A Prototype of Collaborative Virtual Geographic Environments to Facilitate Air Pollution Simulation

XU, Bingli

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in

GeoInformation Science

The Chinese University of Hong Kong

July 2009

This research is funded by two foundations:

1.用于城市群空气污染研究的协同虚拟地理环境系统

Funded by "863" High-Tec program Plan of P.R.C. from 2006.12 to 2009.11 (2006AA12Z207)

2. A Study on the Collaborative Virtual Geographic Environment (CVGE) for Simulating Air Pollution in Pearl River Delta (PRD) Region

Funded by Research Grants Council of Hong Kong--the Competitive Earmarked Research Grant for 2007/2008 (RGC447807)

ABSTRACT

of thesis entitled:

A Prototype of Collaborative Virtual Geographic Environments to Facilitate Air Pollution Simulation

Submitted by Xu,Bingli
for the degree of Doctor of Philosophy
at The Chinese University of Hong Kong
in April 2009

The air pollution that is associated with global economic growth is a global problem. Scientists, governmental officials and the public are focusing on improving understanding, accurately predicting and efficiently controlling levels of air pollution. Air pollution simulation is one method used to achieve these goals. This research will consider a computer supported simulation.

Air pollution simulation has several components, including data preparation, atmospheric circulation and air pollution dispersion modelling and computation, visualization of the model computation results, analysis and model evaluation. Of these components, only atmospheric circulation and air pollution dispersion modelling and computation are mature, while the other components are weak to a greater or lesser extent. To address these weaknesses, this thesis proposes to integrate the data, modelling and analysis into a multi-dimension, virtual geographically referenced environment. In addition, collaboration is used to solve the problem of multi-disciplinary knowledge requirements for conducting air pollution simulation. Based on this model, a concept of collaborative virtual geographic environments (CVGE) is proposed.

The focus of this thesis is two-fold: one is on the development of a conceptual framework and prototype of CVGE from practice of air pollution simulation; the other is on applying this framework to facilitate air pollution simulation. The work of this

thesis can be summarized as follows. (1) Defining the concept of CVGE, developing a conceptual framework for CVGE and discussing primary theories of CVGE. (2) Designing the architecture of a CVGE prototype to facilitate air pollution simulation. (3) Integration and computation of a complex atmospheric circulation model and an air pollution dispersion model based on high performance computation. (4) Geo-visualization of air pollution distribution and dispersion based on calculations using air pollution dispersion models. (5) Geo-collaboration for air pollution simulation. And finally (6) CVGE prototype based air pollution simulation.

The contributions can be drawn from two aspects—CVGE and practice of air pollution simulation. Regarding CVGE, thesis ① develops the conceptual framework of CVGE; ② designs the architecture of a CVGE prototype in order to facilitate air pollution simulation; ③ proposes the concept of a "fuzzy boundary volume object", and designs a solution composed of a particle system wrapped in pollution boxes; and ④ examines the levels of geo-collaboration for air pollution simulation. For air pollution simulation, thesis ① integrates air pollution sources, geo-data, an atmospheric circulation model, an air pollution dispersion model, geo-visualization and analysis into a collaborative virtual geographic environment, which is able to supply a new research methodology and platform for air pollution simulation; in ②, the new platform is scalable and able to free the restrictions of operations on visualization, which paves the way for further extension; ③ couples air pollution dispersion models with geo-information, opening up opportunities for cross studies between air pollution and other research areas, such as the economy, public health and urban planning.

The motivation for future research has two main aspects again—CVGE and practice of air pollution simulation. For the aspect of CVGE, possibilities for future research include: ① more detailed research on the CVGE concept, primary theories and methodologies; ② the efficient integration and management of heterogeneous geo-models with CVGE in standardization; and ③ the efficient rendering of a complex structured object in CVGE. Regarding practice, future research can be conducted into: ① extending air pollution dispersion models; and ② improving the efficiency of air pollutant rendering with a particle system wrapped in pollution boxes.

摘　要

經濟增長伴隨的空氣污染已經成爲全球性問題。如何有效的認識、預測、控制空氣污染，以及如何最大程度降低空氣污染帶來的危害成爲全世界共同面對的課題。計算機支援下的模擬是研究和認識空氣污染的主要手段之一，它具有再現過去和預測未來的雙重功能。

空氣污染模擬包括資料準備、大氣環流與污染擴散建模與計算、計算結果視覺化、分析、以及模型評估等。除了大氣環流與污染擴散建模和計算比較成熟外，其他幾個步驟在功能上存在很大的不足。爲了彌補這些不足，本研究提出將資料、模型以及時空分析集成在一個具有多維地學視覺化平臺下。同時，平臺利用其分佈的特性，將分佈異地的多用戶連接在一起，進行協同研究問題。本研究從地理學的角度出發，將這個平臺定義爲協同虛擬地理環境平臺。

本研究以空氣污染模擬爲背景，構建協同虛擬地理環境的概念框架以及原型系統。並以此爲基礎，研究原型系統支撐下的空氣污染的協同模擬。具體的研究內容包括：① 探討 CVGE 的概念框架、研究內容和基礎理論；②構建基於 CVGE 的空氣污染模擬原型系統框架；③複雜空氣污染模型的高性能計算與集成；④空氣污染模型結果的地學視覺化表達;⑤空氣污染模擬的地理協同；⑥基於 CVGE 原型系統的空氣污染協同模擬。

本研究首次將地理協同應用到空氣污染模擬。創新點體現在理論和應用兩個方面。在理論上：① 比較系統地闡述了 CVGE 的概念和基本理論；② 設計了針對 CVGE 的基於高性能平臺的複雜模型計算問題，設計了高斯模型的平行算法；③ 提出了“模糊邊界體三維”的概念，同時設計了基於包圍盒和粒子系統的表達方式；④ 明確了基於 CVGE 的空氣污染模擬的協同層次，設計了相應的

協同框架，研究了協同的方式，同時探討了基於對比協同衝突探測機制以及基於規則和線上交流的衝突協調機制。在應用上：① 將污染資料、污染模型、污染視覺化表達以及結果分析集成在統一的協同虛擬地理環境中，爲空氣污染模擬提供新的研究思路和研究平臺；② 系統具有可擴展性，同時開放了人機交互的功能限制，爲後續的人機界面的高交互性奠定基礎；③ 將空氣污染的專業模型與地理資訊進行耦合，爲空氣污染的時空分析打下了基礎，同時搭建了空氣污染與其他學科（如經濟學，公共衛生等）交叉研究的橋樑。

本研究的不足之處，同時也是今後的工作，主要體現如下：在理論上：① CVGE 的概念和理論探索尚屬起步階段，還有待進一步昇華；② CVGE 中地學模型集成與管理的通用性理論與方法有待深入開展；③ CVGE 中動態過程的高效表達需要提升；④ 需要開展針對 CVGE 中地理協同的廣度、深度以及可用性與方便性的研究。在應用上：① 擴展空氣污染模型，使之能夠接受更多的空氣污染模型；② 污染視覺化的效率有待進一步提高。

ACKNOWLEDGEMENTS

I am greatly indebted to my supervisor Professor Lin Hui and co-supervisor Professor Chiu LongSang for their continuous help and support during the course of this research and for their expert and careful reading of this thesis. Their insightful suggestions and encouragements have continuously filled me with the much-needed confidence to keep going. Without their support and help, the completion of this thesis would have been impossible and I would never be able to go this far.

My sincere thanks go to Professor Michael Batty from Centre for Advanced Spatial Analysis - University College London, Professor Tsou Jin Yeu and Professor Sun Hanqiu from the Chinese University of Hong Kong (CUHK). They are members of my thesis examination committee and have paid their concern about my work. My thanks also go to Professor Leo Jiaya Jia for his questions on behave of Professor Sun Hanqiu during my oral defense.

Sincere thanks go to Professor Lin Wenshi from Sun Yat-Sen University, Mr. Sammy Tang and Mr. Jimmy Cheung from the Information Technology and Service Center of CUHK, Doctor Zhu Jun and Mr. Hu Ya from the Institute of Space and Earth Information Science, who are also my group members, for their fully and kindly support on MM5 modeling and computation and some technical assistance.

I would like to thank Professor Gong Jianhua from Institute of Remote Sensing Applications of Chinese Academy of Science, Professor Lv Guonian from Key Laboratory of Virtual Geographic Environments of NanJing Normal university, Professor Chen Chongcheng from Key Laboratory of Spatial Data Mining and Information Sharing of Ministry of Education of Fuzhou University, Professor Chen Xi from State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering of HoHai University, Professor You Xiong and Professor Wan Gang from Information Engineering University of the People's Liberation Army, Professor Kong Yunfeng from HeNan University. During my study in CUHK, my academic visiting to their institutes, laboratories, or centers broaden my horizon, which is more than valuable to enhance my thesis research.

I want to show my thanks to Professor Huang Bo from Department of Geography and Resource Management (GRM) of CUHK, Professor Zhang Yuanzhi from CUHK, Professor Guoray Cai from College of Information Sciences and Technology of Penn State University, Professor Christophe Claramunt from Naval Academy Research Institute of France, and Professor Juval Portugali from Tel Aviv Univerisity. The discussion with them on this topic encourages me to enjoy the research.

Special thanks are due to Dr. Hu Mingyuan, Dr. Huang Fengru, Dr. Chen Jinsong, Dr. Zhang Lu, Dr. Eric, Dr. Jiang Liming, Dr. Fang Chaoyang, Dr. Matthew Pang, Dr. Zhao Yibin, Mr. Hu Xianzhi, Miss Zhang Peiyao, Mr. Wu Lei, Mr. Wang Guiwu, Mr. Zhang YongJun, Mr. Zheng Hailong, Mr. Liu Jianbo and other colleagues and friends at Institute of Space and Information Science (ISEIS) and the department of Geography and Resources Management (GRM) for their assistance and advice on my study, and playing with them. I would also like to thank Miss Chloris Yip, Miss Wendy Wong, Mr. Isaac Chan, and Mr. Anthony Wong, who generously provided me with comfortable research environment. I would like to say thanks to Professor Xia Yunqing from Tsinghua University, Dr. Huang Mingxiang from Ministry of Environmental Protection of PR.China for those nice days staying with them.

I would like to say thanks to all the authors whose publications are cited by this thesis.

Most of all, I would like to thank my family. The support and encouragement from my parents and my wife encouraged me to concentrate on my study. Without them, I could not have completed this research. Special thanks to my lovely baby. The birth of him injects me more energy to fulfill this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

2D	Two-dimension
3D	Three-dimension
AERMOD	AMS/EPA Regulatory Model
CAVE	Cave Automatic Virtual Environment
CGVE	Collaborative geographic visualization environments
CSCW	Computer supported cooperative work
CUGrid	Computational grid of the Chinese University of Hong Kong
CUHK	The Chinese University of Hong Kong
CVF	Visualization collaboration followers
CVGE	Collaborative Virtual Geographic Environments
DEM	Digital elevation model
DFBVO	Dynamic fuzzy boundary volume object
DIS	Distributed interactive simulation
DMSO	Defense Modelling and Simulation Office
DRM	Data Representation Model
EDCS	Environmental Data Coding Specification
EPD	Environmental Protection Department
EVS	Environmental Visualization System
FBVO	Fuzzy boundary volume object
FTP	File Transfer Protocol
GC	Grid computation
GDP	Gross Domestic Product
GIS	Geographic Information System
GIScience	Geographic Information Science
GPM	Gaussian Plume Model
GPRD	Greater Pearl River Delta
GrADS	Grid Analysis and Display System
GSI	Grid Security Infrastructure
HCI	Human-computer interaction
HKO	Hong Kong Observatory
HLA	High Level Architecture
HTTP	Hypertext Transfer Protocol
IEC	International Electrotechnical Commission
ISEIS	The Institute of Space and Earth Information Science
ISO	International Organization for Standardization
LES	Lakes Environmental Software
LOD	Level of detail
LTSD	Layer and Time Sequence Data
LULC	Land Use and Land Cover
MA	Mobile agent
MM5	Fifth-Generation NCAR / Penn State Mesoscale Mode

OGSA	Open Grid Services Architecture
OSG	OpenSceneGraph
P2P	Peer to peer
PA	Parallel arithmetic
PBL	Planetary boundary layer
PDA	Personal digital assistant
PHC	High performance computers
PM_{10}	Particulate with an aerodynamic diameter smaller than or equal to 10 micrometers
PPGIS	Public-participation GIS
PRAGMA	Pacific Rim Applications and Grid Middleware Assembly
PRD	Pearl River Delta
RMTP	Reliable multicast transport protocol
RS	Remote Sensing
SCO	Simulation Coordination Office
SDSS	Spatial decision-support systems
SEDRIS	Synthetic Environment Data Representation and Interchange Specification
SO_2	Sulphur dioxide
SRM	Selective reliable multicast
SRM	Spatial Reference Model
SRTP	Selective reliable transmission protocol
S-SCTP	Smooth synchronous collaboration transport protocol
SYSUM	Sun Yat-sen University Model
TC	Tropical Cyclone
TCP/IP	Transmission Control Protocol/Internet Protocol
TIN	Triangulated Irregular Network
TSD	Time Sequence Data
UDP	User Datagram Protocol
VA	Virtual artifact
VCC	Visualization collaboration creators
VE	Virtual Environments
VGE	Virtual Geographic Environments
VPD	Volume pollution data
VR	Virtual Reality
WiFi	Wireless Fidelity
WSRF	Web Services Resource Framework

CHAPTER 1: INTRODUCTION

1.1 Background and Challenges

Air pollution is a global problem and threatens human health in many ways, including: ①the aggravation of respiratory and cardiovascular disease; ② decreased lung function and increased frequency and severity of respiratory symptoms; ③increased susceptibility to respiratory infections; ④effects on the nervous system; ⑤cancer; and ⑥premature death (EPA 2008). More sensitive individuals appear to be at a greater risk of air pollution-related health problems, for example those with pre-existing heart and lung diseases (e.g. asthma, emphysema and chronic bronchitis), diabetics, older adults and children (EPA 2008). Scientists, governmental officials and the public are focusing on achieving an improved understanding, accurate prediction and efficient control of air pollution. Air pollution simulation is one method to achieve the goals. This research will consider a computer supported simulation.

Computer supported air pollution simulation is a good method not only of facilitating scientific research into the law of air pollution, but also of conducting predictions that can influence decision making on air quality management. In the atmospheric and environmental sciences, computer supported air pollution simulation involves a workflow that includes several components, which are the input data preprocess, air pollution modelling and computation, visualization, model evaluation and analysis. If the simulation is used to support decision making, the goal identification and emission control strategy should also be included in the process. The comprehensive computer supported air pollution simulation, as well as the support provided for decision making, are shown in Figure 1.1.

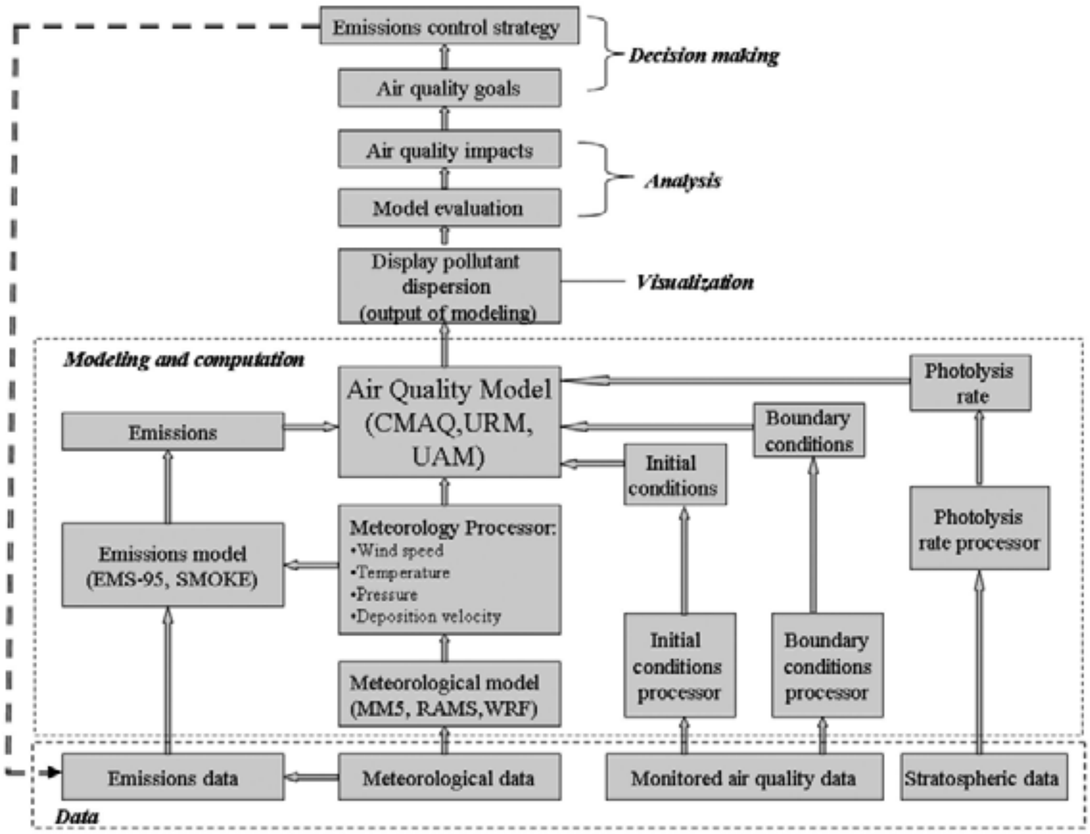


Figure 1.1: The workflow of computer supported air pollution simulation (modified based on (Khan 2003))

Air pollution modelling, which is the core component of air pollution simulation, is now a mature field. Air pollution modelling ranges from short-range dispersion of species (typically accidental release in the case of an industrial hazard), to atmospheric chemistry and climate change. The scales of air pollution modelling include a local scale (accidental release), a regional scale (photochemistry, urban pollution), continental scales (trans-boundary pollution, such as acid rains), and a global scale (for example, atmospheric chemistry in the stratosphere, oxidizing power of the troposphere) (Sportisse 2007). The species that are of interest for air pollution modelling are ozone and volatile organic compounds (photochemistry), trace metals, mercury, methane, carbon monoxide, particulate matter (aerosols), radionuclides and biological species (Sportisse 2007). Numerical models are used to help understand physical processes (assessment of the impact of a given process), environmental forecasts (performed by an emergency centre, for instance), impact studies (of emission sources), sensitivity analysis (with respect to different scenarios of emission) and inverse modelling (of uncertain emissions) (Sportisse 2007).

With the exception of air pollution modelling and computation, the other components of air pollution simulation are rather weak. The arguments for the weaknesses are:

(1) The input data preprocess

Many input data for air pollution modeling are geographically referenced data, or geo-data. To couple the geo-data with air pollution dispersion models, re-project and format transformation for geo-data are always conducted first. Sometimes, the conduction is time consuming for environmental scientists because geo-data are difficult to collect and process using existing technologies from the atmospheric and environmental sciences. Thus, more and more environmental scientists turn to borrow methodologies from GIScience and Remote Sensing to solve this problem. For instance, MM5 (the fifth generation Mesoscale Model) based circulation simulation, which uses high resolution land use and land cover (LULC) data, can improve representation of complicated features of the atmospheric planetary boundary layer (PBL), and consequently affects air quality (Cheng, Kim and Byun 2008, Cheng and Byun 2008). But unfortunately, environmental science does not provide the methodology to process the high resolution LULC data. To solve this problem, environmental scientists always seek support from science of Remote Sensing (RS) because RS can be used to develop the high resolution LULC maps (Kandrika and Roy 2008, Antonarakis, Richards and Brasington 2008, Canty 2009). Another example, the terrain and buildings which are required as the input data for much air pollution modelling, such as AERMOD (AMS/EPA Regulatory Model)(EPA 2006), can be quickly processed and efficiently managed using a geographic information system (GIS).

Even though the geo-data process methodologies are important for air pollution simulation, there is no such a tool integrating them to alleviate burden of environmental scientists.

(2) Visualization of air pollution modelling results

The visualization of air pollution modelling results is useful not only for displaying air pollution distribution and dispersion, but also for facilitating information exploration. Although it is such an important step, the visualization of pollutant distribution and dispersion is currently weak, with a lot of simulation being numerical, and visualization being limited to diagrams. There are many researchers who use 2D and 3D visualization, such as GrADS (Grid Analysis and Display System) (Doty 1995) and Vis5D (Hibbard et al. 1999). However, both 2D and 3D visualizations are scales fixed, and the interaction between operators and program is very low. Apart from the display, 2D and 3D visualizations do not support information queries, visualized analysis, zoom in or out and viewpoint change.

(3) Spatial and temporal analysis

Atmospheric circulation and air pollution dispersion have obvious space-time dynamic characteristics. This can be represented as (x, y, z, t, a), where (x, y, z) represents spatial dimension, t is temporal dimension and a is attribute (this can include many different attributes). Spatial and temporal analysis will thus facilitate the discovery of any spatial and temporal relationship between air pollution and environmental factors, such as transportation and population. However, the analyses that have been conducted by atmospheric and environmental scientists have mainly been focused on model evaluation and spatial and temporal coefficient analysis using a statistical method statistics. Compared with the above analyses, GIS supplies a more efficient and professional kind of spatial and temporal analysis. Nowadays, more and more atmospheric and environmental scientists use GIS to improve their analysis, which shows that the weakness exists in air pollution analysis.

The above three weaknesses pose another three related challenges which are listed below, with sequential numbers to follow the above list.

(4) Integrating the platform with all simulation components

There are a lot of tools/software available today which have been developed to deal with the various steps of air pollution simulation. However, there is still no seamless, efficient and integrated system. The reason for this deficiency is that, in order to create such a platform, multi-disciplinary knowledge is needed, which would include knowledge from atmospheric science, environment science, geography, and computer technologies as a minimum. However, it is not easy for atmospheric and environmental scientists to hold sufficient knowledge and technologies beyond their own domains.

(5) Knowledge sharing and collaboration

This challenge is generated from the above four. As previously discussed, air pollution is a complex phenomenon. To achieve a comprehensive simulation of air pollution, the most reasonable method is to share multi-disciplinary knowledge and engage multi-disciplinary experts in order to conduct a collaborative simulation. In this case, each participant only works on his own research field using his professional knowledge, and leaves any unknown topics to the other participants. This working style can not only improve the rationality and efficiency of a simulation, but can also alleviate the burden on each participant.

The collaboration in air quality management that is conducted by decision makers is also required in daily life. Firstly, air pollution is a regional or even a global problem. If a city government wants to improve the local air quality, it is almost impossible for this city alone to control air pollution emissions, because mass air pollution cannot be blocked by any political boundary. The most effective way is to seek collaborative control with all the neighbouring cities. Secondly, in order to achieve an optimal resolution on air pollution control, decision makers, multi-disciplinary experts and the public should all be encouraged to present their advice, suggestions and comments. This is an example of collaboration between multiple roles.

