

DOCTORAL THESIS

Computational methods in air quality data

Zhu, Zhaochen

Date of Award:
2017

[Link to publication](#)

General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

Abstract

In this thesis, we have investigated several computational methods on data assimilation for air quality prediction, especially on the characteristic of sparse matrix and the underlying information of gradient in the concentration of pollutant species.

In the first part, we have studied the ensemble Kalman filter (EnKF) for chemical species simulation in air quality forecast data assimilation. The main contribution of this paper is to study the sparse data observations and make use of the matrix structure of the Kalman filter updated equations to design an algorithm to compute the analysis of chemical species in the air quality forecast system efficiently. The proposed method can also handle the combined observations from multiple species together. We have applied the proposed method and tested its performance for real air quality data assimilation. Numerical examples have demonstrated the efficiency of the proposed computational method for Kalman filter update, and the effectiveness of the proposed method for NO_2 , NO , CO , SO_2 , O_3 , $\text{PM}_{2.5}$ and PM_{10} in air quality data assimilation.

For the second part, we have proposed and developed an optimization approach for data assimilation by using the gradients of the forecast of state variables. We have studied an objective function consisting of two data-fitting terms. The first term is based on the difference between the gradients of the forecast and the analysis of the state variables while the second term is based on the difference between the observations and the projected analysis of state variables. Here the existence and uniqueness of the analysis solution of the proposed objective function are shown. The solution can be calculated by using the conjugate gradient method iteratively. Experimental results based on the Community Multi-scale Air Quality (CMAQ) are presented. We then showed that the prediction performance of the proposed method is better than that of other testing methods in air quality and weather data assimilation.

Within the third part, we have set up an automatic workflow to connect the management system of the chemical transport model - CMAQ with our proposed data assimilation methods. The setup has successfully integrated the data assimilation into the management system and shown that the accuracy of the prediction has risen

to a new level. This technique has transformed the system into a real-time and high-precision system. When the new observations are available, the predictions can then be estimated almost instantaneously. Then the agencies are able to make the decisions and respond to the situations immediately. In this way, citizens are able to protect themselves effectively. Meanwhile, it allows the mathematical algorithm to be industrialized implying that the improvements on data assimilation have directly positive effects on the developments of the environment, the human health and the society. Therefore, this has become an inspiring indication to encourage us to study, achieve and even devote more research into this promising method.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor Prof. Michael. K. Ng, for not only his perspicacious guidance, inspiring encouragement and financial support throughout the past four years, but also for spending extra time in valuable discussion with me and sharing his ideas on my research. His strict and devoted attitude towards scientific research, his patience and responsibility to students and his kindness and consideration in all matters will be remembered in my heart forever. Receiving an enormous amount of academic knowledge from him in the fields of data assimilation, mathematics and research, and experiencing his excellent personal charisma of sustained attention, persistent efforts and professional attitude are the greatest memories to me and I will be forever grateful to him.

I would also like to thank the Department of Mathematics for providing the excellent studying conditions and facilities and various departmental administrative staffs: Ms. Claudia Chui, Ms. Vicky Hui, Ms. Candy Li and Ms. Jenny Cheung; and the computer technicians, Ms. Tammy Lam and Mr. Cheong Wing Yeung for their kind help and assistance. I would also like to thank my classmates and friends who have given life to my journal because they have brought much happiness to me and made my life more enterprising and entertaining.

Finally, let me express my immense thanks to my family for providing me with unflinching support and continuous encouragement throughout my years of study and during the process of my researching and writing of this thesis. This accomplishment would not possibly have been achieved without the understanding and assistance of the above people.

Table of Contents

Declaration	i
Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Figures	viii
List of Tables	xi
Chapter 1 Introduction	1
1.1 Background	1
1.2 Data Assimilation	3
1.3 The Contribution	6
1.4 Thesis Outline	7
Chapter 2 Air Quality Model Simulation	11
2.1 Air Quality Model	11
2.2 Photochemical Air Quality Model	13
2.3 The Conceptual Model of Community Multi-scale Air Quality (CMAQ)	14
2.3.1 Overview	15
2.3.2 Meteorology-Chemistry Interface Processor (MCIP)	16
2.3.3 Initial and Boundary Conditions (ICON and BCON)	17
2.3.4 Clear-sky Photolysis Rate Calculator (JPROC)	18
2.4 CMAQ Chemistry Transport Model (CCTM)	18

2.4.1	Derivation of the Chemical Transport Model	20
2.4.2	Advection and Diffusion	26
2.4.3	Deposition	28
2.4.4	Gas-phase Chemistry Solvers and Process	29
2.4.5	Aerosol Process	30
2.4.6	Cloud Mixing and Aqueous-Phase Chemistry (CLOUD)	31
2.4.7	Plume-in-Grid (PING)	31
Chapter 3 Data Assimilation Methods		33
3.1	Optimal Interpolation (OI)	34
3.2	Variational Methods	36
3.2.1	3D-Var: Three-Dimensional Variation Method	37
3.2.2	Advantages of 3D-Var compared to the OI	38
3.2.3	4D-Var: Four-Dimensional Variation Method	39
3.2.4	Characteristics of 4D-Var compared to the 3D-Var and OI	41
3.3	Kalman Filter Methods	42
3.3.1	EKF: Extended Kalman Filter	42
3.3.2	EnKF: The Classical (Stochastic) Ensemble Kalman Filter with Perturbed Observations	45
3.3.3	Advantage of EnKF compared to the EKF	47
3.3.4	The Comparison between 4D-Var and EnKF	47
3.3.5	Covariance Inflation of EnKF	48
3.3.6	Localization of EnKF (LEnKF)	49
3.3.7	Characteristics and Advantages of LEnKF compared to the EnKF	51
3.3.8	EnSRF: Ensemble Square-root Kalman Filter	51
Chapter 4 Sparse Matrix Computation		53
4.1	Motivation	53
4.2	Proposed Method	54
4.2.1	Sparse Matrix Computation	54
4.2.2	Multiple Species	56
4.3	Implementation	58

4.3.1	Overview of the Connection between CCTM and DA	58
4.3.2	Configurations on the CMAQ System to meet with Data Assimilation	61
4.3.3	Configurations on the Data Assimilation to fit the CMAQ System	70
4.3.4	Techniques used in the Proposed Data Assimilation Method .	76
4.4	Numerical Results	80
4.4.1	Single Species	81
4.4.2	Multiple Species	98
Chapter 5	Optimization Method	105
5.1	Motivation	105
5.2	Proposed Method	106
5.2.1	Use of Forecast Gradients	106
5.2.2	Multiple Species	108
5.3	Implementation	111
5.3.1	Reduction for the Matrix-Vector Multiplication with Sparse Characteristics	111
5.3.2	Reduction for the Matrix-Vector Multiplication with Geometrical Feature	113
5.4	Numerical Results	115
5.4.1	Single Species	117
5.4.2	Multiple Species	121
Chapter 6	Summary and Future Research	123
6.1	Summary	123
6.2	Future Research	124
	Curriculum Vitae	134