Building and Environment 64 (2013) 169–176

Contents lists available at SciVerse ScienceDirect

Building and Environment

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

Forecasting simulations of indoor environment using data assimilation via an Ensemble Kalman Filter

Cheng-Chun Lin, Liangzhu (Leon) Wang*

Department of Building, Civil and Environmental Engineering, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec H3G1M8, Canada

ARTICLE INFO

Article history:
Received 27 December 2012
Received in revised form
14 March 2013
Accepted 16 March 2013

Keywords:
Data assimilation
Ensemble Kalman Filter
Indoor environment
Forecasting simulation
CONTAM
OpenDA

ABSTRACT

Data assimilation is widely used in weather forecasting and other complex forecasting problems such as hydrology, meteorology, and fire dynamics. Among various data assimilation methods, the Ensemble Kalman Filter (EnKF) is one of the best solutions to large-scale nonlinear problems while the computational cost is relatively less intense than other forecasting methods. In this paper, a new application of EnKF to forecast indoor contaminant concentrations is presented. The first part of the paper introduces the fundamental theories of data assimilation. The second part is a case study of forecasting the concentrations of a tracer gas in a multi-zone manufactured house by using a mass balance model with an EnKF. The benefits of EnKF and several important parameters for EnKF were discussed including numbers of ensemble members and observations, time step of observations, and forecasting lead time. The EnKF method presented is one of the first studies applied to the indoor environment field. It was shown that by using EnKF, the predictability of the simple indoor air model for the multi-zone space was improved significantly.

1. Introduction

The forecasting of indoor environment is of great interests due to its close relationship to occupant's safety [1], thermal comfort [2], and energy efficiency [3,4], for which an accurate prediction of important parameters is often needed such as temperature, relative humidity and contaminant concentration. Forecasting these indoor air properties, especially solving contaminant transport problem in a dynamic environment, is a difficult task since the physical states of the building environment could change rapidly over time [5]. Under such uncontrollable factors as ambient temperature, velocity, humidity, and occupant loads, the contaminant estimation is hard to achieve by conventional methods using steady state analysis of constant model parameters. In addition, sudden release of contaminant, opening doors and windows, change of occupant behavior and the use of electric appliances are a few common examples that may further increase the difficulty of solving the forecasting problem. By using any numerical model to predict future indoor air contaminants, these uncertain events will cause the predicted physical states to depart from reality as the model evolves forward over time. Previous studies showed a few different ways of indoor environment forecasting. Federspiel utilized a method originating from the optimization theory, so-called the Kalman Filter, to estimate the strength of gas sources in buildings successfully [6]. But the model has some restrictions when it applies to multi-zone problems since source strength and air flow rate must be known a priori. Kemajou presented that indoor air temperature and relative humidity could be quickly predicted by using artificial neural network (ANN) [7]. Vukovic [8] and Bastani [9] applied ANN to identify indoor contaminant source successfully and can be used to optimize indoor air sensor networks. But the potential problem of using a black-box type of methods like ANN is that the training of models still relies on trial and error to determine an optimum model structure while the training parameters usually cannot be applied to other buildings. Another limitation is its inefficiency for each new case, in which huge amount of data is required for the training. Sreedharan introduced another model of forecasting simulation using Bayesian Monte Carlo method to quickly analyze measurements from multiple indoor air sensors [10]. This system can monitor real-time indoor environment to help protect occupants by locating the release source of a high-risk pollutant. Follow-up studies [11,12] focus more on using heterogeneous sensor systems such as monitoring door position and mechanic ventilation operating status. Extensive reviews of other methods for locating indoor air contaminants can be found in the

* Corresponding author. Tel.: +1 514 848 2424x5766; fax: +1 514 848 7965.
E-mail addresses: leonwang@bcee.concordia.ca, lzwang@gmail.com (L.(Leon) Wang).

0360-1323/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.buildenv.2013.03.008

article conducted by Liu and Zhai, 2007 [13]. Most of these previous studies focus on locating contaminant source instead of predicting dynamic future evolvement of concentrations in a multizone building. In this paper, the technique of data assimilation is applied to forecast indoor contaminants transport.

The major task of forecasting simulations is to predict future states of physical phenomenon with a certain lead time and accuracy. In order to find optimal state variables, data assimilation provides different algorithms for parameter estimation while taking into account uncertainties of measurements and numerical predictions.

In 1960, a pioneer research of data assimilation theory has been established by R. E. Kalman [14], the Kalman Filter, which provides a recursive solution to find a best possible estimation of the true state. Instead of finding an optimal estimation for one value as best linear unbiased estimation (BLUE), the Kalman Filter can be applied to a dynamic model that evolves over time [15,16]. In order to solve different types of nonlinear problems, a few Kalman Filter variants have been proposed. In the Extended Kalman Filter (XKF), the nonlinear models are linearized by using partial derivatives which is similar to a Taylor series expansion. It has been proven that XKF is effective in many applications but its weakness is the error probability density is not fully considered in the linearization. Although the analysis scheme of XKF is similar to traditional Kalman Filter, the computational requirements of XKF are drastically increased by additional numerical operations, such as linearization. Another mainstream of the data assimilation theories for nonlinear problems is four-dimensional variational assimilation (4D-Var) and is widely used to weather forecasting [17]. Like its counterpart, three-dimensional variational assimilation (3D-Var), 4D-Var is based on minimizing a cost function. In order to calculate the gradient of the cost function for minimization, it is required to manipulate large matrices, which makes 4D-Var computationally intensive. Evensen proposed a more affordable method, Ensemble Kalman Filter, to determine error statistics by using the Monte Carlo method [18]. The method reduces computational requirement of XKF by using ensemble members, which are similar to the samplings in other Monte Carlo methods, to avoid direct calculation and storage of the evolution of the large error covariance matrices. Each ensemble member in EnKF can be calculated separately so it is especially suitable for parallel computing and solving large scale problems. This method has been widely used in weather forecasting, hydrology and fire dynamics predictions [16,19]. No previous studies have been reported to use EnKF for indoor environment simulations. Different from existent studies, the indoor air forecasting model with the EnKF is performed without determining source strength and location but instead depending on the accurate estimation of error statistics including uncertainties from both experiment and numerical model. Therefore the EnKF can perform faster prediction than other methods but relies on rapidly obtaining measurement data.