Traditional collaborations include face-to-face discussions, telephone conversations and video/audio conferences. The advantages of these collaborations are the independent contents, efficient communication and convenience of interaction among the participants. Their disadvantages are obvious, such as the limitation on the number of participants, location conditionality and high cost-effectiveness.

The developments in computer and network communication, especially the emergence of computer supported cooperative work (CSCW), provide the opportunity for long-distance, network-based collaboration. This vastly alleviates the drawbacks of traditional ways of communication, and allows participants to carry out collaboration from "anywhere" where the network can be reached and at "anytime" that the participants prefer.

(6) High performance computation based air pollution modelling

This challenge has existed from the beginning of air pollution modelling and computation. Nowadays, high performance computers (HPC), grid computation (GC) and parallel computation, among other technologies, are rapidly advancing and improving the complexity and speed of model computation. Computer technologies are being pushed forwards. On the other hand, each forwards step in computer technologies will post its own challenges for air pollution modelling.

1.2 A Proposed Solution to the Challenges: Collaborative Virtual Geographic Environments

In order to tackle challenges outlined in 1.1, a solution is now proposed. The workflow of the air pollution simulation remains unchanged, but the components are modified as follows.

(1) Integration of the data preprocess and management

The data process here includes both the input data process and the output data process. The input data process involves converting or transforming input data to the required format that is compatible with the circulation models or air pollution dispersion models. The output data process converts or transforms the output from the air pollution dispersion models into a format which can be used for visualization and analysis.

Both the input and output data of air pollution modelling are geographically referenced with a location (x,y,z). Some data, such as the output data from air pollution dispersion, also has a time dimension. In the geographical community, the geographically referenced and time dependent data are called spatial and temporal data, or geo-data. Spatial and temporal data (geo-data) refers to the full spectrum of the digital geographic data, including digital maps, raster data (such as remote sensing data), vector data (such as roads and buildings) and other geographically referenced data (such as population distribution). Thus, the data process can be re-defined as the conversion or transformation of geo-data into formats that can be used for air pollution modelling, and the conversion or transformation of output data from air pollution modelling into geo-data. Geo-data conversion and transformation have been developed perfectly by many tools, such as ArcGIS for geographic data and the information process, and ERDAS and ENVI for the remote sensing data process. Terrain data and building data can be efficiently processed using ArcGIS, while the

LULC can be precisely calculated using ERDAS or ENVI. The integration of some of these tools' functions will not only alleviate the burden on atmospheric and environmental scientists, but will also save time in preparing air pollution modelling input data.

The input and output data from air pollution modelling and computation are large in volume. Data management attempts to organize, store and maintain the input and output data efficiently, which supplies the basis for data evaluation, analysis and visualization.

(2) Model integration

Circulation models and air pollution dispersion models are the most important bases for air pollution simulation. Many models have been developed for air pollution modelling. Model integration provides methodologies to couple models, with fore-and-after components available to manage these models. As with the database and management, models can be stored into a model base, and model management can be established on the base.

Integration and management of these models will improve the efficiency of air pollution modelling and simulation. When conducting an air pollution simulation, researchers can reference their preferred models against the model base to support their study. They can also develop new models and input these into the model base.

Model computation is also included in this component. Some of the models are complex and computationally intensive, and high performance computation will thus be addressed.

(3) Modifying visualization to geo-visualization

The display function of the traditional 2D and 3D visualizations adopted by atmospheric and environmental scientists are inherited. However, the low interaction between operators and program is modified so that it is highly interactive. The visualization is presented at the geographically referenced virtual locations, where the diagram, table and text can all be attached. This method comes under the domain of geo-visualization, which refers to a set of tools and techniques that present geo-spatial data through the use of interactive visualization.

(4) Extending analysis and model evaluation to include spatial and temporal analysis

With geo-referenced data, spatial and temporal analysis is used to discover relationships between air pollution distribution/dispersion and geographic factors. Model evaluation analysis can also be integrated with spatial and temporal analysis.

(5) Adding geo-collaboration to support multi-disciplinary knowledge sharing and collaboration

As previously stated, air pollution is complex and simulation of air pollution involves multi-disciplinary knowledge from atmospheric science, environmental science, geography, and computer technology. It is unwise for a scientist to step beyond their domain into others and attempt to handle different kinds of knowledge to support their research. A reasonable solution for this complex problem is to engage multi-disciplinary experts to join a virtual group, discuss problems, conduct a collaborative simulation and find an optimal answer. In the virtual group, participants can share their knowledge, present opinions, review others' work, post suggestions and give their comments. If geo-spatial technologies are applied to facilitate this collaboration, the result can be called geo-collaboration.

To summarize this proposed solution to the challenges of air pollution simulation, the new platform should integrate the data and the model, conduct model computation, present air pollution distribution/dispersion using geo-visualization, support spatial and temporal analysis and achieve knowledge sharing and collaboration.

Air pollution is a geography-related phenomenon, or geo-phenomenon for short. Air pollution transport and dispersion are geo-processes. Therefore, the input and output of air pollution modelling are geo-referenced, the results being called geo-data by a geographer. The circulation models and environmental models are also related to geographic locations (x,y,z) and times (t). The models are thus called geo-models. From the viewpoint of geographers, therefore, the proposed solution for air pollution simulation is a platform which, based on geo-data and a geo-model, achieves geo-knowledge sharing and geo-collaboration in a shared virtual environment, with the support of geo-visualization. This kind of platform is termed a collaborative virtual geographic environment (CVGE). The proposed components are CVGE features, which include geo-data integration, geo-model integration, geo-visualization and geo-collaboration.

The proposed solution and the CVGE formed from this proposal are shown as Figure 1.2.

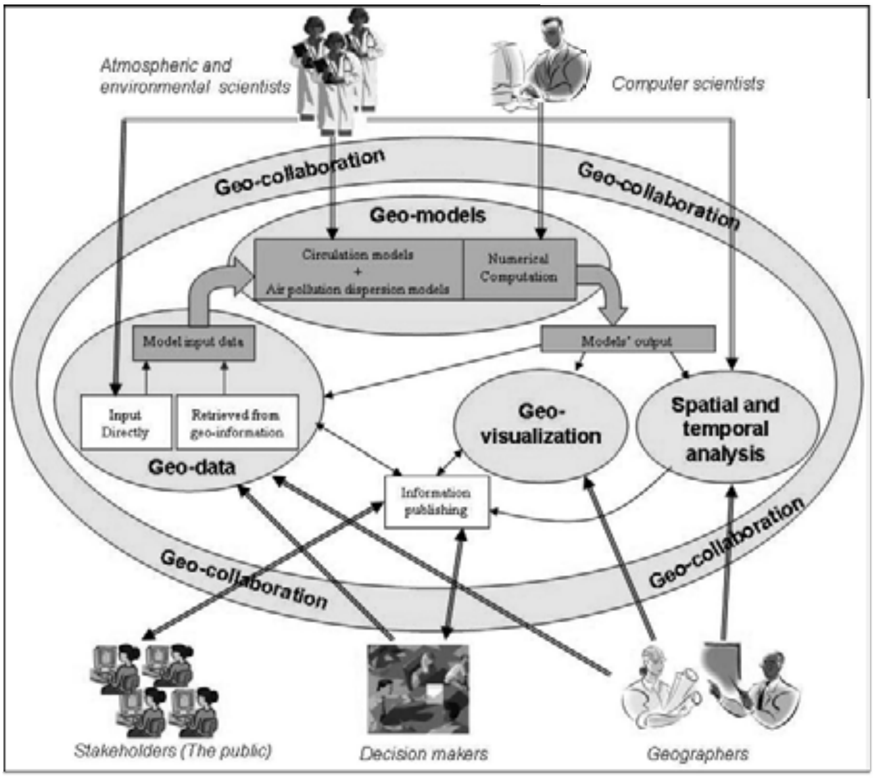


Figure 1.2: The proposed solution to the challenges of air pollution simulation and CVGE framework shadowed by the proposal

1.3 Research Objectives and Significances

1.3.1 Research objectives

The objectives are two-fold: one is on the development of a conceptual framework and prototype of CVGE from practice of air pollution simulation; the other is on applying this framework to facilitate air pollution simulation. The research attempts to integrate geo-data, geo-information and circulation and air pollution dispersion models; apply geo-visualization to existing air pollution distribution and dispersion; adopt geo-collaboration for air pollution simulation and management in its life cycle; and engage multi-disciplinary experts, decision makers and the public to carry out collaborative work from geographically dispersed locations, but connected by the network.

The objectives of this thesis are: ① to develop the conceptual framework of CVGE; ② to design the architecture of CVGE-based air pollution simulation; ③ to study the integration and computation of computational intensive air pollution modelling based on high performance computation; ④ to study geo-visualization of the output of air pollution modelling in CVGE; ⑤ to study the geo-collaboration orienting air pollution simulation in its life cycle; and ⑥ to apply the CVGE prototype and conduct collaborative air pollution simulation in its life cycle. The first objective is the CVGE theory approach. Objectives two to five are developing the prototype of CVGE based on air pollution simulation. Of these four objectives, objective two addresses the overall architecture, which organizes the three key issues contained in objective three, objective four and objective five. Objective six have two functions: one is to serve as a case study to test the feasibility and efficiency of the CVGE prototype; the other one is to serve as an implemented solution to the challenges of air pollution simulation that were highlighted at the beginning of this thesis.

These objectives are selected because they address key issues of both air pollution simulation and the prototype of CVGE. The details of each objective can be described as follows.

(1) Developing the conceptual frameworks of CVGE

The concept of CVGE will be defined; the conceptual framework and system framework of CVGE will be designed; and the theories of CVGE shadowed by the designed frameworks will be discussed.

(2) Designing the architecture of the prototype for CVGE-based air pollution simulation

The prototype system requirements will be analyzed; the functions will be defined to match these requirements; the architecture will be designed to organize the functions; and the key issues for implementing the prototype system will be highlighted.

(3) Integration and computation of air pollution modelling based on high performance computation

Two scales of air pollution modelling will be considered. One is the regional scale, and the other is the local scale. For regional scale air pollution modelling, MM5 (Fifth-Generation NCAR / Penn State Mesoscale Mode) (Dudhia et al. 2005, GA, J and DR 1994) will be selected as the atmospheric circulation model; and SYSUM (the Sun Yat-sen University Model) (Wenshi et al. 2009) is selected as the air pollution dispersion model. SYSUM and MM5 are seamlessly coupled. The output of MM5, together with emissions' source data, will serve as the input for SYSUM. For local scale air pollution modelling, the Gaussian Plume Model (GPM) will be selected as the air pollution dispersion model. The atmospheric conditions, such as wind speed and wind direction, will be input manually into GPM.

MM5 are complicated and computationally intensive. Therefore, the first work towards this objective is to minimize the computation time of MM5. The second task is to integrate MM5 and SYSUM into CVGE.

Although simple, GPM does face the problem of computational efficiency when it is used to calculate air pollution distribution for a large number of pollution sources. Therefore, the third part of this objective is to design parallel arithmetic to decrease the computation time of GPM.

(4) Multi-dimension geo-visualization of geo-data and air pollution distribution/dispersion

Diagrams, 2D and 3D are used to visualize geo-data, geo-information and air pollution distribution/dispersion at geographically referenced locations. The original format of the air pollutant calculated by MM5 and SYSUM cannot be read by CVGE directly. Thus, the first work for this objective is to convert (reformat and re-project) the original air pollution data into a CVGE readable format. The second is to visualize geo-data, geo-information and air pollution distribution/dispersion geographically in 2D and 3D environments. Thirdly, the term Fuzzy Boundary Volume Object (FBVO) is coined to match the features of the air pollution mass, and the method for visualizing FBVO in CVGE is studied in a 3D environment.

(5) Geo-collaboration for air pollution simulation

This objective includes the following work: ① defining what contents may be targeted for geo-collaboration for air pollution simulation; ② studying methods to achieve collaboration on each element; ③ studying specific methods of conflict detection and resolution; and ④ implementing geo-collaboration for air pollution simulation.

(6) Prototype based collaborative air pollution simulation

This objective attempts to conduct simulation not only to test the feasibility and efficiency of CVGE, but also as a response to the problems which were highlighted at the beginning of the thesis.

Two simulations will be conducted. The first is an episode simulation for the Pearl River Delta (PRD) region, which is based on MM5 and SYSUM. This will be used to test a collaborative workflow engaging multiple participants from geographically dispersed locations. The simulation covers all components of air pollution simulation, which include compiling collaborative air pollution sources, operating collaborative MM5 and SYSYM, collaborative geo-visualization of air pollution dispersion and collaborative analysis. Air pollution model evaluation is also studied in this simulation. The second simulation uses the Gaussian plume model to study a hypothetical accident in Hong Kong. This simulation also covers all the components of air pollution simulation. What is more, this simulation will also test conflict and negotiation during model collaborative operation.

1.3.2 Significances

The work is significant from two perspectives: collaborative air pollution simulation and CVGE. From the perspective of air pollution, this research is an innovation in its research methodology, and supplies a new research platform for air pollution simulation. From the perspective of CVGE, this research is a comprehensive research into CVGE. The research touches on four key points of CVGE: definition, framework, theories and implementation. Thus, the research will pave the way for further research in CVGE. The details of these significances are as follows.

(1) **A new platform to facilitate air pollution simulation and support decision making:** The new platform provides a new research method which is collaborative, integrated with geography and oriented towards multiple participants. The integration with geography will benefit the extraction of more geographic factors to support air pollution modelling, geographical visualization

of air pollutants at virtual geographical locations and the application of spatial and temporal analysis to air pollution distribution and dispersion.

(2) **Bridge air pollution to other research areas:** This research combines air pollution simulation with CVGE. The combination stamps geographic location (X,Y,Z) and time (t) on the results of air pollution modelling, which is a powerful link that can bridge air pollution with a lot of other research areas, such as the economy, public health and urban planning.

(3) **The geo-collaboration will benefit not only air pollution simulation, but also other complicated geo-process research and management:** Geo-collaboration for air pollution modelling and management involves multiple roles (experts, officers and the public) and multiple levels (political collaboration, multi-role collaboration and multi-disciplinary collaboration). The geo-collaboration will benefit not only scientific research on air pollution simulation, but also governmental management and control of air pollution emissions. Geo-collaboration is also a reasonable method of finding the optimal solution. Although directed at air pollution simulation, the geo-collaboration studied here may be used to solve similar kinds of complicated geographic processes.

(4) **Although this study is oriented towards a typical problem, the exploration of CVGE, the architecture and the three key issues are all valuable for future research on CVGE.**

(5) **The integration and computation of MM5 based on high performance computation not only contributes to decreasing the computation time for air pollution modelling, but is also valuable in collecting experiences for further research on other geo-models' integration and improving computation efficiency in CVGE.**

(6) **Multi-dimension geo-visualization of air pollution distribution/dispersion can be helpful for users understanding the air pollution phenomenon. The proposed concept of the fuzzy boundary volume object and its high dimension visualization not only make air pollution mass easier to understand for the public, but also extends the possibilities for presenting**

geo-processes in CVGE.

1.4 Thesis Organization

This thesis is composed of nine chapters. The first chapter is the introduction, which includes research background, questions, objectives and significances. The second chapter reviews literatures. The third chapter defines CVGE and develops frameworks of CVGE. Chapter 4 analyzes the requirements of multiple participants, defines the functions to match the requirements, and designs architecture to organize these functions to form a prototype. Following three chapters, chapter 5, chapter 6, and chapter 7, are aimed to study three key issues to support creating the prototype of CVGE, which are air pollution models integration and computation based on high performance computation (chapter 5), multi-dimensional geo-visualization on output of air pollution modeling (chapter 6), and geo-collaboration for air pollution simulation (chapter 7) . Chapter 8 conducts the simulation to test the feasibility and efficiency of CVGE prototype. The geo-collaboration for air pollution simulation is also discussed in chapter 8. The final chapter (chapter 9) is the conclusion. The organization of this thesis can be summarized as figure 1.3.

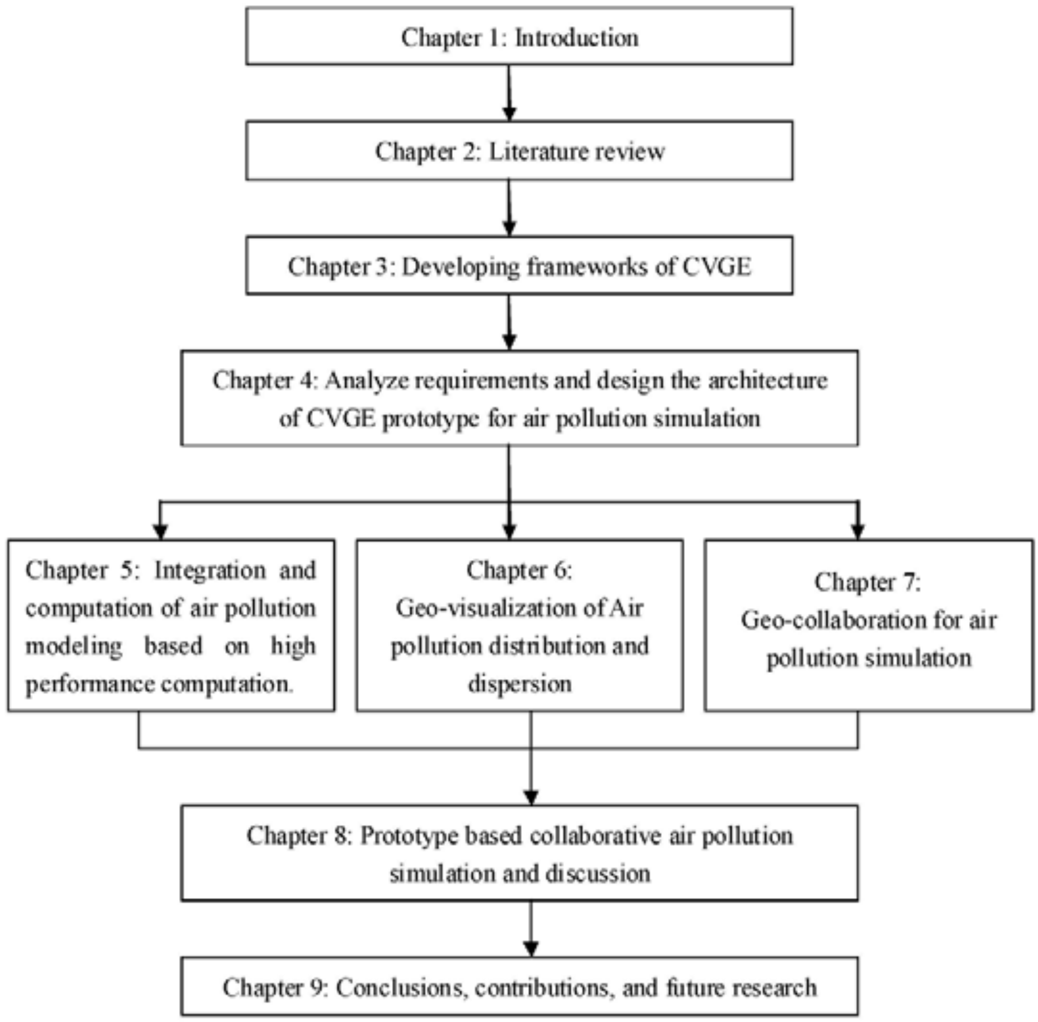


Figure 1.3: Thesis organization

CHAPTER 2: LITERATURE REVIEW

The research objectives of this thesis involve three research areas: air pollution simulation, support for computer technologies and geography. The literature review will therefore cover these three areas. The literature review on computer technologies will focus on those which can be helpful for creating a CVGE system, particularly looking at Virtual Reality, virtual environments, network communication, computer supported cooperative work (CSCW), collaboration and Synthetic Environment Data Representation and Interchange Specification (SEDRIS). The literature review on air pollution simulation will focus on current tools which are applied to facilitate air pollution visualization, modeling, and analysis, and reveal the shortcomings that exist in these tools. The literature review on geography will focus on GIS-based modelling, geo-visualization and geo-collaboration.

2.1 Related Computer Technologies

This part reviews the state-of-the-art computer technologies which will benefit creating a CVGE system.

2.1.1 Virtual reality and virtual environments

Virtual Reality (VR) is one key technology to support creating CVGE. VR allows a user to interact with computer-simulated environments, and is often used in wide variety of applications, commonly associated with their immersive, highly visual, 3D environments. Most current virtual reality environments are primarily displayed either on a computer screen (Oppenraaij. et al. 2009) or through special or stereoscopic displays (Sousa Santos et al. 2009, DeFanti et al. 2009). Some advanced haptic systems include tactile information(Reif and Walch 2008). Users can interact with a virtual environment or a virtual artifact (VA) either through the use of standard input devices such as a keyboard and mouse, or through multimodal devices such as a wired glove(Reif and Walch 2008). The research contents of VR can be concluded as ① Interface investigation and ② virtual scene generation, rendering and display.

Currently, there are many famous VR systems. In 1991, Cave Automatic Virtual Environment (CAVE) was invented at the Electronic Visualization Lab in University of Illinois (www.evl.uic.edu/pape/CAVE/). More recent VR systems based on the CAVE are the ImmersaDesk (www.evl.uic.edu/pape/CAVE/idesk/) and the IWall. Another display device created for virtual reality is GeoWall, which is a practical 3D display system that combines advanced graphics PCs and specially calibrated digital projectors to display three-dimensional presentations with ease and versatility(Johnson et al. 2006b). Using 3D and high-resolution technology, GeoWall allows audiences to view and interact with spatially immersive content.

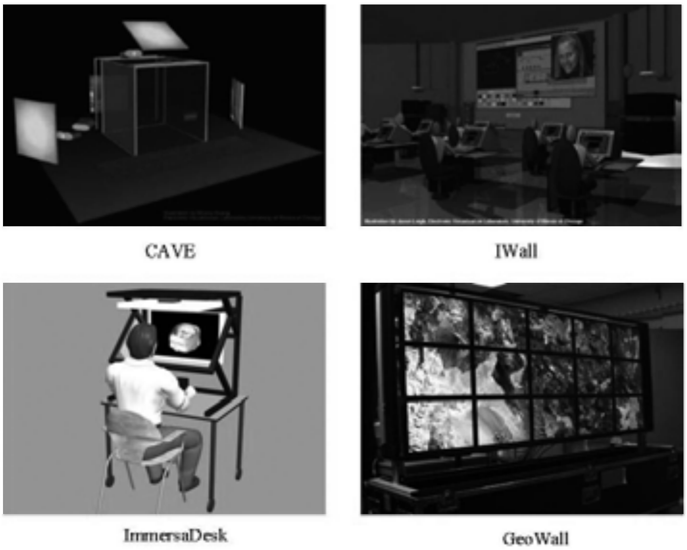


Figure 2.1: Virtual reality system (http://www.evl.uic.edu/, access date 2009.4)

The immersion of VR depends on not only the display, but also the interaction between computer and human. The interaction devices have evolved from mouse, keyboard, to joy stick, digital glove, digital helmet, and now even brain (Lecuyer et al. 2008).

