13] or complex urban areas [14]. Bayesian inference is also known for its flexibility in handling extra unknowns. Senocak et al. [15] estimated the diffusion coefficient in their Gaussian dispersion model in addition to estimating source location and strength. Ristic et al. [16] considered wind speed as an unknown variable and located a source using binary readings.

The STE problem, as mentioned earlier, is an ill-posed inverse problem. Poor model predictions of concentrations often cause heavy deviations in estimated values. The accuracy of the dispersion model is crucial. On flat terrain, Gaussian dispersion models are generally preferred due to their simplicity and rapid calculations [11,13,15]. However, these models cannot characterize the complex flow caused by the presence of buildings and other geometries in an urban area [17,18], where computational fluid dynamics (CFD) modeling is often employed to reproduce the realistic flow and pollutant transport [8,14,19].

In the vast majority of STE investigations performed using CFD modeling, the Reynolds-averaging approach is employed to model the turbulent flow and the turbulent passive scalar transport, assuming the gradient diffusion hypothesis in order to close the turbulent scalar-flux term $-\overline{u'c'}$, as follows

$$-\overline{u'c'} = D_t \nabla c \tag{1}$$

where u' and c' are the fluctuating components of the velocity and concentration, respectively, D_t is the turbulent mass diffusivity, and ∇c is the mean concentration gradient. Combest et al. [20] published a comprehensive review of the above content. By far, the simplest and most popular way to account for D_t is to assume that there is a similarity between turbulent mass diffusivity and turbulent momentum diffusivity (i.e., eddy viscosity ν_t) by assigning a global turbulent Schmidt number

$$Sc_t = \nu_t / D_t \tag{2}$$

Sc_t plays an important role in the scalar transport modeling, as its specific value has a dominant impact on the accuracy of predictions. Tominaga and Stathopoulos [21] reviewed relevant investigations performed over the past several decades and concluded that the optimum Sc_t value is problem dependent and should be selected carefully. In STE studies, despite the fact that Sc_t have been widely recognized as a dominant effect in urban dispersion simulation [21] and consequently be one of the most important factors in STE, the influence of Sc_t has been rarely addressed. In addition, to date, the selection of Sc_t has been quite arbitrary; surprisingly, in several papers, the value of Sc_t is not even reported.

Therefore, it is important to investigate the role of Sc_t in the STE problems and the determination of its optimum value in order to obtain better estimation results. In this paper, the impact of Sc_t in the STE problems is examined, and the possibility of using Bayesian inference to improve the estimation of the source location and strength by treating Sc_t as an extra unknown parameter is explored. In Section 2, the source term estimation problem and its solution are formulated using the Bayesian framework. Section 3 presents the calculation method of the source-receptor relationship using the CFD approach and the adjoint equations. Section 4 describes the experimental and simulation settings of the demonstration case, which is a wind tunnel scenario of a point tracer source with a constant releasing strength in an urban mock-up. The results and discussion are provided in Section 5. First, the impact of Sc_t on the estimation results is analyzed; then, we present the estimation results obtained using the proposed method and compare its performance to that of an existing method. The conclusions are described in Section 6.

2. Bayesian inference

Bayesian inference was first applied to atmospheric STE problems by Chow et al. [10] and Keats et al. [14], who provided the fundamentals of the method. In this paper, based on their work, the influence of Sc_t is taken into consideration in the Bayesian STE framework, as the flowchart shown in Fig. 1. The proposed method is demonstrated using a basic case of a single point source with a constant releasing strength; however, by combining this with other Bayesian STE methods, it can be generalized to address multiple point sources [11,12,22], as well as variable releasing strengths [13,23,24].

2.1. Problem formulation

Bayesian inference addresses parameter estimation problems in a probabilistic way. Assume that we are interested in estimating a parameter set, $\boldsymbol{\theta}$, which may include source location, strength, or other unknown quantities, given the measurement information, $\boldsymbol{\mu}$, obtained from a network of sensors. Bayesian theory provides a rigorous way to make an inference based on all of the information given in this problem. More specifically, the estimation results are obtained as the posterior probability, which is given by Bayes' theorem as

$$p(\boldsymbol{\theta}|\boldsymbol{\mu}) = \frac{p(\boldsymbol{\mu}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{\mu})} \propto p(\boldsymbol{\mu}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{3}$$

where $p(\boldsymbol{\mu}|\boldsymbol{\theta})$ is the likelihood function, $p(\boldsymbol{\theta})$ is the prior probability reflects the a priori knowledge about the unknown parameters, $p(\boldsymbol{\mu})$ is the evidence, acting as a normalizing constant in order to obtain the posterior distribution, $p(\boldsymbol{\theta}|\boldsymbol{\mu})$ is the posterior probability, which is the quantity of interest in the STE problems and contains information required for the inference. To calculate $p(\boldsymbol{\theta}|\boldsymbol{\mu})$ requires assigning the appropriate function forms for the likelihood function $p(\boldsymbol{\mu}|\boldsymbol{\theta})$ and the prior probability $p(\boldsymbol{\theta})$.

2.2. Likelihood function

In this study, we consider the STE problem of a single point source releasing with a constant strength and an undefined Sc_t, represented by $\boldsymbol{\theta} = (\boldsymbol{x_s}, q, Sc_t)$ where $\boldsymbol{x_s}$ is the source location and q is the source strength. Given a set of parameters $\boldsymbol{\theta}$, assuming there are M sensors deployed in the field, the relation between observations and predictions is denoted by an additive noise model:

$$\boldsymbol{\mu} = q\boldsymbol{h}(\boldsymbol{x_s}, Sc_t) + \boldsymbol{\varepsilon} \tag{4}$$

where $\boldsymbol{\mu} \in \mathbb{R}^M$ is the vector of concentration measurements, $\boldsymbol{h}(\boldsymbol{x_s}, Sc_t) \in \mathbb{R}^M$ is the source-receptor relationship vector representing the predicted mean concentrations at the sensors if a