The objective of this paper is to explore the applications of EnKF to forecasting indoor air environment and discuss the key parameters involved in the accuracy of EnKF for indoor air contaminants transport forecasting. This paper applied an EnKF to a case study of forecasting the concentrations of a tracer gas in a multi-zone manufactured house by using a mass balance model. The benefits of EnKF and several important parameters for EnKF were discussed including numbers of ensemble members and observations, time step of observations, and forecasting lead time. The EnKF method presented is one of the first studies applying a weather forecasting model to indoor environment field. In this paper, all numerical operations relating to data assimilation are based on a generic toolbox for data assimilation, OpenDA [20], developed by Delft University of Technology, Netherlands.

2. Methodology

The detailed explanations on data assimilation and the fundamental theories of EnKF can be found in many references [14,15,18and20] so they are not covered in this paper to avoid repetition. Instead of copying down the math of EnKF, this paper will focus on how the fundamental methodology of EnKF is applied to the indoor air modeling of a multi-zone manufactured house.

2.1. Multi-zone manufactured house

The experimental data for data assimilation come from a series of tracer gas measurements in a manufactured house conducted by National Institute of Standards and Technology (NIST) in 2007 [21,22]. The house includes living room, family room, kitchen and three bedrooms without people during the time of the test and with mechanical ventilation running under normal control and schedule as shown in Fig. 1. An attached garage was added after the measured data were collected so it was not included in this study. During the experiment, tracer gas, sulfur hexafluoride (SF_6), was injected about every six hours in the living room and the concentration of SF_6 was measured at various indoor locations for every ten minutes by gas chromatography. The system is capable of measuring SF_6 concentration over a range of 3–300 ppb, with an uncertainty of about 5% of the reading. In this study, bedroom 3 was excluded because its sampling line was moved to the outside for another research project. Please note that the measurement of each room starts from different time step as shown in Table 1. For example, master bedroom was measured at $t = 0$, 10 min and so on, and bedroom 2 is measured at $t = 1$, 11 min etc. The average air change between indoors and outdoors was calculated based on the tracer gas decay method, which was about 0.1–0.4 air changes per hour (ACH). The SF_6 test has been continuously running and it covers large variants of test conditions so the test house was selected. In this study, we took twelve hours measurement data from four different locations as observation while two injections are included. The first one is at the beginning of the experiment and the other one is about six hours after the first injection. The injection rate is intended to achieve an average initial gas concentration of around 120 ppb but is assumed to be unknown in the simulations. The instruments were calibrated regularly and believed to be accurate enough. 5% error is thus assigned to the observation errors accounting for the mentioned uncertainties. The uncertainties associated with the mixing conditions are considered in the numerical model in the following sections.

2.2. Tracer gas concentration model

The modeling of indoor environment here is to forecast the SF_6 concentrations in the different rooms of the house by a mass

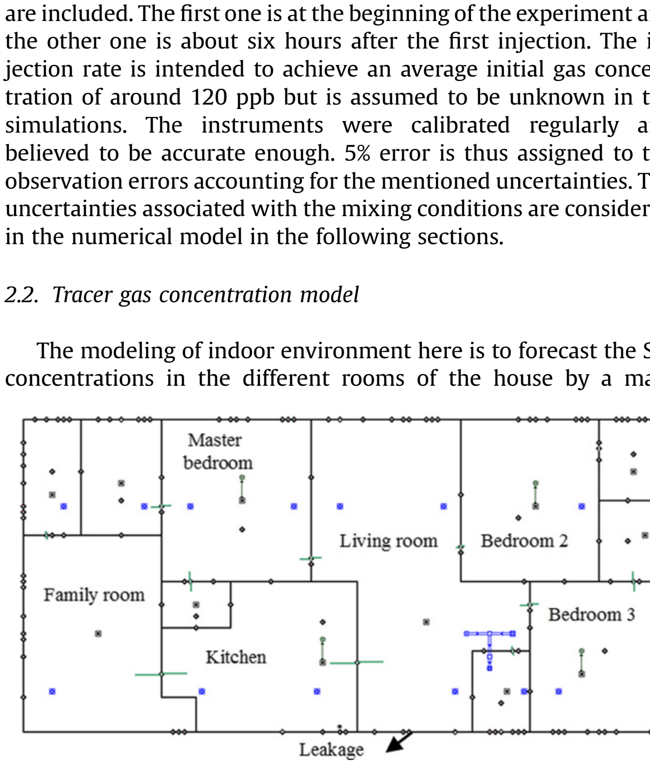

Fig. 1. CONTAM model of the manufacture house [20].