Figure 2.2: Virtual reality interaction devices

Virtual reality has been widely used, such as in entertainment (Wang 2007, Fraser 2006, Hansson 2000), education(Sun 2008, Chen et al. 2007, Sampaio, Henriques and Ferreira 2006, Odeyale and Balogun 2006, Kaufmann and Schmalstieg 2006), medicine (Chou and Handa 2006, Battersby, Brown and Powell 2000), simulation (Ruthenbeck, Owen and Reynolds 2008, Chai et al. 2008, Bal and Hashemipour 2007, Zhou, Blum and Eskandarian 2005), scientific visualization(Peterka et al. 2006, van Dam, Laidlaw and Simpson 2002, Barcelo 2001), military(Thrush and Bodary 2000), aerospace (Freund, Rossmann and Schluse 2000), manufacturing (Manesh, Hashemipour and Bal 2008, Bal and Hashemipour 2007, Okulicz 2004, Ou, Sung and Hsiao 2002), art (Larsson, Vastfjall and Kleiner 2001), and so on.

Nowadays, Virtual Environments (VE) are popular for high dimensional visualization applications. VE are created based on VR technologies and sciences of applied fields. VE have been widely used in scientific research, such as exploring data (Ai et al. 2005), remotely testing complex systems(Cristaldi et al. 2005), training apprentice(Record 2005, Beavis et al. 2006, Boejen et al. 2007, Limniou, Roberts and Papadopoulos 2008), treating patient(Edmans et al. 2006), to help design(Fortner 2006), supporting decision making (Varela and Soares 2007, Bishop, Stock and Williams 2009), and so on. The online games based on virtual environments are flourishing. The most popular among them are Active Worlds(www.activeworlds.com, access Jun. 2009), Second Life (www.secondlife.com, access Jun. 2009), Barbie Girls (www.barbiegirls.com, access Jun. 2009), Club Penguin(www.clubpenguin.com, access Jun. 2009), Forterra Systems (www.forterrainc.com, access Jun. 2009), Gaia Online (www.gaiaonline.com, access Jun. 2009), Habbo Hotel (www.habbo.com, access Jun. 2009), Kaneva (www.kaneva.com, access Jun. 2009), Neopets (www.neopets.com, access Jun. 2009), Teen Second Life (www.teen.secondlife.com, access Jun. 2009),The Sims Online (thesims.ea.com, access Jun. 2009), There (www.there.com, access Jun. 2009), Webkins (www.webkinz.com, access Jun. 2009), Whyville (www.whyville.net, access Jun. 2009), Zwinktopia (www.zwinky.com, access Jun. 2009) and so on.

2.1.2 Network communication

The second important technology to support creating CVGE is net communication, especially the efficient transport protocols. There are several network protocols widely used to transfer data on network. HTTP (the Hypertext Transfer Protocol) and FTP (the File Transfer Protocol), which are based on TCP/IP （Transmission Control Protocol/Internet Protocol) direct connections, are the well established protocols for reliable file transfers over the Internet. In contrast to TCP/IP based protocols, UDP (User Datagram Protocol) datagram, which is called a connectionless service, does not establish a connection between the sender and the recipient of a network package. IP multicasting is the protocol to be scalable to a large number of distributed users.

Based on the above protocols, some deducted protocols come forth to meet specific requirements, such as selective reliable transmission protocol (SRTP) for large scale distributed interactive simulation (DIS), reliable multicast transport protocol (RMTP) to meet a multicast tree, selective reliable multicast (SRM) for distributed whiteboard application, smooth synchronous collaboration transport protocol (S-SCTP) to address the delay and the jitter problems (Boukerche, Maamar and Hossain 2008).

To optimize network structure and decrease data running online, some technologies are adopted, such as peer to peer (or P2P) and mobile agent (MA). P2P network does not have the notion of clients or servers but only equal peer nodes that simultaneously function as both "clients" and "servers" to the other nodes on the network. P2P networks are typically used for connecting nodes via largely ad hoc connections. An important goal in P2P networks is that all clients provide resources, including bandwidth, storage space, and computing power. Thus, as nodes arrive and demand on the system increases, the total capacity of the system also increases. There are some cases to discuss using P2P in virtual environments (Zhu et al. 2007b, Romain, Christian and Jerome 2006, Gupta and Kaiser 2005). Mobile agent (MA) can be used to decrease data transferring on network because, unlike transforming huge data from data server, MA is able to migrate (move) from one computer to another autonomously and continue its execution on the destination computer (Lin et al. 2006b, Neo, Lin and Liew 2005, Loo, Ly and Tang 2000). But MA will bring threat to system security (Topaloglu and Bayrak 2008).

Currently, Wireless Fidelity (WiFi) based communications have caught researchers' attentions. More and more WiFi based applications are launched. For example, WiFi is for indoor positioning and navigation (Ocana et al. 2005, Evennou and Marx 2006, Lim et al. 2007), disaster management(Andrade, Palmer and Lenert 2006), traffic control (Daoud et al. 2007), and so on. WiFi is valuable for forming a mobile geo-collaboration for CVGE.

2.1.3 Computer supported cooperative work and Groupware

The term computer supported cooperative work (CSCW) was first coined by Irene Greif and Paul M. Cashman in 1984 (Grudin 1994). CSCW is the study of how people work together using computing and communication technologies, and is broadly interdisciplinary, drawing from computer science, management information systems, information science, psychology, sociology, and anthropology(Olson and Olson 1999).

Groupware, which is often considered synonymous with CSCW, is computer software that functions to provide a means for human collaboration across time and space (Lococo and Yen 1998). Groupware is typically classified according to two dichotomies, viz. (1) whether its users work together at the same time (synchronous) or at different times (asynchronous) and (2) whether they work together in the same place (co-located) or in different places (dispersed) (ter Beek et al. 2009). The functionalities of Groupware were presented by Khoshafian and Buckiewicz as Figure 2.3.

	Same Place	**Different Places**
Same Time *Synchronous*	Electronic Meetings Team Rooms Group Decision Support Systems Electronic Whiteboards	Video Conferencing Teleconferencing Screen Saving Document Sharing Electronic Whiteboards
Different Times *Asynchronous*	Shared Containers Mailboxes Electronic Bulletin Boards Virtual Rooms, Kiosks Document Management Systems	Electronic Mail Workflow Form Flow Messaging Routing & Notification

Figure 2.3: Groupware functionalities (Khoshafian and Buckiewicz 1995)

A complete groupware infrastructure has four dimensions: communication (pushing or pulling information), collaboration (shared information leading to shared understanding), coordination (delegation of task, sequential sign-offs, etc.), and control (management of conflict) (Pendergast and Hayne 1999). Besides these four dimensions, another two issues are also important to achieve the success of groupware. One is the concurrency control and user awareness, which the latter should be understood as users having a sense of the (past, current, future) activities of other users (ter Beek et al. 2009). The other one is psychological safety, which is a sense of interpersonal trust and being valued in a work team, is an important determinant of groupware technology adoption in an educational setting. (Schepers et al. 2008)

2.1.4 Collaboration

Collaboration is defined by WiKi "is a recursive process where two or more people or organizations work together toward an intersection of common goals by sharing knowledge, learning and building consensus" (http://en.wikipedia.org/wiki/ Collaboration, access date 25th Apr. 2009). WiKi goes on that "Collaboration does not require leadership and can sometimes bring better results through decentralization and egalitarianism. In particular, teams that work collaboratively can obtain greater resources, recognition and reward when facing competition for finite resources". David Coleman and Stewart Levine draw a conclusion on definitions of collaboration, and give their opinion that "collaboration is primarily about people, about trust, and about the willingness to share information and work in a coordinated manner to achieve a common goal"(Coleman and Levine 2008). Collaboration acts in almost all domains, such as economy, management, education, research, and has several types, for instance face to face, telephone, and network. This review is only focused on the collaboration related with virtual environments based on medium of network.

Both as entertainment and as a problem-solving tool, collaboration in technology encompasses video games(Rhee, Park and Kim 2006), distributed computing, knowledge sharing(Arora, Dans and Lansang 2007, Santoro, Borges and Rezende 2006, Kotlarsky and Oshri 2005) and communication(Austerlitz and Sachs 2006) tools. Collaboration in the technology sector refers to a wide variety of tools that enable groups of people to work together, and encompasses both asynchronous(Dewiyanti, Brand-Gruwel and Jochems 2005) and synchronous(Tandler 2004) methods of communication and serves as an umbrella-term for a wide variety of software packages.

Collaboration has been widely used. Collaboration based education and training between teachers and students take big proportion of collaboration applications (Mockus 2008, Zampardi et al. 2007, Urquhart, Cox and Spink 2007, Phillips 2007, Mockus 2006, Bouras et al. 2006). Another big proportion of collaboration applications is in industrial design (Pappas et al. 2006, Tann and Shaw 2007, Kao et al. 2004, Eynard et al. 2005). Other applied collaborations are in medical research (Alberola et al. 2000), conferences(Fukui, Honda and Okada 2004, Jin et al. 2005), tour(Strouse 2004), scientific researches(Kim, Kuester and Kim 2005), and so on.

For the trends of collaboration, David Coleman and Stewart Levine give ten points conclusions(Coleman and Levine 2008), which are ① Convergence of audio, video, and web conferencing; ② Presence everywhere; ③ Integration of synchronous and asynchronous; ④ Collaborative consolidation in enterprise; ⑤ Pushing collaboration into the infrastructure; ⑥ Market consolidation; ⑦ Driving collaboration into verticals and critical processes; ⑧ Changing distribution channels for collaboration; ⑨ Changing buyer for collaboration; and ⑩ Mobile collaboration.

2.1.5 Synthetic Environment Data Representation and Interchange Specification (SEDRIS 2008)

Data in CVGE is heterogeneous and distributed. Data interoperation, data collaboration, data sharing, data representation, and data interchange are all core technologies to implement CVGE. These five functions can get valuable references from Synthetic Environment Data Representation and Interchange Specification (SEDRIS).

SEDRIS main sponsors are Defense Modelling and Simulation Office (DMSO) and Simulation Coordination Office (SCO). It was initiated in 1994 and currently has been standardized by International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC).

SEDRIS is an infrastructure technology that enables information technology applications to express, understand, share, and reuse environmental data. SEDRIS technologies provide the means to represent environmental data (terrain, ocean, air and space), and promote the unambiguous, loss-less and non-proprietary interchange of environmental data.

SEDRIS is fundamentally about: (a) the representation of environmental data, and (b) the interchange of environmental data sets. SEDRIS offers a data representation model, augmented with its environmental data coding specification and spatial reference model, so that one can articulate one's environmental data clearly, while also using the same representation model to understand others' data unambiguously. Therefore, the data representation aspect of SEDRIS is about capturing and communicating meaning and semantics. For the interchange part, the SEDRIS API, its format and all the associated tools and utilities play the primary role, while being semantically coupled to the data representation model. The representational aspect of SEDRIS is much like a language or a method for unambiguously describing the environment, independent of whether the environment is geo-specific, geo-typical or completely fictitious. And the interchange aspect is a mechanism for sharing the described environmental data.

The objectives of SEDRIS are (a) to provide a powerful method for articulating and capturing the complete set of data elements, and the associated relationships, needed to fully represent environmental data; (b) to provide a standard interchange mechanism to distribute environmental data and to promote database reuse among heterogeneous systems; and (c) to support the full range of applications across all environmental domains that span ocean, terrain, atmosphere and space.

SEDRIS has five core technology components, which are the SEDRIS Data Representation Model (DRM), the Environmental Data Coding Specification (EDCS), the Spatial Reference Model (SRM), the SEDRIS interface specification (API), and the SEDRIS Transmittal Format (STF). DRM, EDCS, and SRM are used to achieve the unambiguous representation of environmental data. The DRM, the EDCS, and the SRM are used to capture and communicate meaning and semantics about environmental data. The SEDRIS API and the STF allow the efficient sharing and interchange of the environmental data represented by the other three components.

2.2 GIS-based Modelling, Geo-Visualization, and Geo-collaboration

2.2.1 GIS-based modelling

Michael F. Goodchild listed three reasons for urging that GIS evolve into an effective platform for spatial modelling(Goodchild 2005). First, GIS provides an excellent environment for representing spatial variation in the initial and boundary conditions of models and their outputs. GIS also includes numerous tools for acquir-ing, preprocessing, and transforming data for use in modelling, including data management, format conversion, projection change, re-sampling, raster-vector conversion, etc. It also includes excellent tools for displaying, rendering, querying, and analyzing model results and for assessing the accuracies and uncertainties associated with inputs and outputs. Secondly, much progress has been made recently in the handling of time in GIS. Third, and perhaps most important, many of the techniques used in GIS analysis would be much more powerful if they could be coupled with an extensive toolkit of methods of simulation.

GIS has been widely used by a lot of research areas for supporting their modelling. GIS can be used to modelling static phenomena with static models whose input and the output both correspond to the same point in time, or dynamic process with dynamic models whose output represents a later point in time than the input (Goodchild 2005). The static spatial modelling supported with GIS can be referenced from works done by Thompson, Carozza, and Sandoval-Ruiz (Thompson, Carozza and Zhu 2008, Sandoval-Ruiz, Zumaquero-Rios and Rojas-Soto 2008), while the dynamic modelling can be referenced from works done by Batty and Balram (Balram and Dragicevic 2003, Batty 2005). Models can be discrete and continuous, which are also reflected in GIS supported spatial modelling. Agent based modelling coupled with GIS(Benenson 2003) is a example of discrete modelling. The GIS supported fluvial erosion modelling (Finlayson and Montgomery 2003) is a example of continuous modelling.

The one-size-fits-all approach that is inherent in GIS and in systems is unlikely ever to address all possible needs, and instead much attention has been devoted to coupling GIS with packages that are more directly attuned to the needs of modelling (Goodchild 2005). Coupling is also widely used to link standalone models to GIS, and there are three types of coupling(Goodchild 2005): ① a standalone package might be coupled with GIS by exchanging files;② coupling may take the form of integrating the GIS with the modelling packages using standards that allow a single script to invoke commands from both packages; and ③ the entire model may be executed by calling functions of the GIS, using a single script. Maguire et al. (Maguire, Batty and Goodchild 2005) suggests three similar approaches as loose coupling(employs common file structures, file translators, and more recently, Web services messaging), moderate integration (uses techniques such as remote procedure calls and shared database access) and tight integration(can be achieved by, for example, object-component calls, or function calls).

There are some challenges for GIS supported spatial modelling as Goodchild highlighted (Goodchild 2005). First, “static modelling and the calculation of indicators are classic GIS applications and are well suited to traditional GIS architecture”, but dynamic modelling is not well solved. Second, “modelling in GIS raises a number of important issues, including the question of validation, the roles of scale and accuracy, and the design of infrastructure to facilitate sharing of models”.

2.2.2 Geo-visualization

The term geographic visualization, which is currently shortened as geo-vocalization and as well as the related cartographic visualization, was prompted by a 1987 National Science Foundation report on visualization in scientific computing (MacEachren et al. 2004). In 1995, the International Cartographic Association (ICA) established a Commission on Visualization, which expanded its focus in 1999 to visualization and virtual environments. This commission has played an important role in stimulating geo-visualization research and in articulating an international, interdisciplinary research agenda(MacEachren et al. 2004). MacEachren and Kraak (2001) explained the reasons of “Why focus on “geo” visualization” as follows (Maceachen and Kraak 2001): ① Estimates suggest that 80% of all digital data generated today include geospatial referencing (e.g., geographic coordinates, addresses, postal codes, etc.); ② But these magnitude and complexity of data sets that can be brought together through their common geospatial links present an extraordinary challenge for information science. Meanwhile, existed visualization is not being taken advantage of to exploit the full potential of geospatial data; ③ Geo-visualization has the potential to provide 'windows' into the complexity of phenomena and processes involved, through innovative scene construction, virtual environments, and collaboration, thus prompting insight into the structures and relationships contained within these complex, linked datasets.

Geo-visualization was defined by MacEachren and Kraak that integrates approaches from visualization in scientific computing, cartography, image analysis, information visualization, exploratory data analysis, and geographic information systems to provide theory, methods, and tools for visual exploration, analysis, synthesis, and presentation of geospatial data (any data having geospatial referencing) (Maceachen and Kraak 2001). Dodge et al. (2008) viewed geo-visualization as the application of any graphic designed to facilitate a spatial understanding of things, concepts, conditions, processes or event in the human world (Dodge, McDerby and Turner 2008).

The representation ways to achieve geo-visualization are variety, such as traditional map (Xavier et al. 2007, Chertov et al. 2005), tables (Kraak 2003), diagrams, and 3D or even high-dimensional environments (Ohno and Kageyama 2007). These ways can be summarizes as Figure 2.4. The devices which are used to support geo-visualization have been extended from desktop to PDA and mobile phone(Lin et al. 2006a).

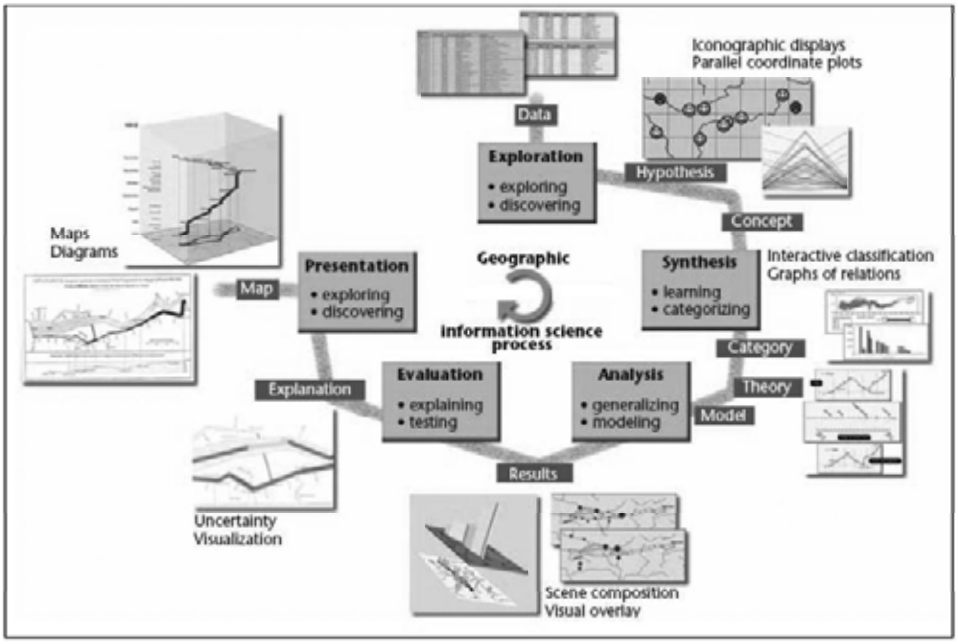


Figure 2.4: The functionalities and present ways of geo-visualization (Kraak 2006)

The applications of geo-visualization are as widely as those of GIS, such as in public health (Bhowmick et al. 2008, Xavier et al. 2007), history, economy, human behaviors, decision making (Andrienko et al. 2007) and so on. The applications also promote challenges for further research, some of which are illustrated in paper of "Research Challenges in Geovisualization" (Maceachen and Kraak 2001) on four primary themes: representation, integration with knowledge construction and geo-computation, interface design, and cognitive-usability issues.

2.2.3 Geo-collaboration

In 2000, MacEachren published his research discussed on maps and GISystems as mediators of human–human same-place collaboration at both same and different times by review on spatial decision-support systems (SDSS), public-participation GIS (PPGIS), and new software and display forms to facilitate group work (MacEachren 2000). In the succeeding year, he published another paper to delineate the potential and challenges of different-place collaboration facilitated by geo-information technologies (MacEachren 2001). In this paper, MacEachren summarized that "most of the work on environments to support asynchronous geo-collaboration has occurred within the context of PPGIS", while "synchronous different-place geo-collaboration can be subdivided into two categories (having potential for overlap): geospatial group decision support" which is "dominated by geographers" and "support for science" which is "cross-disciplinary (with geographers in the minority)". Three human aspects for collaboration at a distance were pointed either in this paper (MacEachren 2001) that: ① perspectives and negotiation; ② representing participants and facilitating joint behavior; ③ usability of tools and environments. Based on above discussions, in 2003, MacEachren et al defined geo-collaboration from two aspects of activity and field of research as those: ① "As an activity", geo-collaboration is "group work about geographic scale problems facilitated by geospatial information technologies"; ② "As a field of research", geo-collaboration is "the study of these group activities, together with the development of methods and tools to facilitate them"(MacEachren et al. 2003).

The research and activities before 2001 had been reviewed in a pair of papers (MacEachren 2001, MacEachren 2000). After that, geo-collaboration is directly applied as research topic or research support by many researchers. Cai et al. (Cai et al. 2004, Cai et al. 2005, MacEachren et al. 2006) used maps to mediate EOC(emergency operation center)–mobile team collaboration for crisis management. MacEachren et al. (MacEachren et al. 2003) developed same-time, same-place group work environments that mediate distributed thinking and decision-making through use of large-screen displays supporting multi-user, natural interaction. MacEachren and Brewer (MacEachren and Brewer 2004) developed a conceptual framework for visually-enabled geo-collaboration with six dimensions identified from human components (problem context, collaboration tasks, and perspective commonality) and computing infrastructure (spatial and temporal context, interaction characteristics, and tools to mediate group work). MacEachren and Brewer also pointed out the challenges on geo-collaboration as ① to develop a theoretical understanding of the cognitive and social aspects of both local and remote collaboration mediated through display objects in a geospatial context; ② to develop approaches to multi-user system interfaces that support rather than impede group work; ③ to understand the ways in which characteristics of methods and tools provided to support collaboration influence the outcome of group work; and ④ to initiate a parallel, concerted effort focused on integrating, implementing, and investigating the role of the visual, geospatial display in collaborative science, education, design, and group decision support system. Schafer et al. presented a software architecture that facilitates the development of geo-collaboration solutions(Schafer, Ganoe and Carroll 2007). Christopher carried a research on "utility assessment of a map-based online geo-collaboration tool"(Sidlar and Rinner 2008).

2.3 Tools to Facilitate Air Pollution Simulation

Many tools have been applied to visualize air pollution distribution and dispersion. Even though the functions are limited, these tools have done contributions to facilitate air pollution modelling, visualization and analysis.

2.3.1 Grid Analysis and Display System(Doty 1995)

The Grid Analysis and Display System (GrADS) is an interactive desktop tool that is used for easy access, manipulation, and visualization of earth science data. The format of the data may be either binary, GRIB, NetCDF, or HDF-SDS (Scientific Data Sets). GrADS has been implemented worldwide on a variety of commonly used operating systems and is freely distributed over the Internet.

GrADS uses a 4-Dimensional data environment: longitude, latitude, vertical level, and time. Data sets are placed within the 4-D space by use of a data descriptor file. GrADS interprets station data as well as gridded data, and the grids may be regular, non-linearly spaced, Gaussian, or of variable resolution. Data from different data sets may be graphically overlaid, with correct spatial and time registration. Operations are executed interactively by entering FORTRAN-like expressions at the command line. A rich set of built-in functions are provided, but users may also add their own functions as external routines written in any programming language.

Data may be displayed using a variety of graphical techniques: line and bar graphs, scatter plots, smoothed contours, shaded contours, streamlines, wind vectors, grid boxes, shaded grid boxes, and station model plots. Graphics may be output in PostScript or image formats. GrADS provides geophysically intuitive defaults, but the user has the option to control all aspects of graphics output.

GrADS has a programmable interface (scripting language) that allows for sophisticated analysis and display applications. Use scripts to display buttons and drop menus as well as graphics, and then take action based on user point-and-clicks. GrADS can be run in batch mode, and the scripting language facilitates using GrADS to do long overnight batch jobs.

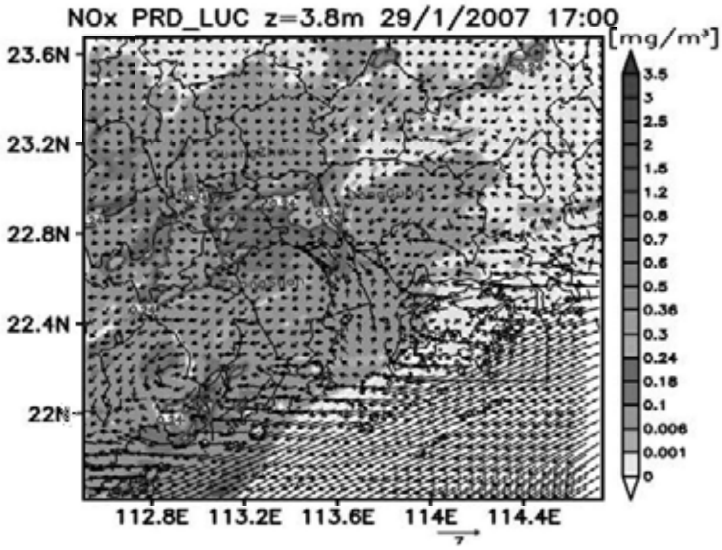


Figure 2.5: An example of air pollution (NO_x) distribution display in GrADS(Zeng et al. 2009)

2.3.2 Vis5D(Hibbard et al. 1999)

Vis5D is a software system that can be used to visualize both gridded data and irregularly located data. Sources for this data can come from numerical weather models, surface observations and other similar sources. Vis5D can work on data in the form of a five-dimensional rectangle. That is, the data are real numbers at each point of a "grid" which spans three space dimensions, one time dimension and a dimension for enumerating multiple physical variables. Of course, Vis5D works perfectly well on data sets with only one variable, one time step (i.e. no time dynamics) or one vertical level.

However, data grids should have at least two rows and columns. Vis5D can also work with irregularly spaced data which are stored as "records". Each record contains a geographic location, a time, and a set of variables which can contain either character or numerical data.

A major feature of Vis5D is support for comparing multiple data sets. This extra data can be incorporated at run-time as a list of *.v5d files or imported at anytime after Vis5D is running. Data can be overlaid in the same 3-D display and/or viewed side-by-side spread sheet style. Data sets that are overlaid are aligned in space and time.

In the spread sheet style, multiple displays can be linked. Once linked, the time steps from all data sets are merged and the controls of the linked displays are synchronized. The Vis5D system includes the vis5d visualization program, several programs for managing and analyzing five-dimensional data grids, and instructions and sample source code for converting your data into its file format.

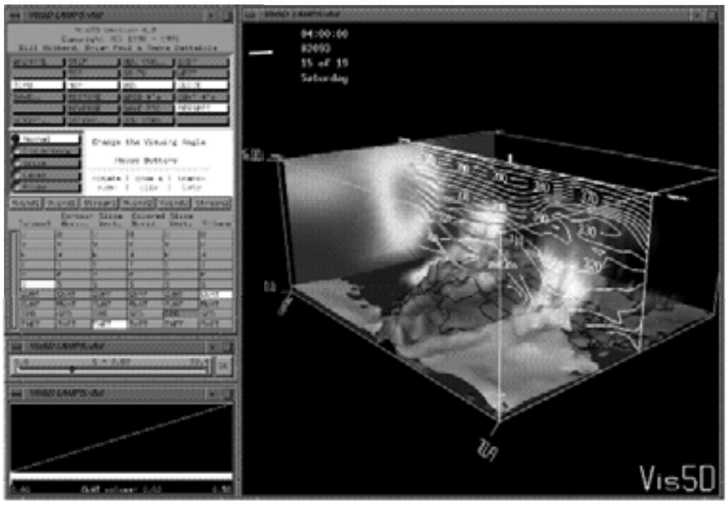


Figure 2.6: Visualization effects in Vis5D (www.ssec.wisc.edu/~billh/vis5d.html, access date May 2009)

2.3.3 Environmental Visualization System

Environmental Visualization System (EVS), or currently version of EVS Pro, is C Tech's most popular product for not only analysis, visualization and animation, but also advanced gridding, model building, output options, geostatistics capabilities, animation and GIS functions to accommodate litigation support, public relations and the more demanding requirements of earth science professionals. EVS Pro supports volume rendering. Following figure shows an example created by Prof. Carlos Borrego that a true 3D street-canyon model coupling boundary layer flow with Lagrangian dispersion to simulate urban air pollution in city centers.

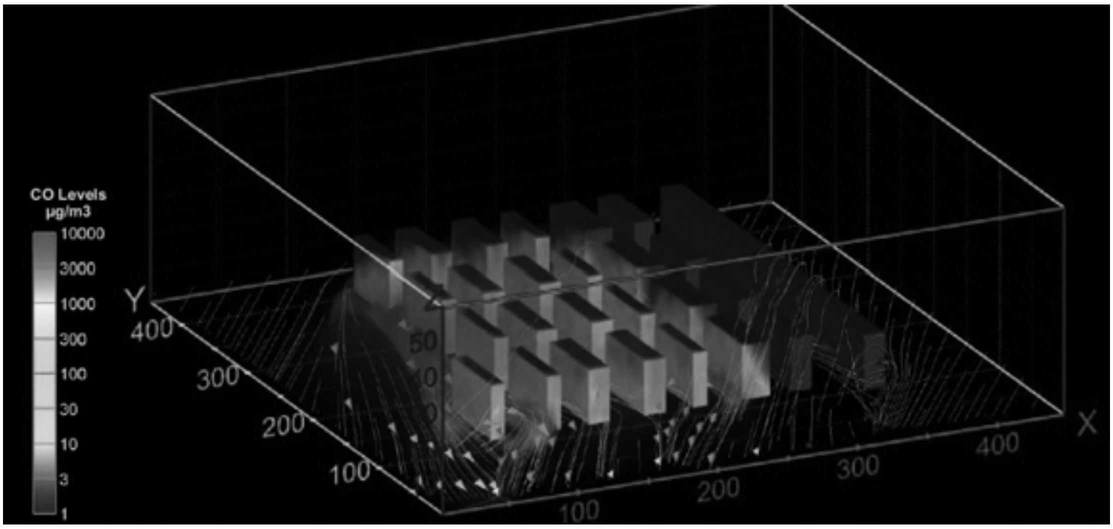


Figure 2.7: Air pollution visualization in EVS (www.ctech.com/index.php?page=pollution, access date May 2009)

2.3.4 Lakes Environmental Software

Lakes Environmental Software (LES) tries to supplying robust and easy-to-use air dispersion modelling software to consulting companies, industries, governmental agencies and academia. LES includes many tools for air dispersion modelling, risk assessment, emergency release, emissions management, compliance assurance model, and some free utilities. The air dispersion modelling has implemented ISC-AERMOD View for Gaussian plume air dispersion model, CALPUFF View for puff long range air dispersion Model, AUSTAL View for Lagrangian particle tracking air dispersion model, CALRoads View for traffic air dispersion model, and Screen View for air dispersion screening model. Risk assessment realizes two components, which are IRAP-h View to assess human health and EcoRisk View to make ecological risk assessment. Emergency release is composed of two tools, which are SLAB View

Suite and SEVEX View. The emission management is in fact GIS based emission information management tool. CAM View is compliance assurance monitoring model to allow managers to see correlations between contaminant emission rates and process parameters.

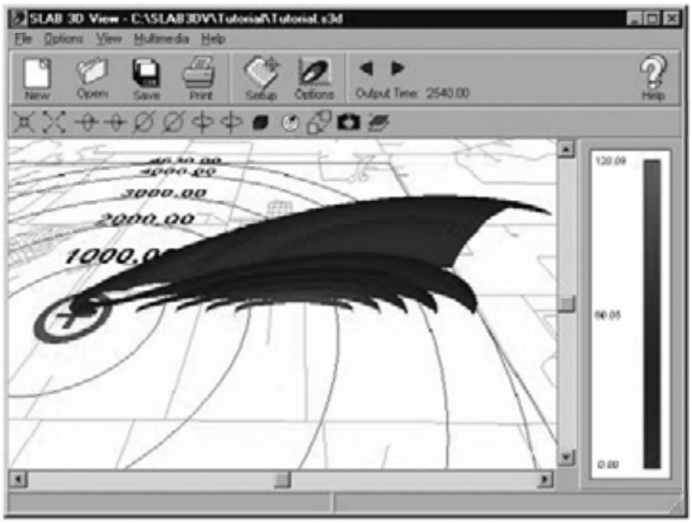


Figure 2.8: Air pollution dispersion visualization in SLAB 3D View (www.scisoftware.com/products/slab_3d_details/slab_3d_details.html, access date May 2009)

2.4 Conclusions of Reviews

(1) Summary of technologies advances

The advance in technologies, especially in computing and communication, actually prepares the basic techniques for CVGE development. The high dimension visualization and multi-channel interface between human and virtual environments that have been developed by VR will help CVGE to achieve multi-dimensional geo-visualization and interaction between participants and virtual environments. The practices of virtual environments, CSCW, groupware and collaboration supply referenced frameworks, architectures and methodologies for CVGE construction. Physical high-speed networks and soft algorithms of data transmission allow online communication which is the base for the implementation of geo-collaboration in CVGE.

Although advanced, the reviewed technologies cannot be termed CVGE simply by being piled together because the platform for CVGE integrates not only the techniques, but also the geo-knowledge. For example, if supporting air pollution simulation, CVGE should integrate circulation models and air pollution dispersion models. Otherwise, the vivid environments created are nothing more than man-made animation games. For instance, some of the VR systems and VE systems are attracted because they make high dimension visualizations that can approach a feeling of immersion. They can show perfect visualizations, such as simulating vivid cloud movement, water flow and so on. If the supporting processes are based on physical geographic rules, the VR and VE can be viewed as a scientific platform to some extent. Otherwise, these geographic phenomena simulations are nothing more than cartoon animation. In this condition, VR and VE are questionable, even though they have beautiful visualizations, because there is no scientific knowledge at their base.

This summing up of the reviewed technologies indicates that modern technologies have advanced to a great extent and certainly supply solid techniques for CVGE development. However, the technological advances still leave challenges as well as the chances for developing CVGE. One of these is the need to combine these technologies to establish an efficient and seamless platform which not only supports geo-collaboration and high dimension geo-visualization, but also integrates geo-knowledge. Because of a lack of integrating geo-knowledge with geo-visualization and geo-collaboration, none of the current virtual environments can be called CVGE.

(2) Summary of GIS supported modelling, geo-visualization and geo-collaboration

Although GIS supported spatial modelling has been widely used in other research areas, GIS itself is still regarded as a tool for preparing the input parameters of models and visualizing the output of models' computations. In a lot of situations, researchers – especially those who do not focus their research on areas in GIScience – use GIS tools (such as ArcGIS) to transform and re-project data or to retrieve geo-information from the data. The after-processed data will then be copied to folders which can be controlled by a program of model computation. After computation, the output will be copied back to the location controlled by the GIS tools for geo-visualization. The moderate integration and tight integration, as Maguire (Maguire et al. 2005) shows, are mainly carried out by geographers because both of these integrations depend greatly on current GIS tools. The challenges of GIS supported modelling, which are highlighted by Goodchild (Goodchild 2005) as dynamic modelling, validation and the sharing of models, are almost impossible to solve using only GIS domains because problem-oriented modelling must be based on professional knowledge, which is beyond GIS. GIS has never handled time particularly well (Goodchild 2005), which puts a hurdle in front of anyone attempting dynamic process modelling with GIS.

Geo-visualization, which was defined by MacEachren, has broken through the traditional meaning of geographic visualization, which focused on spatial representation using geographic graphics, such as maps, tables, diagrams and 3D virtual environments. It has extended to include spatial cognition, knowledge integration and geo-computation. These extensions are also considered to be geo-collaboration as defined by MacEachren.

The combination or even seamless integration of geo-modelling, geo-visualization and geo-collaboration has created a trend for GIScience. In this trend, GIS cannot solve all of the problems anymore, but it can serve as a method and a mediator in both geo-visualization and geo-collaboration.

(3) Summary of the tools to support air pollution simulation

The reviews of the supporting tools for facilitating air pollution modelling also show some challenges. Firstly, researchers know that air pollution is complex and that collaborative work among people from multiple disciplines will enhance the rationality of air pollution modelling. But unfortunately, all of the reviewed tools are designed for use by an individual, and do not support group work. Secondly, spatial and temporal analysis is weak among the supporting tools. As a type of geographic process, air pollution dispersion will be better studied if it is merged into a geographic environment because the professional geo-knowledge can be seamlessly inputted into the air pollution process. Thus, spatial and temporal analysis can be used to discover geographic factor related rules. Thirdly, the supporting tools are almost all closed. Even the zoom in/out functions, which is the most basic operation for geographic information visualization, are blocked. High-dimension visualization is used by these tools, but it is scale-fixed in a low-resolution box, which is a long way from virtual reality. The interaction is low, and the immersion is lacking.

(4) Further conclusions

From the above discussion, further conclusions can be drawn as follows:

- Advanced technologies, especially the VR, VE, CSCW, collaboration, network communication and SEDRIS, have prepared the techniques for CVGE development.
- GIS supported modelling evidences the importance and usability of spatial modelling. However, it also reveals the shortcomings of dynamic modelling that are caused by a GIS inherent framework that aims to deal with geo-data, but not geo-models.
- Geo-visualization and geo-collaboration integrate geo-models with geo-visualization, geo-computation and geo-collaboration into a seamless platform for geo-knowledge sharing and generation, spatial cognition and the interaction of multiple participants to approach cross-disciplinary research into complicated geo-phenomena and geo-processes.

- CVGE, which borrows from advanced modern technologies, has taken the trend towards the integration of geo-data, geo-models, geo-visualization, geo-computation and geo-collaboration to build a shared and distributed virtual environment which engages multiple participants from geographically dispersed locations to carry out collaborative research on geo-phenomena and geo-processes.
- Air pollution simulation is complex and cross-disciplinary, and a feasible technological infrastructure and tools are not yet available to support collaborative group work. At the same time, the current tools also lack geographic information integration, spatial and temporal analysis, scalable visualization and a high interactive interface.
- CVGE-based air pollution simulation is valuable because the merits of CVGE are not implemented by the current tools in the following aspects: ① one feature of CVGE as a method of geo-model integration is that it allow integration of circulation models and air pollution dispersion models into CVGE; ② the integration provides the opportunity to retrieve geographic factors from CVGE, such as the input from circulation models and air pollution dispersion models; ③ multi-dimension geo-visualization can produce air pollution distribution and dispersion at geographic locations in 2D, 3D and even higher dimensions, which can clearly present and facilitate understanding of air pollution; and ④ geo-collaboration provides the possibility of cross-disciplinary simulation of air pollution by engaging multi-disciplinary experts from geography, atmospheric science, environmental science and computer science.

The above conclusions form the primary motivations for this thesis, which attempts to study a prototype of collaborative virtual geographic environments in order to facilitate air pollution simulation. The reasons for carrying out this study are: ① in the community of GIScience, CVGE is both a trend and a challenge; ② CVGE is a new, valuable platform that can facilitate air pollution simulation; and ③ the development of modern technologies allow us to turn the above two possibilities into reality.

CHAPTER 3: DEVELOPING FRAMEWORKS OF COLLABORATIVE VIRTUAL GEOGRAPHIC ENVIRONMENTS

3.1 Introduction

In 2001, a book was published named "Virtual Geographic Environments-A Geographic Perspective on Online Virtual Reality" (Gong and Lin 2001). It is the first time to systematically discuss virtual geographic environments (VGE). Before and after that, researches are carried out for VGE not only on scientific researches but also applications. In China, VGE has attracted the attention of scholars. Lin argues that VGE can be viewed as new generation knowledge of geography(Lin, Gong and Shi 2003, Lin and Zhu 2005a) and the core has extended from geo-data to geo-data and geo-model comparing with GIS(Lin and Xu 2007). Lin also argues that VGE can become one of the scientific methods and advanced technologies of modern experimental geography study (Lin, Huang and Lv 2009). Gao et al carries out researches on VGE based virtual battle terrain and VGE modeling(Gao, Xia and You 1999). Sun leads a group to do research on "Innovation virtual environments for resource environment" and "virtual multiple dimension information generation system for earth system science"(Sun 1999). Gong applies VGE to facilitate researches in public health and digital drainage area (Gong et al. 2008, Zhang 2006). Li and Zhu focus their researches on Cyber-City with the feature of virtual reality(Li, Zhu and Li 2000). Guonian Lv establish the first lab of virtual geographic environments to study cognitive theory in virtual environments, the fast way to collect and assimilate 3D data, geographic analysis modeling and integration into VGE, distributed cluster computers based spatial collaboration and decision making. Chen carries out research on forest modeling, generation and real time rendering for VGE(Shu, Zhu and Chen 2004). Outside of China, there are also many researches related with VGE. MacEachren and Cai research and develop a collaborative geo-visualization environment(MacEachren et al. 2004). Dykes from University of Leicester established a Virtual environment for student fieldwork using networked components (Dykes, Moore and Wood 1999). Electronic visualization laboratory at University of Illinois creates two famous immersion systems: CAVE and GeoWall, which have been used in geography research and education(Johnson et al. 2006a). Batty from University College London does modeling in virtual environments and applies it in urban planning(Hudson-Smith et al. 2007). At the same time, more and more commercial companies turn to research on virtual environments with geographic referencing. The most famous one is Google Earth which was launched in 2005. According to Jones, who spoke at the Fifth International Symposium on Digital Earth, more than 200 million people have downloaded Google Earth (Jones 2007). That's just under the population of Indonesia at roughly 233 million and the United States at just more than 301 million. To compete with Google company, Microsoft puts forth their product—Virtual Earth. There are other companies who show their virtual environments with geographic referencing, such as Leica Virtual Explorer from Leica company, TerraExplorer from Skyline company, TerrainView from ViewTec company, and so on.

Compared with VGE, collaborative virtual geographic environments (CVGE) are little tackled. The CGVE (collaborative geographic visualization environments) corned by MacEachren and Brewer (MacEachren 2001) can be regarded as indirect discussions on some of the issues of CVGE. In that paper, the authors proposed a comprehensive conceptual framework for research on CGVE in the context of geo-collaboration with six components as follows: ① problem context – a distinction among four contexts: decision support, design, knowledge construction and refinement, and training-education; ② collaboration tasks – four pairs related to group decision-making and scientific work, respectively: generate-explore, negotiate-analyses, choose-synthesize, execute-present; ③ commonality of perspective – two continua related to group decision-making and science: conflict to cooperation and different paradigm to same paradigm; ④ spatial and temporal context

– same time same place, same time different place, different time same place, different time different place; ⑤ interaction characteristics – three components of interaction: typology of connections among actors, size and aggregation of the group(s), and constraints on information transfer among actors; and ⑥ environment as mediator – many issues with how to represent both the information forming the basis of collaboration and the participants in collaboration and their behaviors. Gong and Lin (Gong and Lin 2006) examine CVGE by developing a CVGE framework to facilitate "the collaborative planning of silt dam systems in watersheds". They also defined CVGE as "a 3-D, distributed, and graphical world representing and simulating geographic phenomena and processes to enable geographically distributed users to explore geo-problems and theories and generate hypotheses, and to support geo-model building and validation and collaborative ecological planning". On the aspect of system developing, Zhu et al. tried a technique approach of CVGE construction with P2P (Peer-to-Peer) and grid technologies (Zhu et al. 2007a).

CVGE are in its infancy. Previous researches on CVGE are valuable for discovering CVGE. This research does agree partly with some of the opinions developed by the precursors. This chapter will re-define CVGE, develop a conceptual framework and a system framework of CVGE to match new definition, and discuss the theories behind frameworks.

3.2 CVGE Definition

CVGE can be simply regarded as VGE with extended function of geo-collaboration. Before defining CVGE, VGE will be explored first.

Lin defined VGE as follows (Lin and Zhu 2005b). Virtual Geographic Environment (VGE) is a virtual representation of the natural that enables a person to explore and interact, in cyberspace, with the vast amount of natural and cultural information gathered about the physical and cultural environment. In other words, VGE consists of the avatar-based human society and the surrounding objective environments, such as the hardware of computer network and sensors, and the software of data and culture environments. In particular, the avatar-based humans here are a combination of the human in the real word and the 3D avatar in the virtual world. From this point of view, users become part of the data set, part of the VGE, where they can explore and interact with the virtual world. Compared with the data-centered GIS, VGE is, therefore, also defined as a human-centered environment that represents and simulates geographic environments (physical and human environments) and allows distributed multi-user to implement exploratory geospatial analysis, geo-computation, and geo-visualization, and to control collaborative work for supporting design and decision.

After several years' advancement, VGE adds many new perspectives. The most important one is about geo-model integration and management in VGE. Just as GIS manage geo-data, applications oriented VGE should have the ability to manage geo-models(Lin and Xu 2007). It is the geo-models that make simulation, prediction, and virtual dynamic presentation possible. The second perspective on VGE is that VGE will become the new generation language of geography(Lin and Zhu 2005a). The third one regarded VGE as a human centered environment(Gong and Lin 2001, Lin and Gong 2002, Lin and Zhu 2005c), and the forth one argues that VGE can serve as a tool for experimental geography(Lin et al. 2009).

Based on this discussion, this thesis defines VGE as follows: Virtual Geographic Environments, which are virtual representations of physical geography, culture geography and imaginary geography, have two cores (geo-data and geo-models) centred by virtual environments established with the support of modern technologies for information collection (such as remote sensing, global navigation satellite system and photogrammetry), huge volume data storage, wide bands network, high performance computation, high fidelity graphics rendering and multiple channel human-computer interaction for the representation, computation, simulation and analysis of geo-phenomena and geo-processes, not only to explain "what" the geographic phenomena are, but also to discover "why" and "how" the phenomena act. Unlike the previous definition (Lin and Zhu 2005b), this emphasizes the importance of the management of geo-models and the representation, simulation and prediction of geo-models based geographic phenomena. The new definition weakens the features of high dimension visualization and human centeredness because these are not the core elements of VGE. Visualization is the only method for geo-knowledge representation in VGE, and the dimension should be multiple dimensional, and not only high. Under some conditions, the transfer data and information and the running of models only take place within a virtual system, while a human is outside of the system. This system could also be called VGE, but without a human at the centre.

Because it is VGE, with the extended function of geo-collaboration, CVGE on the one hand inherits the characteristics of geo-data and geo-model integration and multi-dimensional visualization from VGE, while on the other hand extending this with new features such as engaging with multiple participants, distributing functions and operations, sharing geo-knowledge and supporting geo-collaboration. CVGE is still centred on geo-data and geo-model, as VGE is. But the differences between CVGE and VGE are also obvious. Firstly, CVGE must be oriented towards multiple users, while VGE can only be operated by an individual. Secondly, CVGE should contain the functions and operations needed to combine geographically dispersed participants, while VGE can be situated on one desktop. Thirdly, CVGE supports geo-collaboration, while VGE does not do this yet. This thesis therefore defines collaborative virtual geographic environments as follows.

Collaborative virtual geographic environments, which are VGE with the extended function of geo-collaboration and which are based on the two cores of geo-data and geo-models, are a geo-knowledge sharing and generation platform that engages multiple participants from geographically dispersed places to study and manage geo-processes and geo-phenomena in the multi-dimension virtual environments. Compared with previous definitions, this contains the "six components" of CGVE coined by MacEachren and Brewer (MacEachren and Brewer 2004), as well as highlighting the functions which are tackled by Gong and Lin (Gong and Lin 2006).

3.3 Conceptual Framework of CVGE

A conceptual framework is developed in this section according to definition of CVGE given in 3.2. The following discussions will tackle the important, though not exhaustive, dimensions of CVGE.

3.3.1 The two cores of geo-data and geo-models

CVGE inherits the merits of a geographic information system (GIS), such as geo-data management, spatial and temporal analysis and geographic visualization. Geo-data still has a core position in CVGE, but the formats and contents have been broadly extended. Besides the widely-used raster and vector, which were developed by GIS, other types of data are also required to create high dimension CVGE, such as 3D objects, textures, video, audio and instant messaging. These extended data types improve the heterogeneity of geo-data and thus present more complex challenges for geo-data integration and management. On the flip side, the broad geo-data provides abundant information to support spatial and temporal analysis.

Geo-data can be viewed as one way of recording geographic phenomena. However, in geographic environments, a huge range of geo-processes cannot be recorded as geo-data, such as wind flow, air pollution dispersion and human behaviours, because they are dynamic and the shapes are changing from time-to-time. Fortunately, geo-models can be created as geo-knowledge to represent these geo-processes. Geo-models almost cover the geographic problems that cannot be recorded by geo-data. If geo-data based geographic presentations can be used to show "what" the recorded phenomena are, geo-models based geographic presentations can reveal the mechanisms of "why" and "how" behind the phenomena.

Currently, there are more and more requirements relate to rules behind phenomena. These requirements are stringent because they do not want to discard geo-data in order to grasp geo-models, but want to possess both. The widely-used GIS is designed, centred on and work perfectly well on geo-data. Unfortunately, GIS isolates geo-models from their domain. CVGE takes these requirements and adds geo-models as a second core.

Geo-data integration and management in CVGE can receive references from GIS and a huge number of database systems. However, the integration and management of geo-models in CVGE are lacking, and their challenges are obvious. Some of these challenges are listed below.

(1) Unifying geo-models

A CVGE does not aim to study how it is possible to create professional geo-models of geo-phenomena and geo-processes, but rather tries to integrate multiple types of geo-models. Geo-models come from many disciplines, such as oceanography, atmospheric science, geology, biology, demography, economics, mathematics and physics. A discipline creates its own professional models according to its own research customs. In order to integrate multi-disciplinary models into CVGE as geo-models, they must first be unified in format. This problem can be explained by taking GIS as an example. GIS unifies geo-data by location (x,y,z) and time (t). Given the location and time, all of the geographic data and information can be associated in GIS. Returning to geo-models in CVGE, what approach can be adopted to unify the huge number of heterogeneous geo-models that are created by multiple disciplines?

(2) Matching geo-models

The relationship among models can be simplified as output-input, which means one model's output serves as the others' input. The output-input relation can be types of one-one, many-one, one-many, and many-many.

One-one: If output of model A can only be accepted by model B, and model B only accepts model A's output, we call it one-one type relation.

One-many: If output of model A can be accepted by many models and these multiple models only accept the output of model A, we call the relation one-many relation.

Many-one: If outputs from multiple models can be accepted by model A and model A only accepts these models' output, we call the relation many-one relation.

Many-many: If multiple models' output can be accepted as inputs of other multiple models and vice versa, we call the relation many-many.

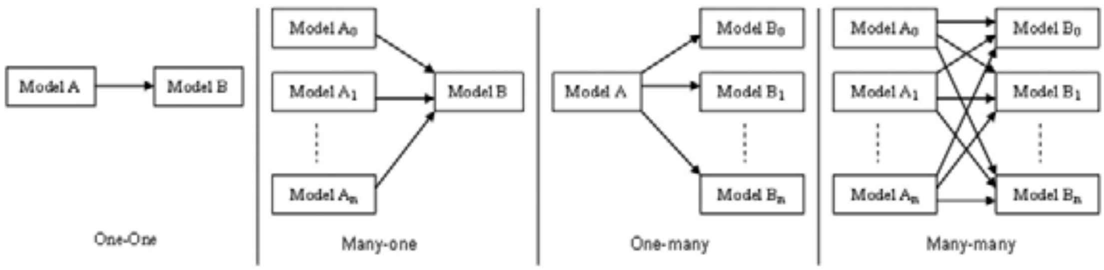


Figure 3.1: Model relations

Matching geo-models means searching a geo-model's input from another geo-models' output. The rules to check if two geo-models are matching are based on parameters' restrictions. Only if an input parameter has the same physical meaning, the same unit, and the same scale as an output parameter do the two parameters match. If part or all of the parameters among two geo-models are matched, the two geo-models are called as partly matching or totally matching.

(3) Efficient computation on geo-models

For synchronous geo-collaboration in CVGE, geo-models should respond to the invoking as soon as possible. Unfortunately, a lot of geo-models are complex and computation intensive. For instance, the Fifth-Generation NCAR / Penn State Mesoscale Mode (MM5), which is widely used for atmospheric circulation modeling and computation, often requires several hours or even several days for a two-week period simulation. This long waiting-time will definitely lower the efficiency of geo-collaboration. Thus, special measures must be taken to minimize computation time of geo-models.

There are two way to approach geo-models computation efficiency. The first one is to optimize geo-models with high efficient arithmetic. It is hard for CVGE developers to solve this problem because a lot of geo-models are established beyond geo-knowledge. The second way is to use high performance computation technologies, such as Computational Grid and parallel computation. The second way is also hard because developers must know how to perform high performance computation. But comparing with the first way, the second way is easy and possible.

(4) Integrating and managing geo-models

The three types of modeling coupled with GIS (loose coupling, moderate integration, and tight integration) (Maguire et al. 2005) can also be applied into CVGE to coupling geo-models. Geo-models are in fact many program codes, and the management of geo-models can be referenced from commercial program developing platforms, such as Visual Stdio.Net 2005, Java, Delphi and so on. In these platforms, models are coded as functions and stored in function libraries (.lib files) or dynamic link libraries (DLL). The platforms also supply description files (such as .h files for VC++ platform), which can be regarded as the metadata of models, to tell users what functions (models) are used for, what the meanings for input and output variables of models are, and so on.

(5) Coupling geo-data and geo-models

Geo-data and geo-models in CVGE are not separated, but coupled. Geo-data can be used to initiate geo-models input parameters. The output of geo-models will enrich geo-data.

3.3.2 Geo-visualization

Two issues are important for geo-visualization in CVGE.

(1) Multi-dimension geo-visualization

Geo-visualization in CVGE must consider usability. Because of the different research backgrounds and different assigned works, participants of CVGE may prefer different dimensions to be available in geographic visualization. This can be explained by examining collaborative air pollution simulation. Suppose a collaborative air pollution simulation is carried out by one geographer, one environmental scientist and one computer scientist. During the geographic visualization, all three are inclined to work according to their custom. The environmental scientist prefers a 2D display, like he has on his familiar GrADS. The geographer will choose a 2D map, but with more interaction between human and the map. For the computer scientist, because of his lack of knowledge of air pollution, he would like the system to present pollution mass in a more natural way, in 3 dimensions or even high dimension.

Geo-visualization aims to present geo-phenomena and geo-processes in an intuitive way, which will improve usability and benefit the overall geo-collaboration. Generally speaking, a higher dimension visualization will have higher usability and consequently lead to more efficient geo-collaboration. However, high dimension geo-visualization will cost more time in system development and require higher performance from the computer source and network communication. Thus, when designing geo-visualization, the balance between dimension and cost should be borne in mind. Two rules can be used to achieve this balance: one is that the geo-visualization must satisfy the overall geo-collaboration; the other one, based on the first rule, is to adopt the lowest dimension.

(2) Geo-data based geo-visualization and geo-models based geo-visualization

CVGE have two cores: geo-data and geo-models. The geo-visualization should tackle both cores, which are thus called geo-data based geo-visualization and geo-models based geo-visualization.

Geo-data is geo-information shot at a certain time point. Once geo-data has been generated, the resulting geo-information is fixed. Using the geo-data, we can only perform geo-information exploration by means of geo-data analysis, mining and so on. Compared with geo-data, geo-models are more affluent in the geo-information provided and can thus be called geo-information generation machines. Geo-models are not only geo-knowledge in themselves, but can also generate new geo-knowledge by changing the input initiate parameters. Therefore, geo-models will become the important engine that can push CVGE forwards. Another reason why geo-models based geo-visualization is important is that geo-models can present dynamic geo-processes. In geographic environments, everything is changing. Things change their positions, shapes, contents, as well as other changeable attributes. The better way of presenting these changes is through geo-models, not geo-data. Thus, the most reasonable way of visualizing the changes is geo-models based visualization rather than geo-data based visualization.

Geo-data based geo-visualization is a popular pattern for a lot of VE work. This pattern is composed of two steps: modelling terrain and objects using model creation tools; and driving them using graphic engines such as OpenGL in virtual environments. Unlike geo-data based geo-visualization, geo-models based geo-visualization automatically models geo-phenomena and geo-processes when the program is running. In this way, modelling and geo-visualization are seamlessly integrated and alternately working.

3.3.3 Geo-collaboration

Geo-collaboration in CVGE support all of the six dimensions defined by MacEachren and Brewer (MacEachren and Brewer 2004). Table 3.1 gives the summary of these six dimensions in CVGE.

Table 3.1: Six dimensions of geo-collaboration in CVGE

Dimensions	Supported by CVGE
Problem context	① Knowledge construction and refinement; ② Design; ③ Decision-support; ④ Training and education
Collaborative task	① Generate; ② Negotiate; ③ Choose; ④ Execute; ⑤ Explore; ⑥ Analyze; ⑦ Synthesize; ⑧ Present
Commonality of perspective	① Shared understanding; ② Resolution of disputes among competing points of view
Spatial and temporal context	① Same time same place; ② Same time different place; ③ Different time same place; ④ Different time different place
Interaction characteristics	① Free group size; ② Connection depends on requirements; ③ Two-way communication information (text, instant messaging, video, audio, files, and so on).
Mediator	① Table; ② Diagram; ③ 2D digital map; ④ 3D environments;

3.4 A System Framework of Collaborative Virtual Geographic Environments

Gong and Lin designed a system framework of CVGE with five levels as network level, data level, modeling level, graphics level, and user level(Gong and Lin 2006). These five levels are absolutely the key domains for CVGE system development. In this research, another system framework of CVGE is developed to highlight the geo-collaboration. The new framework keeps those five levels as the collaborative base. This thesis designs geo-collaboration of CVGE with four levels, which are geo-data collaboration, geo-modeling collaboration, geo-visualization and analysis collaboration, and decision-making collaboration. The four levels of geo-collaborations in CVGE can be explained as Figure 3.2.

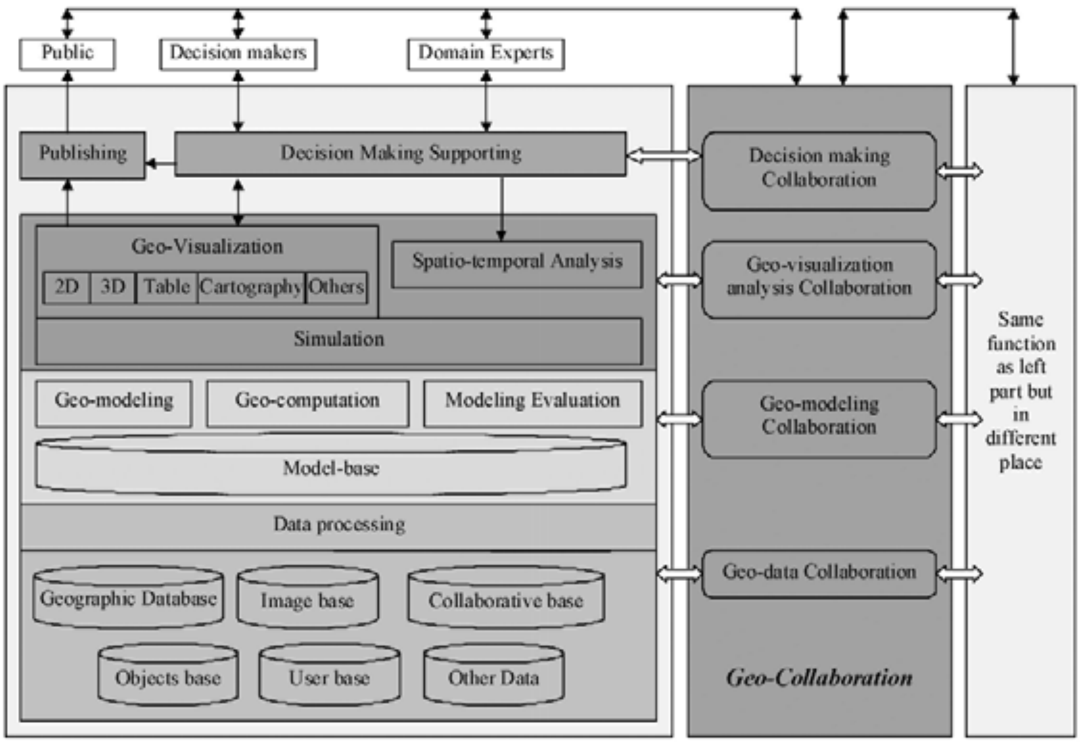


Figure 3.2: System framework of CVGE

3.4.1 Geo-data collaboration

Geo-data is one core of CVGE, which is record of geographic phenomena with spatial and temporal features. Geo-data collaboration is collaboration on geo-data reading, editing, updating, deleting, inserting, and backing-up.

There are several key issues for geo-data collaboration. First one is data consistency. For a single user oriented data management system, data consistency can be ensured by database restriction rules. But for multiple participants' collaborative environments, not only database restriction rules for single computer but also restriction rules among multiple computers are all needed to keep data consistency. The second key issue for data collaboration is about data storage. Three kinds of data storage frameworks for geo-data collaboration, which are centered data storage framework (Figure 3.3 (A)), distributed data storage framework (Figure 3.3 (B)) and hybrid data storage framework (Figure 3.3 (C)). Three kinds of data storage frameworks have their advantages and disadvantages. Framework A is easy for maintaining data consistency and implementing data collaboration, but hard for synchronization when users need to transport data from centered server to user clients for data rendering. Framework B distributes data to users' computers completely, which may get data rendering synchronization easily because there is no data needed to be transported from server to client. But on the other hand, framework B makes data management, date consistency and collaboration more complex. Framework C is a balance between A and B. The advantages and disadvantages from A to C can be summarized in table 3.2.

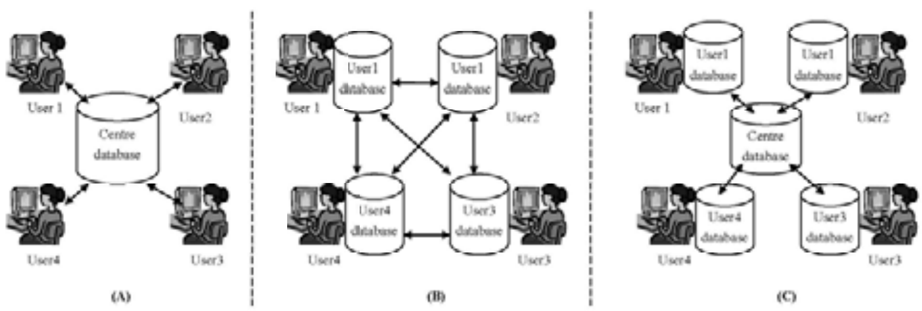


Figure 3.3: Data storage frameworks for geo-data collaboration

The third key issue for data collaboration is real-time and reliability of data transport for network communication. In synchronous collaboration, real time communication calls for fast transportation on data. When huge data, such as terrain data, need to be transported, more bandwidth will be occupied, which causes data delay. At the same time, reliability of data transport should be assured.

Table 3.2: Features for data storage frameworks (A, B and C)

Framework	Advantages	Disadvantages
A	• Easy for data management; • Easy for maintaining data consistency; • Easy for system implementation; • Simple rules for collaboration.	• Hard for data synchronized rendering; • Increasing conflict.
B	• Improving data rendering synchronization; • Less conflict; • Easy for system implementation.	• Complex rules for collaboration; • Each user must maintain his database.
C	• System is adaptive; • Decrease data transportation on network and improve data rendering synchronization; • Conflictions only happen when users make operation on centre database.	• It is very complex for system implementation, collaboration and database maintaining.

There are two ways to implement geo-data collaboration, which are copy/paste methodology (CPM) and lock/unlock methodology (LUM). If geographically distributed participants only use geo-data for rendering and analysis, but not edition and updating, they may copy the geo-data from centre database and paste it to their local database. During collaboration, participants only need to communicate with several bit messages, which can make collaboration smoothly. But if participants need to edit or update geo-data, the lock/unlock methodology will be reasonable. For instance, participant A wants to edit a geo-data set, he needs to check if the geo-data is locked. If yes, A should wait until the geo-data is unlocked. If A has the authority to edit the geo-data, he will lock it first and then does what he wants to do. There are four types of locks can be used for geo-data collaboration, which are intent lock, share lock, update lock and exclusive lock. Intent lock can be used for intention to lock some resource. Share lock allows geo-data users to share but not edit the resource. If a user wants to update a geo-data set, he should use update lock. Exclusive lock is adopted to monopolize data source which can not be edit, update and query by others.

3.4.2 Geo-modeling collaboration

Geo-modeling collaboration includes two aspects, one is geo-model creation collaboration and the other one is geo-model operation collaboration. Geo-model creation collaboration or collaborative geo-modeling is collaboration among multiple participants to create a shared geo-model (such as building, road and so on). Geo-model creation collaboration can get reference from widely used model creation collaboration in product design (Wang and Chou 2008, Mendikoa et al. 2008, Chen, Chen and Chu 2008, Zha and Du 2006). Geo-model operation collaboration includes collaborations on geo-model selection, geo-model parameters setting, geo-model running, geo-model output, processing on geo-model output and so on.

There are two key issues for geo-model collaboration, one is conflict detection and the other one is geo-model management. Conflict of geo-model collaboration happens along with geo-model selection, geo-model parameters setting, geo-model running controlling and geo-model output processing. For example, participant A (PA) wants to use geo-model A (MA) to present a geographic problem. At the same time in the collaborative group, participant B (PB) wants to use another geo-model B (MB) to present the same problem. So, the conflict occurs. The way to solve geo-model collaboration conflicts is saved for future discussion. The second key issue for geo-model collaboration is geo-model management. Geo-model collaboration must be established on the share geo-model pool. When huge number of geo-models fills the pool, an efficient mechanism for geo-model management is pivotal. In GIS, geo-data is organized by location (x,y,z) and time (t). All the geo-data can be queried by spatial temporal index (x,y,z,t). With spatial temporal index, GIS data is standardized as raster or vector which can be stored and management in geo-database management system. Like spatial and temporal data in GIS, geo-models in VGE should be indexed by some features, such as model code.

3.4.3 Geo-visualization collaboration

If classified by visualized contents, visualizations in VGE include scene visualization, geo-data visualization, geo-model operation visualization, and all kinds of computation output visualization. If classified by visualization dimension, visualization in VGE can be classified as 2D visualization, 3D visualization, and high dimension visualization. So, visualization collaboration in CVGE is collaboration on scene visualization, geo-data visualization, geo-model operation visualization and all kinds of visualization in 2D, 3D or even higher dimension environments.

Currently, some visualization collaborations have been tackled, such as scene visualization collaborations for virtual scene and geo-data (Taesombut et al. 2006, Hutanu et al. 2006, Wegman 2000). Scene visualization collaboration means editing and updating on shared scene by one participant can be detected and viewed by other participants. During editing and updating, the other participants can give their suggestions or even prevent the operation. View point navigation is one common way in scene visualization collaboration.

3.4.4 Decision making collaboration

Decision making is always the top and final goal for scientific researches, so does in CVGE. Decision making collaboration is based on geo-data collaboration, geo-model collaboration and visualization collaboration. Decision making is always a work flow step by step. Some of the step may be parts or entire of geo-data collaboration, geo-modeling collaboration and geo-visualization collaboration.

Decision making collaboration may be based on comparison of multiple similar scenarios on same objectives or collaboration on each step of one shared scenario, which is shown as Figure 3.4.

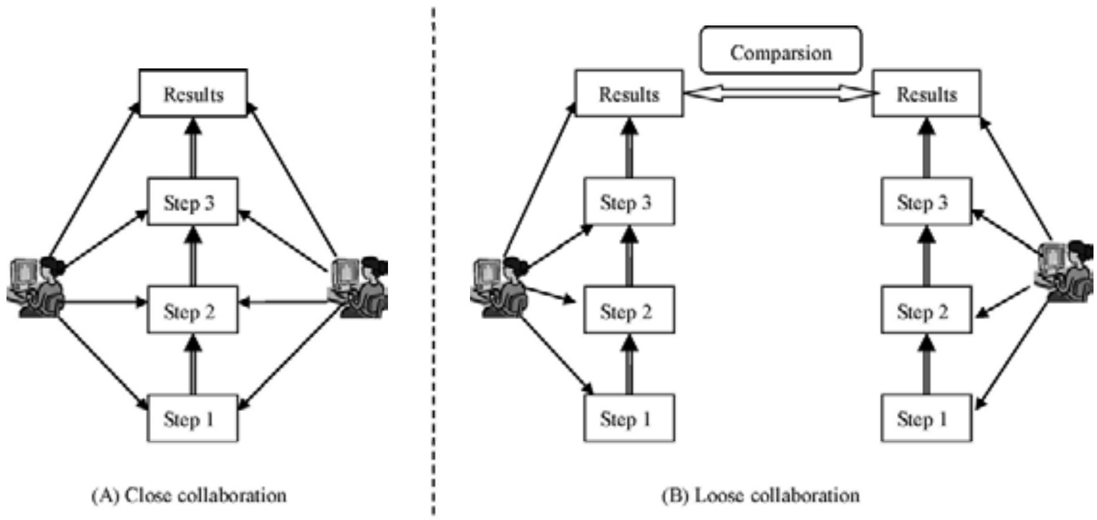


Figure 3.4: Two types of decision making collaborations ((A) is close collaboration which happens in all decision making steps, while (B) is loose collaboration which only compares each participant's result.)

3.5 CVGE Theories behind the Frameworks

Based on the above discussion, CVGE involve a mixture of disciplines from geography/earth system science, social science, information science and computer science, as well as sciences from the fields of application.

Geography, especially given the three sub-disciplines of mapping, cartography and GIS, has possessed one word place in the term of CVGE and does much to contribute to basic theories of CVGE. The scale and projection of maps have been inherited by CVGE. Many basic theories of cartology, including cartographic information theory, cartographic communication theory, cartographic cognition theory, mathematical cartography principles, cartographic linguistics, cartographic visualization principles, cartographic perception theory, cartographic generalization principles, synthetic mapping theory and geo-spatial information pattern theory, are also inherited and are assigned new features in CVGE. GIS is one of the strongest elements that is pushing CVGE forwards. Geo-data management, spatial and temporal analysis and digital cartography, which were developed by GIS, are also bases that support CVGE construction. Earth system science, which has the same position as geography in regard to the basic theories of CVGE, brings mature geo-models and methodologies that deal with complex systems from a comprehensive view.

The geo-collaboration in CVGE involves participants living in real society. The human-computer interaction (HCI) often considers participant characteristics that might influence the success of CVGE. These include physical abilities, user disabilities, cognitive and perceptual abilities, personality differences, cultural and international diversity, age and user experience (MacEachren and Brewer 2004). Human behaviours in CVGE are influenced by the knowledge they have acquired from living in society. The levels of cooperation and conflict among the participants who carry out collaborative work in CVGE depend on a knowledge of social science.

Information science and computer science supply technologies that help in implementing CVGE. A CVGE is a kind of information system. Information science prepares ways of data/information processing, analysis and representation for CVGE. Computer science also lends strong pillars to CVGE development, especially in terms of virtual reality, virtual environments, CSCW, groupware, collaboration and network communication.

The final goal of CVGE is to solve geo-spatial problems. These problems may lie within the domain of geography, earth system science or GIScience. They are also possible beyond the pillar sciences of CVGE. If this is the case, the science where the problems are rooted should also be included as a basis for an applied CVGE.

To sum up, there are four pillar sciences that support CVGE (shown in Figure 3.5). The intersection of the four sciences is specialized as the domain of CVGE. Some of the theories behind CVGE are briefly listed below.

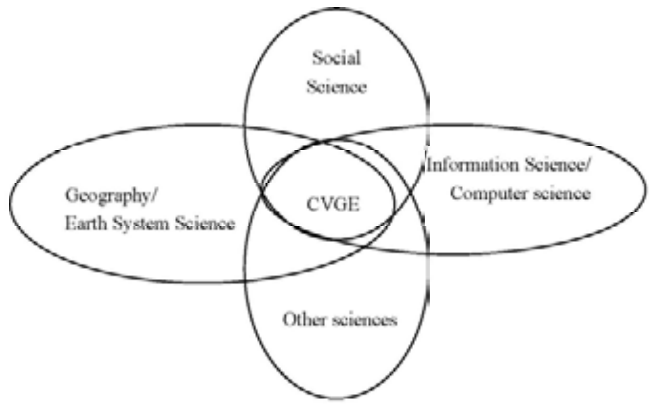


Figure 3.5: Pillar sciences of CVGE

(1) Spatial cognition in CVGE

The study of spatial cognition in geography sheds light on issues such as how spatial knowledge and beliefs are acquired and develop over time; the nature of the structures and processes of spatial knowledge; how people navigate and stay oriented in space; how people use language to communicate with each other about space; and how aspects of spatial knowledge and reasoning are similar or differ between individuals or groups (Montello, Neil and Paul 2001). Spatial cognition in geography includes two aspects: the location and the characteristics of geographic phenomena. It is based on the common cognitive process, i.e. perception, geographic mental imagery, geographic conceptualization, the mental representation of geographic knowledge and geo-spatial reasoning (Wang, Liu and Zhang 2005a). In CVGE, geo-spatial cognition may have typical characters, unlike in the real world and GIS, because there is the possibility of making predictions and high dimension visualizations, or even immersion visualizations.

(2) Spatio-temporal representation in virtual environments

There are two types of representations in CVGE: data-based representations and model-based representation. Model-based representation is most different from both cartology and GIS.

Cartology and GIS are data oriented, and present spatial-temporal phenomena using a symbol system. The representation in both GIS and cartography is static. This is one reason why GIS does not solve the time issue well. However, in CVGE, geo-models couple space and time perfectly. Thus, the geo-model based representation is a real spatio-temporal representation.

(3) Similarity theory in virtual environments

Similarity theory is one of main ways in which CVGE can be applied to represent geo-phenomena and geo-processes in high dimension visualization. Information representation based on similarity theory is easier to understand for information receivers. In CVGE, similarity theory must answer two key questions: one is how to approach the similarities between real geo-phenomena and geo-processes and their virtual counterparts in geometry, temporal characters, dynamics and attributes; the other is the extent to which these can be regarded as similarities.

(4) Spatial and temporal scale theory

Spatial and temporal scales are always key issues in geography. However, spatial and temporal scales have new features in CVGE. Once they are surveyed and stored into the geo-database, the scales of geo-data are fixed in both space and time. Unlike geo-data that has this fixed location and time, geo-models free them as various input parameters (x,y,z,t) and produce different output results. Therefore, with the support of geo-models, CVGE can provide information about a point anywhere and anytime, which blurs the fixed scale view found in geography. CVGE creates a combination of a scale fixed "body" from the geo-data and a scale free "spirit" supported by geo-models.

(5) Integrated modelling in CVGE

Modelling in CVGE includes not only natural geo-phenomena and natural geo-processes such as atmosphere modelling, hydrology modelling and terrain modelling, but also social problems such as urban expansion and pedestrian evacuation. The models in CVGE can also be classified into two types, according their dynamics. If a model never changes its position, shape and content in a virtual environment, it is called a static model. For example, in most virtual environments, a tree is modelled using some kind of model creation tools, and is then imported into the virtual environments. In these virtual environments, the tree will remain static. Compared with the static models, dynamic models involve processes where the shapes, position and contents change with time. Wind flow models, air pollution dispersion models, water flow models and human behaviour models are all examples of dynamic models. Integrated modelling means that, when creating a model, the surrounding models should be considered so that all of the models included in the CVGE couple well.

(6) Theories evolved from information science

A CVGE is an information system. Therefore, CVGE must involve information collection, information encoding, information integration and information processing. Although information science is a mature field and can be used to support CVGE creation, it cannot solve all of the problems related to geographic characteristics of spatio-temporal geo-information in CVGE. CVGE information theory should pay special attention to modern information collection methods such as remote sensing, global navigation satellite systems and photogrammetry for the integration, processing, efficiently use and transformation of geo-information.

(7) Human behaviour theory in CVGE

CVGE inherits a "bridge" character from geography in connecting natural science and social science. Agents and avatars are virtual humans in VGE whose actions should accord with their behaviour in real society. Human behaviour theory in

CVGE attempts not only to study how agents should be shaped, using social science, but also to identify real people behaviour in a virtual environment using avatars.

CHAPTER 4: ARCHITECTURE OF A CVGE PROTOTYPE FOR AIR POLLUTION SIMULATION

This chapter will firstly analyse the requirements that result from multiple participants who are also the potential users. After this, the functions are defined that will fulfill these requirements. The third task will be to design the architecture of a CVGE prototype for air pollution simulation, which will integrate air pollution source data, atmosphere circulation models, air pollution dispersion models, geographic visualization of the output of air pollution modelling and the accompanying geo-collaborations. The key issues in this architecture are identified in the last section of this chapter.

4.1 Requirements Analysis

In the proposed system, Figure 1.2, the various participants who are also the users will have different requirements. The participants can be classified as three roles (domain experts, decision makers and the public). The experts include geographers, atmospheric and environmental scientists and computer technicians. The decision makers are governmental officials, who own the most accurate pollution source data. The following analyses the requirements of the various groups, which are the primary basis for the architecture design.

(1) Requirements from Experts

- **Atmospheric and environmental scientists**

The atmospheric and environmental scientists have requirements as:

- Having high priority to access circulation modeling and air pollutions dispersion modeling;

- Having right of setting some simulation parameters, such as simulation time length, simulation time step, resolution for model computation, configuration of model computation output;
- Having right to query or retrieve some boundary conditions from CVGE;
- Having right of pollution sources evaluation;
- Having right of air pollution result interpretation;
- Having right of air pollution analysis;
- Having right of conducting evaluation on circulation modeling and air pollution dispersion modeling;
- Having ability to receive data and information from other participants;
- Having the right to publish their results and receive others' results;
- Having right to communicate with other participants;

- **Geographers**

The requirements which geographers have are:

- Having the right to integrate and process geo-data and geo-information;
- Having the ability to evaluate any geo-referenced data;
- Having the ability to geographically visualize the output of air pollution modeling;
- Having the ability to conduct spatial and temporal analysis on air pollution distribution and dispersion;
- Having the right to publish their results and receive others' results;
- Having right to communicate with other participants;

- **Computer technicians**

The requirements from computer technicians include:

- Having priority to maintain system security;
- Having right to operate high performance computation;

- Supplying technical support;
- Monitoring operations on system;
- Having right to communicate with other participants;

(2) Requirements from decision makers

The requirements from decision makers are:

- Having the right to propose scenarios;
- Having the priority to compiling air pollution sources, such as adding a new pollution source, deleting a pollution source, update the pollution sources attributes;
- Having the right to join analysis;
- Having the right to collect advice and comments from experts and public;
- Having the priority to make the final decision;
- Having right to communicate with other participants;

(3) Requirements from the public

The public have the requirements as follows:

- Having right to access the published information, such as pollution source distribution, simulated air pollution distribution/dispersion;
- Having the right to vote the decision on air pollution control;
- Having the right to submit their collected information about air pollution sources;
- Having right to communicate with other participants;

4.2 Function Definition

According to requirements analysis, the key functions are defined as follows.

(1) Role identification

This function tries to give a participant a role, such as geographer, atmospheric scientist, environmental scientist, decision makers, and public. When a participant enters into the system, a dialog will be popped up. On the dialog, the roles are listed for choosing. Once the role being chosen, the participant can only do operations as what the role's requirements listed in 4.1.

(2) Pollution source compiler

The pollution sources compiler is used to add, delete, and update air pollution source data. The compiler supplies the tool to generate sources based scenarios.

The compiler can be invoked by decision makers to add/delete/update one pollution source. The compiler also supports batch compiling, such as to open or close all air pollution sources in a city.

(3) Model server and model operation interface

Model server is used to running circulation models, air pollution dispersion models, and some other models. The model server is automatically running behind the visualized interface. It has the ability to detect the model commands from network, chose the right models, launch the calculation, and reply results to commands' generators.

The model operation interface is portable and can be plug in user interface. The model operation interface supplies function of input parameter finalization, model running commands generation, state monitor of model running, and result transformation.

(4) Geo-visualization

The geo-visualization is used to display air pollution sources, output of air pollution modeling (including stationary distribution and animated dispersion), and analysis visualization. The methods adopted to support geo-visualization include table, diagram, 2D visualization, and 3D visualization.

(5) Network communication

CVGE system is distributed. Thus, the function of network communication is required to connect system components and geographically dispersed participants.

The information and data which are transferred on network include constant messaging, pollution sources' adding/deleting/updating information, exchanged files, view point information for geo-visualization, video data, audio data, and so on.

The transferring speed of network communication should be as fast as possible. But at the same time, the online information quality should be ensured.

(6) System coordinator

System coordinator is a self-control program, or can be called an agent. System coordinator is the backbone of the whole system to connect the user, model server, and visualization platform. Besides connection, the functions of system coordinator also include filtering online information, maintaining system security, and relaying information among participants.

Another important function of system coordinator is holding the conflict detector and mediator for negotiation, which will be defined as the next function.

(7) Conflict detector and mediator for negotiation

When conducting geo-collaboration, the conflicts among participants are unavoidable. So, this function is defined to detect the conflict automatically and setup a mediator for negotiation to solve the conflicts.

The conflict detection and negotiation can be installed in system coordinator to form a central style of geo-collaboration. The function can also be distributed in each participant to form a distributed style of geo-collaboration.

4.3 Designing the Architecture of a CVGE Prototype for Air Pollution Simulation

CVGE based air pollution simulation contains several modules, including a geo-data module, a geo-modelling and computation module, a geo-visualization module, a network communication and geo-collaboration module and an analysis and decision-making module. These are put into five tiers as shown in Figure 4.1. Each tier is restricted to suit the CVGE construction and air pollution simulation. The bottom tier includes geo-data, including CVGE development data (DEM, texture, 3D objects, video and audio), air pollution source data and meteorological data (wind filed, temperature, air pressure) and collaboration data (messages, files, rules and shared metadata). The lower geo-data comes from different sources, and has different types, different formats and different scales. The geo-data processing tier is used to validate and transform the geo-data into a CVGE readable format. Above the geo-data processing tier is the geo-modelling and computation tier. The terrain models in this tier are mainly optimizing the arithmetic to improve large-scale terrain data loading, rendering and editing. Unlike the static object models, which can be created before a system is running and stored in a model base, the object models here are dynamic, and are used to create dynamic objects (such as air pollution mass) during the running time. The numerical weather prediction models are used to calculate environmental conditions, and this output will be input into the air pollution dispersion model for air pollution distribution and dispersion computation. The collaboration models will be used to ascertain conflicts and open the way for negotiation. The visualization models are arithmetic, dealing with coordinate conversion, finding the locations of objects in the virtual environment, level of detail, and so on. Analysis models integrate geography-related spatial and temporal analysis arithmetic with domain-specific (air pollution) analysis arithmetic. The visualization tier is located above the modelling and computation tier. This tier is also the interface between computer and operators.

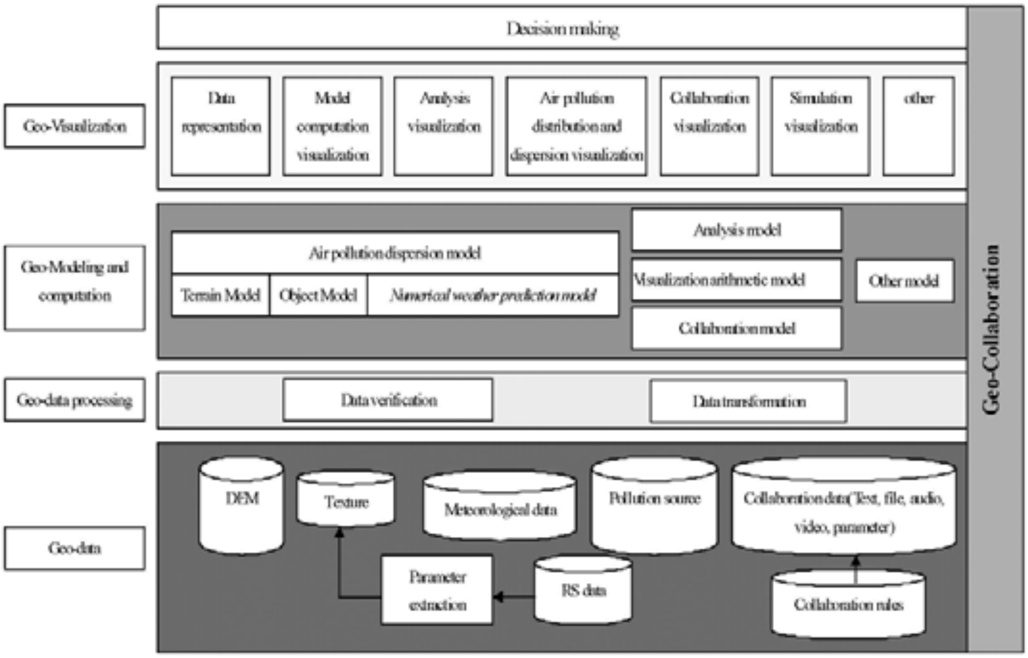


Figure 4.1: Tier architecture of a CVGE prototype for air pollution simulation

The visualization tier includes the functions of geo-data representation, geo-model computation visualization, analysis visualization, air pollution distribution and dispersion visualization, collaboration visualization and simulation visualization. At the top of the architecture is decision making, although collaboration will take place on each tier of geo-data, geo-models and visualization to aid decision making.

Geo-data, geo-models and geo-visualization are not installed on one computer, but are distributed as shown in Figure 4.2. The whole system is structured as a star topology, with the centre of system management node surrounded by a computational grid node to support MM5/SYSUM based modelling and computation, PC cluster node to support Gaussian Plume Model parallel computation, a 2D visualization node, a 3D visualization node and users connected by wide area network. The system management node takes many roles in the overall system, such as: ① middleware to connect users and models in the computation environment; ② security guard to prevent attacks on the computational grid; ③coordinator detecting and negotiating conflicts between the multiple users; and ④ storage of shared data (both pollution source and model computation result). MM5, which is computation intensive, is installed on the computational grid of the Chinese University of Hong Kong (CUGrid). The environmental data to support the running of MM5 is also loaded on the CUGrid. The Gaussian plume model is calculated using parallel computation based on a PC cluster. MM5 and the Gaussian plume model are both controlled by a windows style of model operation interface, which can be integrated by using 2D visualization and 3D visualization. Two visualization nodes integrate not only geo-data, but also geo-model operations, pollution result transformations and visualizations. The visualizations in 2D and 3D environments include geo-model computation visualization, air pollution result (both distribution and dispersion) visualization, analysis visualization and collaboration visualization. The 2D visualization environment and 3D visualization environment are actually the system interfaces, and can be distributed on a wide area network. Therefore, users both on campus and abroad can select one interface to connect to the system, generate running commands for the geo-models, run geo-models on the CUGrid or PC cluster, collect computation results from the geo-models and visualize them and, finally, make geo-collaborations during modelling and visualizations. In this architecture, geo-collaborations happen when multiple users conduct collaborative compilations on air pollution sources, operate on the same models or carry out collaborative visualizations. The module of conflict detection and negotiation on the system management node is a key part of geo-collaboration.

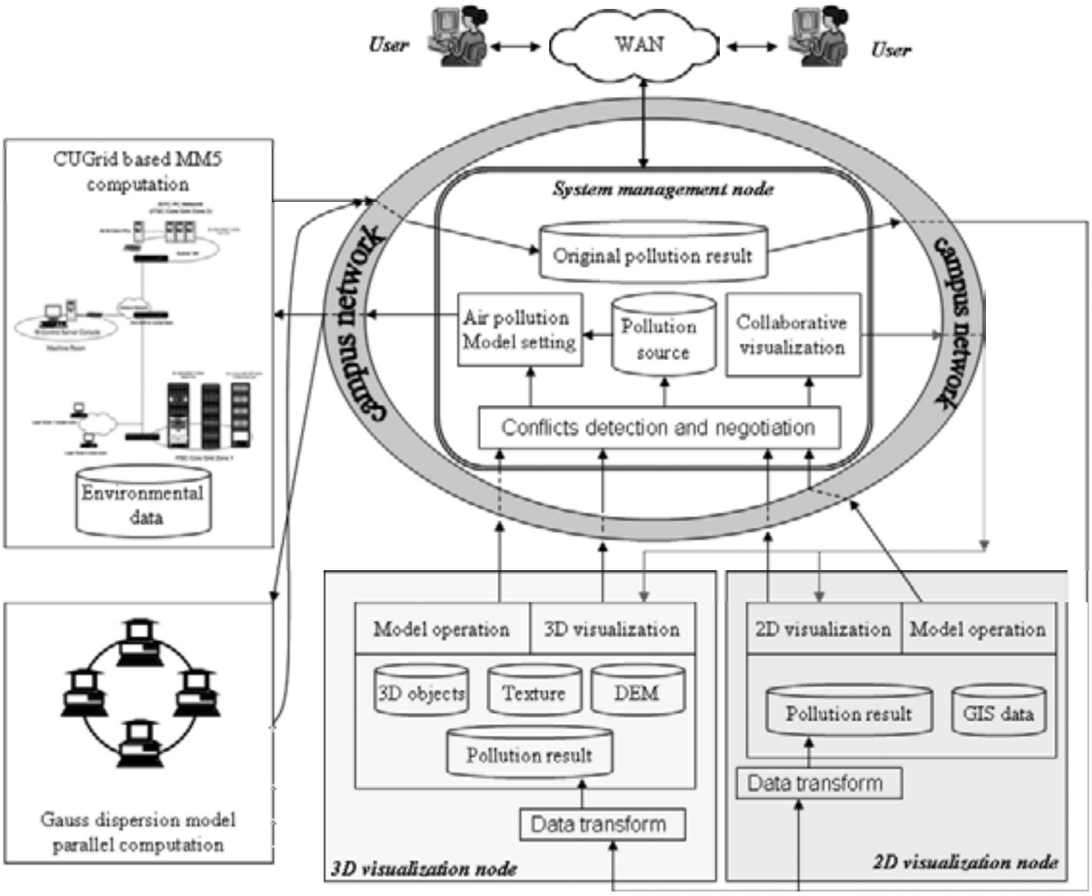


Figure 4.2: Functional architecture of a CVGE prototype for air pollution simulation

In order to meet the distributed functions of Figure 4.2, the hardware architecture is designed as shown in Figure 4.3. The central system is established using facilities from the Chinese University of Hong Kong (CUHK). The computational resources in the CUGrid include hybrid computers such as PC clusters and high performance computers. The parallel computation PCs for the Gaussian Plume Model computation will take place in idle computers in the Institute of Space and Earth Information Science (ISEIS) in CUHK. The visualization nodes are either desktops or notebooks. The bandwidth of the campus network in CUHK is 1 Gigabite (GB). The system management is undertaken by a server computer which has two network cards. One of these is a set local area IP address to connect all campus resources, such as the CUGrid, paralleled computational PCs and visualization interface users. The other card is allocated a wide area IP address so that wide area users can connect. Users both on and out of campus interchange information through a system management node. The hardware is scalable throughout the entire system. This means that computers can be easily added into the CUGrid and paralleled PCs without changing the existing architecture. At the same time, the user interface of the 2D/3D visualization nodes can also be extended to accommodate any number.

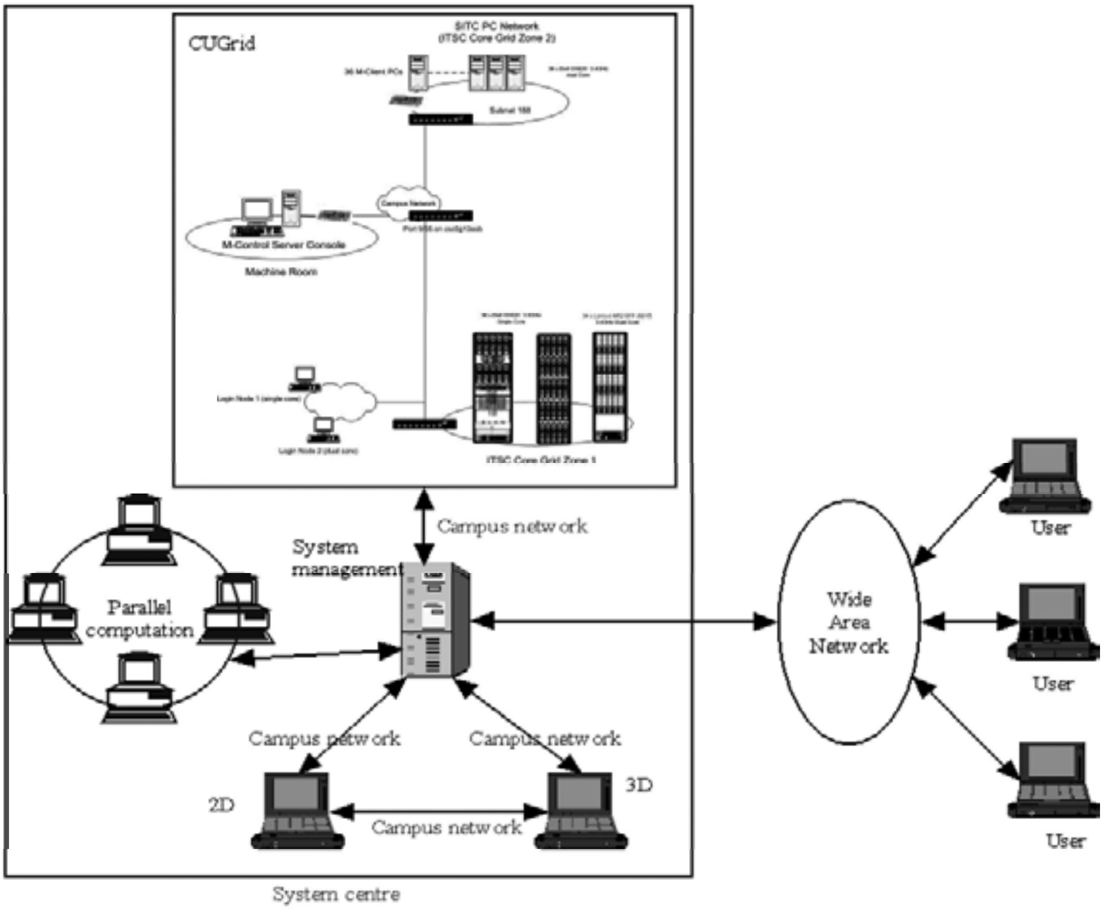


Figure 4.3: Hardware architecture of a CVGE prototype for air pollution simulation

4.4 Key Issues in Architecture

To support the prototype architecture, several key issues must be solved. These include: multiple protocol network communication, multi-dimension geo-visualization of the output of air pollution modeling, high performance computer-supported model computation and geo-collaboration among multiple participants. The following section will discuss the each of these issues. There will be further detailed discussion in the next chapters.

(1) Multiple protocol based network communication

The above architecture is distributed. The network connection will be a key issue to link the geographically distributed users. In order to support geo-collaboration, the network must have the capability to ① transmit a large volume of data as fast as possible and communication information in real time; and ② ensure transmission reliability, especially for command messages.

In order to allow online discussion, the network communication must be able to transmit data and information in real time. When the amount of data and information is small, the real time requirement can be easily implemented. However, the transformed data in CVGE can be very large indeed. For instance, in this research, the air pollution result from model computation could be several GB. It is impossible to send this amount of data to users in a few seconds in order to achieve real time effect. There are two ways of implementing real time communication for large-scale data. One is to transmit only changed data, like that in a Distributed Interactive Simulation (DIS) or High Level Architecture (HLA) (Xu, Lin and Gong 2006). The other way is to create a data pyramid and split the large volume of data into small parts. During online discussion, the core data can be transmitted first, with other information going later. The second mechanism is like that used in Google Earth. Another problem which must be solved for network communication is reliability. In this research, some of the transmitted data can be lost once or even several times, such as texture, but other data must arrive at the destination in time, for instance model running commands and collaboration data.

Multiple transport protocols are a valuable way of meeting these requirements. Transmission Control Protocol (TCP) and Internet Protocol (IP) based connective communication can be applied to send the required data reliably. A User Datagram

Protocol (UDP) may facilitate large data transfers over modern, high-speed, fibre-optic computer networks. At the same time, multicast and Peer-to-Peer (P2P) communications, which are currently used by an increasing number of large-scale, online games and distributed simulations, can be adapted to achieve more flexible transmission.

(2) CUGrid based MM5/SYSUM integration and computation

MM5 and SYSUM, which are used for atmospheric circulation modelling and air pollution dispersion modelling, are both complex and computationally intensive. The computational grid in the Chinese University of Hong Kong (CUGrid) is used to minimize computation time and to try and facilitate online discussion. In order to integrate MM5 and SYSUM into the CUGrid and make them computable, there are several problems that need to be solved.

Firstly, the eight modules of MM5 must be reorganized. Some modules, which will never be used, are discarded. At the same time, other modules are added in order to meet the demands of air pollution simulation, such as a pollution source compiling module and a computational grid based model computation module.

Secondly, MM5 should be distributed in order to make it possible to control remotely. According to the distributed architecture of the whole system, MM5 is separated into an executable part and an operational part. The executable part is installed at the CUGrid to execute commands from users, while the operational part is integrated through a window style user interface for remote control. The executable part should be self-controlling, which means that it can automatically receive commands from users, execute them and send the results back to users.

Thirdly, the user interface should be friendly and usable. The common remote operation terminal for MM5 is SSH (Secure Shell), and the operation is based on command lines. It is unwise to integrate SSH into CVGE directly as the MM5 operation interface because it is very difficult for laymen such as geographers to understand it. A reasonable method is to visualize operations using a windows style interface with simple click functions.

Finally, the connection between the executable part and the user interface should be independent of the operating system. The executable MM5 is running on the CUGrid with a Linux operation system, which is different from the users' windows system. Therefore, the problem of network communication for heterogeneous platforms must be solved.

(3) Parallel computation of Gaussian Plume Model

The computational grid is efficient for high performance computation, but it is also complex and difficult to maintain. Compared with grid computation, parallel computation is easier to control. Therefore, for the Gaussian plume model, PC cluster based parallel computation is used to minimize computation time.

The features of the Gaussian plume model, including the multiple input pollution sources, multiple layers for calculation and multiple grids on each layer, make it suitable for parallelization. The calculation efficiency of the Gaussian plume model is closely related to these features, which are parameterized as a number of point sources, a number of layers and a number of grids on each layer. Thus, the parallel computation issue for the Gaussian plume model is how these three parameters should be combined to achieve fast computation.

(4) Multi-dimension geo-visualization of the output of air pollution modeling

The diagrams, tables, 2D and 3D should all be supported. Of these four methods of visualization, 3D geographical visualization of the air pollution distribution/dispersion is the most difficult.

An air pollution mass does not have clear boundaries like roads and buildings. It can be viewed from outside or inside. For this research, an air pollution mass is called a fuzzy object. Air pollution is also dynamic: it moves here and there and keeps changing shapes as it is driven by the wind. Therefore, an air pollution mass cannot be pre-created using model creation tools, such as 3DMas, and then imported into CVGE. This kind of mass should be established on running time, which is called dynamic modelling. Meanwhile, an air pollution mass is not a surface enwrapped 2.5 dimensional object, but a volume object. Because of these features, an air pollution mass is called a dynamic fuzzy boundary volume object.

The above explanation covers the pollutant's real features. However, the visualization is not based on real conditions, but on an air pollution model output. The output is discrete in slitting the dynamic volume into layers, grids and times. This kind of slitting can easily be rendered in a 2D environment as raster, with different colors and a different transparency for each grid to achieve good view effects. However, in a 3D environment, vivid visualization becomes very hard. In a 3D environment, the discrete pollution space should revert into the original volume style. Some special measures should thus be taken in order to render each point in the volume cube, such as a particle system. To achieve smooth visualization, optimizing arithmetic must be adopted in order to decrease rendering time.

(5) Geo-collaboration among multiple participants

The geo-collaboration on air pollution simulation engages many participants, including experts, officers and the public. This geo-collaboration includes four separate elements: geo-data collaboration, geo-modeling collaboration, geo-visualization collaboration and decision making collaboration. The overall geo-collaboration for the air pollution simulation should satisfy the following requirements.

Firstly, this geo-collaboration should support contributions from different locations. Secondly, this geo-collaboration should support both same time and different time collaboration. Thirdly, participants who connect to collaborative environments should have the ability to detect when others are making changes (such as position changing, shape changing) and update their CVGE contents automatically. Fourthly, regarding conflict detection and negotiation during the geo-collaboration, conflicts should be detected automatically; if a conflict is detected, negotiation should be supplied to resolve this.

CHAPTER 5: INTEGRATION OF AIR POLLUTION MODELING BASED ON HIGH PERFORMANCE COMPUTATION

The system is a combination of different possible air pollution models, but in this research, two specific scales of air pollution modelling are integrated. The first is a regional scale, which is composed of MM5 for atmospheric circulation modelling and coupled with SYSUM for air pollution dispersion modelling. The second is for local scale air pollution modelling, using the Gaussian plume model. The two models can be selected by users from interface. In order to minimize the computation time of regional scale modelling, the CUGrid is adopted. However, for the Gaussian plume model, parallel computation is used to decrease computation time.

This chapter will first introduce MM5, SYSUM, CUGrid and the Gaussian plume model. Then, in the second part, the methodology of integrating MM5 with CVGE will be studied. The implementation of this integration will also be discussed in the second part. In part three, parallel arithmetic for the Gaussian plume model is designed.

5.1 Introduction

5.1.1 MM5 (Dudhia et al. 2005, GA et al. 1994)

MM5 is evolved from a mesoscale model used by Anthes at Penn State in the early '70's. Since that time it has undergone many changes designed to broaden its applications. These include (i) a multiple-nest capability, (ii) nonhydrostatic dynamics, (iii) a four-dimensional data assimilation (Newtonian nudging) capability, (iv) increased number of physics options, and (v) portability to a wider range of computer platforms, including OpenMP and MPI systems. There are eight modules for MM5, which are TERRAIN, REGRID, LITTLE_R/RAWINS, INTERPF, INTERPB, NESTDOWN, MM5, and GRAPH/RIP.

Terrestrial and isobaric meteorological data are horizontally interpolated (programs TERRAIN and REGRID) from a latitude-longitude grid to a mesoscale, rectangular domain on either a Mercator, Lambert Conformal, or Polar Stereographic projection. Since the interpolation of the meteorological data does not necessarily provide much mesoscale detail, the interpolated data may be enhanced (program LITTLE_R/RAWINS) with observations from the standard network of surface and rawinsonde stations using a successive-scan Cressman or multi-quadric technique. Program INTERPF then performs the vertical interpolation from pressure levels to the σ-coordinate of the MM5 model. After a MM5 model integration, program INTERPB can be used to interpolate data from σ-levels back to pressure levels, while program NESTDOWN can be used to interpolate model level data to a finer grid to prepare for a new model integration. Graphic programs (GRAPH/RIP) may be used to view modeling system output data on both pressure and σ-levels. Above description can be shown as Figure 5.1.

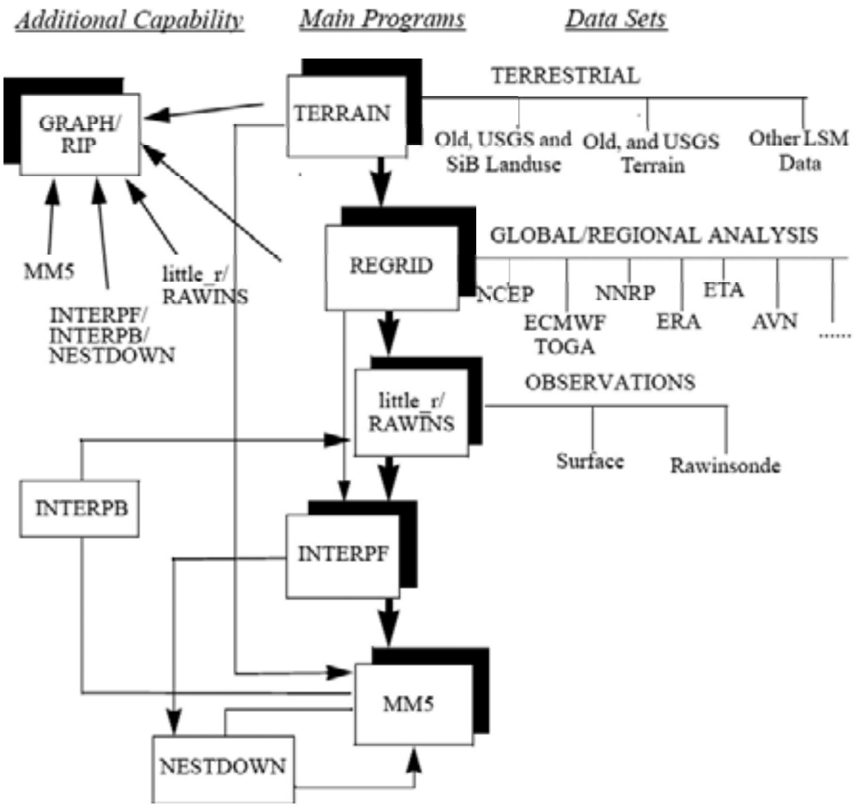


Figure 5.1: The MM5 modeling system flow chart (GA et al. 1994)

The modeling system usually gets and analyzes its data on pressure surfaces, but these have to be interpolated to the model's vertical coordinate before being input to the model. The vertical coordinate is terrain following (see Figure 5.2 (B)) meaning that the lower grid levels follow the terrain while the upper surface is flat. Intermediate levels progressively flatten as the pressure decreases toward the chosen top pressure. A dimensionless quantity σ is used to define the model vertical coordinate.

$$\sigma = (P_0 - P_t)/(P_{s0} - P_t) \qquad \text{..............................(4.1)}$$

Where P_0 is the reference-state pressure, P_t is a specified constant top pressure, and P_{s0} is the reference-state surface pressure.

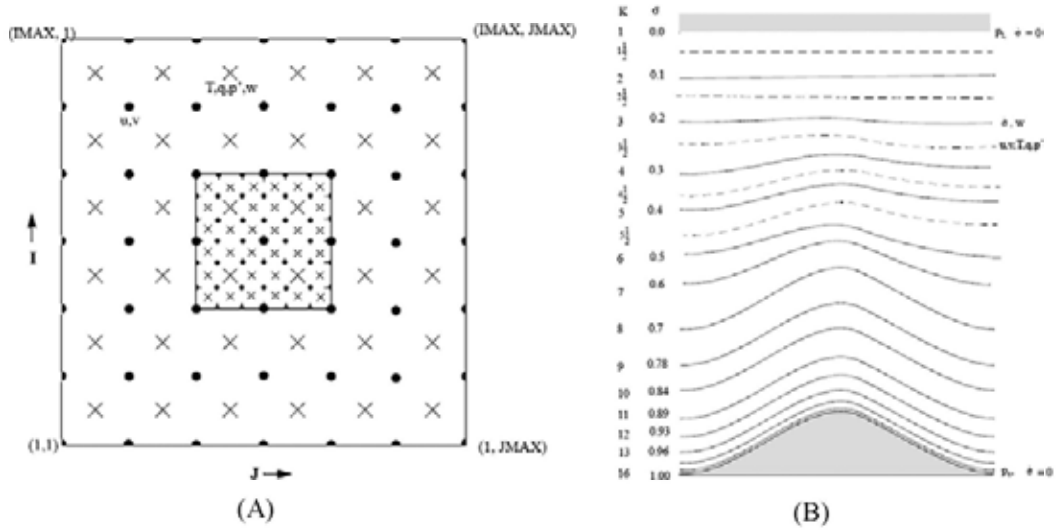


Figure 5.2: The MM5 Model Horizontal and Vertical Grid ((A): Schematic representation showing the horizontal Arakawa B-grid staggering of the dot (l) and cross (x) grid points. The smaller inner box is a representative mesh staggering for a 3:1 coarse-grid distance to fine-grid distance ratio. (B): Schematic representation of the vertical structure of the model. The example is for 15 vertical layers. Dashed lines denote half-sigma levels, solid lines denote full-sigma levels.)(GA et al. 1994)

The horizontal grid has an Arakawa-Lamb B-staggering of the velocity variables with respect to the scalars. This is shown in Figure 5.2 (A) where it can be seen that the scalars (T, q etc.) are defined at the center of the grid square, while the eastward (u) and northward (v) velocity components are collocated at the corners. The center points of the grid squares will be referred to as cross points, and the corner points are dot points. Hence horizontal velocity is defined at dot points, and when data is input to the model the preprocessors do the necessary interpolations to assure consistency with the grid.

5.1.2 SYSUM

SYSUM (Sun Yat-sen University Model) is designed by Professor Wenshi Lin in Sun Yat-sen University through modifying MM5 by adding an equation (see equation 5.2) of air pollution transportation and dispersion. SYSUM is a comprehensive, three-dimensional, multi-scale, non-hydrostatic, Eulerian atmospheric chemical transport model, which is designed to simulate single air pollutant dispersion in the troposphere. SYSUM is not self-governed but integrated into MM5. SYSUM get the input parameters from MM5 results. MM5 provides the time-dependent three-dimensional wind, temperature, pressure, and specific humidity fields. The air pollution transportation and dispersion is driven "on-line" by the meteorological output of the MM5. The time-varying trace gas concentrations are predicted by solving numerically a species continuity equation that includes advection, diffusion, and deposition processes. The equation adopted by SYSUM for air pollution dispersion modeling is:

$$\frac{\partial c}{\partial t}+\frac{\partial uc}{\partial x}+\frac{\partial vc}{\partial y}+\frac{\partial wc}{\partial z}=\frac{\partial}{\partial x}(K\frac{\partial c}{\partial x})+\frac{\partial}{\partial y}(K\frac{\partial c}{\partial y})+\frac{\partial}{\partial z}(K\frac{\partial c}{\partial z})+Q-L \quad \text{............(5.2)}$$

Where:

c — Concentrations of an air pollutant (kg/kg);

u、v、w — Wind velocity component in three dimension(m/s);

K(x,y,z,t) — Eddy diffusivity (m^2/s);

Q — Emission (1/s);

L — Deposition fluxes (1/s).

※ From here to the end of thesis, if there is no distinguishing, MM5 includes original MM5 for atmospheric circulation modeling and coupled SYSUM for air pollution dispersion modeling.

5.1.3 CUGrid

Grid computing is a form of distributed computing whereby a "super and virtual computer" is composed of a cluster of networked, loosely-coupled computers, acting in concert to perform very large tasks. This technology has been applied to computationally-intensive scientific, mathematical, and academic problems through volunteer computing, and it is used in commercial enterprises for such diverse applications as drug discovery, economic forecasting, seismic analysis, and back-office data processing in support of e-commerce and web services. What distinguishes grid computing from typical cluster computing systems is that grids tend to be more loosely coupled, heterogeneous, and geographically dispersed. Also, while a computing grid may be dedicated to a specialized application, it is often constructed with the aid of general purpose grid software libraries and middleware.

The facilities of the Grid of the Chinese University of Hong Kong (CUGrid) are provided for high-performance computing needs and for supporting computation-intensive research projects. It currently consists of 36 single CPU nodes in one logical batch, 70 dual core nodes in a second logical batch and another 48 dual core nodes in another batch. 56 of the quad-core nodes are added into the CUGrid recently. All these nodes are physically interconnected and managed with a grid middleware Globus ToolKit Version 4. The total CPUs available to user via OpenPBS queue is about 490 with total computational power of 2.5TFLOPS (Tera Floating point Operations Per Second) and its processing power is expected to be increased to 3TFLOPs to 4TFLOPs in the shortly future. Figure 4.3 of the CUGrid architecture helps to depict the overall architecture of the CUGrid.

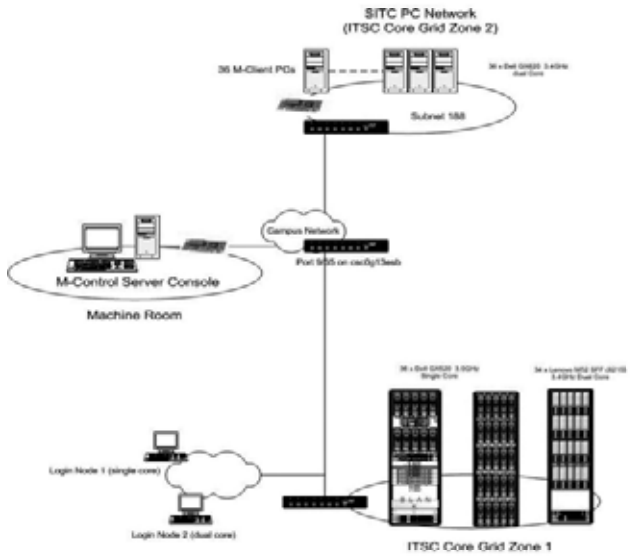


Figure 5.3: CUGrid architecture

CUGrid is constructed to adapt and comply to the standards from the Open Grid Services Architecture (OGSA), its related OGSA Grid Security Infrastructure (GSI) components, W3C's Web Services Resource Framework (WSRF) services and the Globus ToolKit(GTK Version 4). These standards guide the building, verifying, and providing the fundamental computational facilities for all the computing needs in CUHK. Topology of the CUGrid is scalable. It is based on the gigabit network technology that uses mainly commonly available Cisco switches. Because it is built with standard components, the underlying network can easily be expanded using flexible, inexpensive switches to accommodate additional servers for increased horizontal and overall performance on bandwidth and computational capacity.

CUGrid is connected to the Grid Testbeds of the Pacific Rim Applications and Grid Middleware Assembly (PRAGMA) via G/farm at the CUGrid PRAGMA Testbed extension, this effectively allowing CUHK research projects to be collaborated with other institutions in PRAGMA (Pacific Rim applications and Grid Middleware Assembly, e,g, UCSD and NCSA in US, AIST in Japan, KISTI in Koera, CNIC in China etc,). The resources of computational and data resources could then be further expanded and utilized by the researcher for further needs which beyond the CUGrid capacity.

Figure 5.4: PRAGMA Grid Testbed

For this research, the parallel version of the MM5 was configured and domain parameters for NS (Parallel MM5 parameter, for PROCMIN_NS, the minimum number of processors allowed in North/South Dimension) and EW (Parallel MM5 parameter, for PROCMIN_EW, the minimum number of processors allowed in East/West Dimension) domain in MM5 were specifically tuned for optimizing the parallel processes of the project's MM5 data on the servers of the CUGrid. Majority of the server is then available and can be used to support the model computation and distribute the work loads to the desired computational nodes. This helps to achieve strong linearity of the parallel processes for the pre-process of the project's data.

5.1.4 Gaussian plume model(Jiang 2004)

Based on statistic theory, in steady turbulence environment, particle dispersion and transport behaviors accord with Gaussian plume model, which is also called Gaussian dispersion model. Gaussian plume model does not include chemical reaction among pollutants and wet and dry deposition, and is an idealized simplified dispersion model. Gaussian plume model is simple and easy for calculation, which is the reason that it is still widely used. From literature review, many applications based on Gaussian have been carried out, such as point emission dispersion (Arystanbekova

2004, Wu et al. 2004), line emission dispersion(Wang et al. 2004), area emission dispersion, and so on(Collett and Oduyemi 1997). At the same time, Gaussian dispersion model is also used to find the emission location from pollution distribution, such as in work done by ISLAM (ISLAM 1999).

In this research, the classical Gaussian plume model is adopted. The classical Gaussian plume is a steady-state model that requires a continuous emission. This model can be shown as Figure 4.5. The parameters are listed in table 4.1.

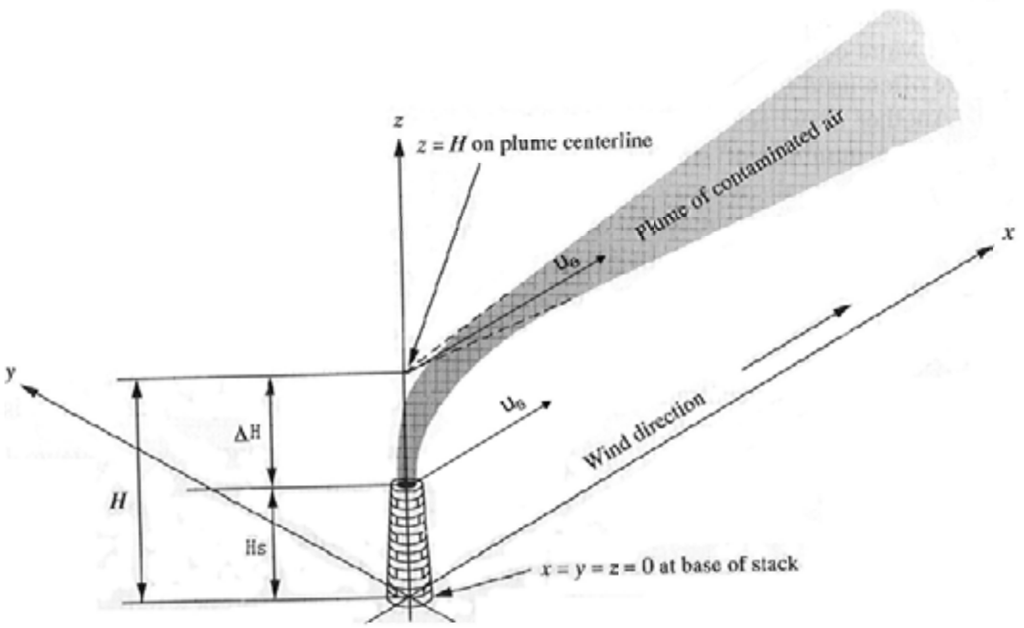


Figure 5.5: Gaussian plume model parameters

Table 5.1: Gaussian plume model parameters

Name	Physical meaning	Unit
H	Effective height of emission	m
H_s	Height of stack.	m
ΔH	Height of smoke rising.	m
C(x,y,z,H)	Pollution concentration at point (x,y,z).	mg/m^3
C(x,y,0,H)	Pollution concentration on terrain surface.	mg/m^3
C(x,0,0,H)	Pollution concentration on terrain axis surface.	mg/m^3
u_s	Wind speed at stack height (H_s+ΔH).	m/s
u_e	Wind speed at the effective height of emission (H_s+ΔH).	m/s
x	Downwind distance.	m

y	Horizontal distances from plume centerline at *x* meters.	m
z	Vertical distances from plume centerline at *x* meters.	m
F_t	Pollutant residence half-time.	h
$\delta_x=\delta_y$	Horizontal diffusion coefficient.	m
δ_z	Vertical diffusion coefficient.	m
T	Integral time (Low wind duration).	s
Q	Emission rate.	mg/s
L_d	The height of mixing layer.	m

The validity of Gaussian plume model depends on the accuracy of above parameters, especially $\delta_x(=\delta_y)$, δz and ΔH. For above listed parameters in table 5.1, u can be gotten from weather observation data; Q can be gotten by calculation or measure; $\delta_x(=\delta_y)$, δ_z and ΔH are closely correlated with weather and terrain surface which can be calculated with reference (Jiang 2004).

Gaussian plume model can be classified according to surface wind speed. It is also related with mixing layer height.

(1) Wind speed (u) is higher than 1.5m/s

- **The height of mixing layer (L_d) is higher than 2000m**

The reflection by top of mixing layer is not considered.

$$\begin{cases} C(x,y,z,H)=\dfrac{0.5Q}{\pi u_e\delta_y\delta_z}\exp(-\dfrac{0.5y^2}{\delta_y^2})\bullet \\ \qquad [\exp(-\dfrac{0.5(z-H)^2}{\delta_z^2})+\exp(-\dfrac{0.5(z+H)^2}{\delta_z^2})]\exp(-\dfrac{0.693x}{3600F_tu_e}) \\ C(x,y,0,H)=\dfrac{Q}{\pi u_e\delta_y\delta_z}\exp(-\dfrac{0.5y^2}{\delta_y^2})\exp(-\dfrac{0.5H)^2}{\delta_z^2})\exp(-\dfrac{0.693x}{3600F_tu_e}) \\ C(x,0,0,H)=\dfrac{Q}{\pi u_e\delta_y\delta_z}\exp(-\dfrac{0.5H)^2}{\delta_z^2})\exp(-\dfrac{0.693x}{3600F_tu_e}) \end{cases} \quad \text{........(5.3)}$$

- **The height of mixing layer (L_d) is not higher than 2000m**

Influence of mixing layer top is considered. Gaussian plume model is classified into three conditions.

First condition: $L_d \geq H_s+1.5\Delta H$. Pollutant is totally reflected by mixing layer.

$$\begin{cases} C(x,y,z,H)=\dfrac{0.5Q}{\pi u_e \delta_y \delta_z}\exp(-\dfrac{0.5y^2}{\delta_y^2})\exp(-\dfrac{0.693x}{3600F_t u_e})\bullet \\ \qquad \sum_{n=-4}^{4}[\exp(-\dfrac{0.5(z-H+2nL_d)^2}{\delta_z^2})+\exp(-\dfrac{0.5(z+H-2nL_d)^2}{\delta_z^2})] \\ C(x,y,0,H)=\dfrac{Q}{\pi u_e \delta_y \delta_z}\exp(-\dfrac{0.5y^2}{\delta_y^2})\exp(-\dfrac{0.693x}{3600F_t u_e})\sum_{n=-4}^{4}\exp(-\dfrac{0.5(H-2nL_d)^2}{\delta_z^2}) \\ C(x,0,0,H)=\dfrac{Q}{\pi u_e \delta_y \delta_z}\exp(-\dfrac{0.693x}{3600F_t u_e})\sum_{n=-4}^{4}\exp(-\dfrac{0.5(H-2nL_d)^2}{\delta_z^2}) \end{cases} \quad \ldots\ldots(5.4)$$

Second condition: $H_s+0.5\Delta H \leqslant L_d \leqslant H_s+1.5\Delta H$. Pollutant is partly reflected by mixing layer.

$$\begin{cases} C(x,y,z,H)=\dfrac{0.5Q(1-P)}{\pi u_e \delta_y \delta_z}\exp(-\dfrac{0.5y^2}{\delta_y^2})\exp(-\dfrac{0.693x}{3600F_t u_e})\bullet \\ \qquad [\exp(-\dfrac{0.5(z-H)^2}{\delta_z^2})+\exp(-\dfrac{0.5(z+H)^2}{\delta_z^2})] \\ C(x,y,0,H)=\dfrac{Q(1-P)}{\pi u_e \delta_y \delta_z}\exp(-\dfrac{0.5y^2}{\delta_y^2})\exp(-\dfrac{0.693x}{3600F_t u_e})\exp(-\dfrac{0.5H^2}{\delta_z^2}) \\ C(x,0,0,H)=\dfrac{Q(1-P)}{\pi u_e \delta_y \delta_z}\exp(-\dfrac{0.693x}{3600F_t u_e})\exp(-\dfrac{0.5H^2}{\delta_z^2}) \\ P=1.5-(L_d-H_S)/\Delta H \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(5.5)$$

Third condition: $L_d \leqslant H_s+0.5\Delta H$. All pollutants penetrate mixing layer. So, the pollution concentration under mixing layer is zero.

$$C(x, y, z, H)=0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(5.6)$$

(2) Wind speed is low (u≤1.5m/s)

- **The height of mixing layer (L_d) is higher than 2000m**

The reflection by top of mixing layer is not considered.

$$\begin{cases} C(x,y,z,H) = \int_0^T \frac{Q}{(2\pi)^{1.5}{\delta_y}^2\delta_z} \exp(-\frac{0.5(x^2+y^2)}{\delta_y^2}) \bullet \\ \qquad [\exp(-\frac{0.5(z-H)^2}{\delta_z^2}) + \exp(-\frac{0.5(z+H)^2}{\delta_z^2})] \exp(-\frac{0.693x}{3600F_t u_e}) dt \\ C(x,y,0,H) = \int_0^T \frac{2Q}{(2\pi)^{1.5}{\delta_y}^2\delta_z} \exp(-\frac{0.5(x^2+y^2)}{\delta_y^2}) \exp(-\frac{0.5H^2}{\delta_z^2}) \exp(-\frac{0.693x}{3600F_t u_e}) dt \\ C(x,0,0,H) = \int_0^T \frac{2Q}{(2\pi)^{1.5}{\delta_y}^2\delta_z} \exp(-\frac{0.5x^2}{\delta_y^2}) \exp(-\frac{0.5H^2}{\delta_z^2}) \exp(-\frac{0.693x}{3600F_t u_e}) dt \end{cases} \qquad \text{.......(5.7)}$$

- **The height of mixing layer (L_d) is not higher than 2000m**

Influence of mixing layer top is considered. Gaussian plume model is classified into three conditions.

First condition: $L_d \geq H_s + 1.5\Delta H$. Pollutant is totally reflected by mixing layer.

$$\begin{cases} C(x,y,z,H) = \int_0^T \frac{Q}{(2\pi)^{1.5}{\delta_y}^2\delta_z} \exp(-\frac{0.5(x^2+y^2)}{\delta_y^2}) \bullet \\ \qquad \sum_{n=-4}^{4} [\exp(-\frac{0.5(z-H+2nL_d)^2}{\delta_z^2}) + \exp(-\frac{0.5(z+H-2nL_d)^2}{\delta_z^2})] \\ \qquad \bullet \exp(-\frac{0.693x}{3600F_t u_e}) dt \\ C(x,y,0,H) = \int_0^T \frac{2Q}{(2\pi)^{1.5}{\delta_y}^2\delta_z} \exp(-\frac{0.5(x^2+y^2)}{\delta_y^2}) \\ \qquad \sum_{n=-4}^{4} \exp(-\frac{0.5(H-2nL_d)^2}{\delta_z^2}) \bullet \exp(-\frac{0.693x}{3600F_t u_e}) dt \\ C(x,0,0,H) = \int_0^T \frac{2Q}{(2\pi)^{1.5}{\delta_y}^2\delta_z} \exp(-\frac{0.5x^2}{\delta_y^2}) \\ \qquad \sum_{n=-4}^{4} \exp(-\frac{0.5(H-2nL_d)^2}{\delta_z^2}) \bullet \exp(-\frac{0.693x}{3600F_t u_e}) dt \end{cases} \qquad \text{...........(5.8)}$$

Second condition: Hs+0.5ΔH⩽Ld⩽Hs+1.5ΔH. Pollutant is partly reflected by mixing layer.

$$\begin{cases} C(x,y,z,H) = \int_0^T \frac{0.5Q(1-P)}{\pi u_e \delta_y \delta_z} \exp(-\frac{0.5(x^2+y^2)}{\delta_y^2}) \exp(-\frac{0.693x}{3600F_t u_e}) \bullet \\ \qquad [\exp(-\frac{0.5(z-H)^2}{\delta_z^2}) + \exp(-\frac{0.5(z+H)^2}{\delta_z^2})]dt \\ C(x,y,0,H) = \int_0^T \frac{Q(1-P)}{\pi u_e \delta_y \delta_z} \exp(-\frac{0.5(x^2+y^2)}{\delta_y^2}) \exp(-\frac{0.693x}{3600F_t u_e}) \exp(-\frac{0.5H^2}{\delta_z^2})dt \\ C(x,0,0,H) = \int_0^T \frac{Q(1-P)}{\pi u_e \delta_y \delta_z} \exp(-\frac{0.5x^2}{\delta_y^2}) \exp(-\frac{0.693x}{3600F_t u_e}) \exp(-\frac{0.5H^2}{\delta_z^2})dt \\ P = 1.5 - (L_d - H_s)/\Delta H \end{cases} \quad \ldots(5.9)$$

Third condition： Ld≤Hs+0.5ΔH. All pollutants penetrate mixing layer. So, the pollution concentration under mixing layer is zero.

C(x, y, z, H)= 0 …………………..(5.10)

5.2 MM5 Integration into CVGE Based on the CUGrid

5.2.1 MM5 integration framework based on the CUGrid

The framework for MM5 integration into CVGE based on the CUGrid is shown in Figure 5.6. The framework includes three parts: the CUGrid based model server, system management and users. The framework is in three tiers and is based on the "Server-Client/Server-Client" mode. The users are clients who generate commands and send these to the server through the management node. The management node serves as middleware to relay in two directions the commands/messages between the distributed users and the CUGrid server. Therefore, the management node can be viewed as a server for the users and a client for the CUGrid server. CUGrid security is very high: no users are permitted to connect to the CUGrid directly. This is another reason why a management node is added to mediate between the CUGrid server and users. The CUGrid server executes commands and harmonizes all of the computers on the CUGrid.

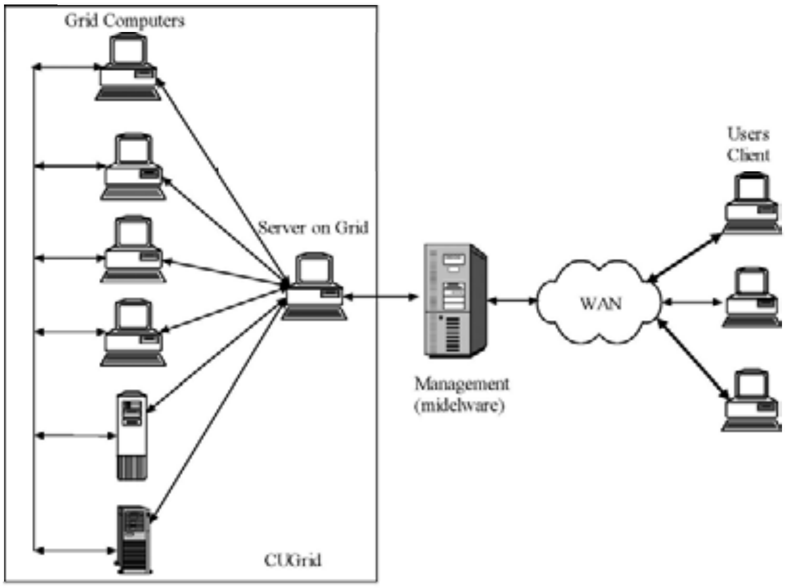


Figure 5.6: The framework of CUGrid based MM5 integration

In the following paragraphs, the three tiers are discussed. Online messages, which are standardized for communication among the users, the management and the CUGrid server, are also studied.

(1) Users

Users are located all over the world, wherever can be reached by the network. Users are command and message generators. By way of management, users can query information received from the server or send out a command to server. With an implemented windows interface, users can operate the MM5 simply.

(2) Management

Management serves as the coordinator for the whole system. This not only relays model operation commands, but also maintains system security. If a user wishes to use MM5 on the CUGrid, they would send a command to the network. The management node receives the commands, transforms their format from the user's operating system (Windows or Linux) into the CUGrid format (Linux) and sends them onto the CUGrid. After the model computation, the result is sent to the management node from the CUGrid. The management node will then perform a reverse procedure, having received the result from the CUGrid, transforming the result from the CUGrid format to the user format, and sending the result to users. The management node thus separates users from grid while connecting them both. The management node therefore improves the security of the CUGrid because users cannot disturb its environment.

There are several merits of this structure. The first is that users do not need to know how the grid calculates MM5 and performs the actual processing, but only how to validate and accept the computed results. The second merit is that the system is easy to maintain because the three tiers are relatively independent.

The management functions are shown as Figure 5.7.

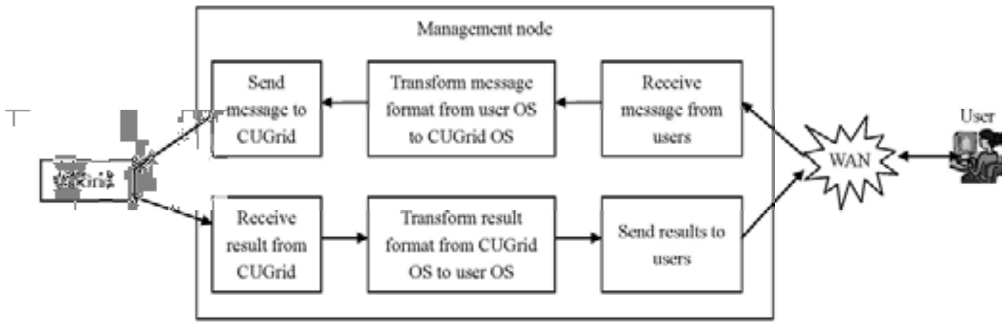


Figure 5.7: Functions of the management node

(3) Server on the CUGrid

The server is one node of the CUGrid computers. The server receives and executes commands, allocates jobs to other computers on the CUGrid, collects the computation output and sends this output and information to users. The operating system of the server in which MM5 is installed is Linux. MM5 runs through an executable CSH file. When commands are received by the server, they are written into a CSH file. After they are received, the file will be closed and then executed. When CSH files are running, information about the current state will be sent to users.

(4) Online commands for MM5 operations

The online commands for MM5 operations are used for communication between users and the server on the CUGrid through the management. The commands are standardized so that they are understood by all tiers. According to the functions, commands can be classified into three types, which are file transferring commands, executing commands and message strings. When a user orders the server to send back a file, he can send out a file transferring command to the server on the CUGrid. The server will identify which file should be transferred and then send this back to the user. Executing commands are used to request the server to execute some command lines, such as running a TERRAIN deck file. A message string is used by the server to reply to the current state of a process. For instance, when the server on the CUGrid runs "make intel", many middle processing messages will be generated. Using message strings, these processing messages may be sent back to users to inform them of the current state of the command being executed. The commands structures are enumerated in Figure 5.8.

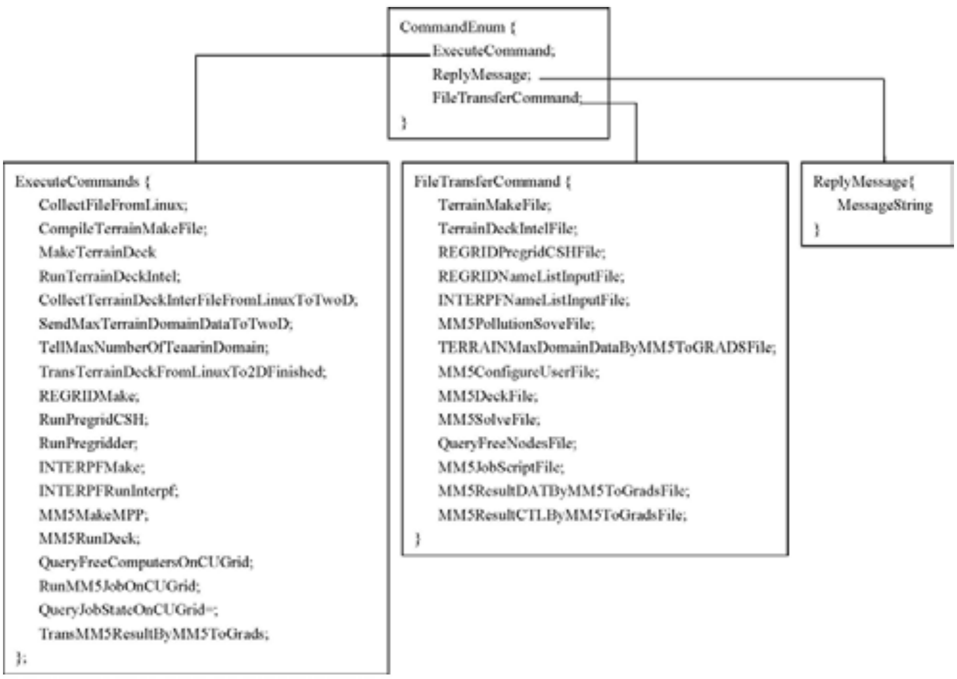


Figure 5.8: Command enumeration for MM5 operations

5.2.2 MM5 modules integration

There are eight modules which are used for preparing and running MM5: TERRAIN, REGRID, LITTLE_R/RAWINS, INTERPF, INTERPB, NESTDOWN, MM5 and GRAPH/RIP. In this research, there is no observation data used for mesoscale grid analysis, and so the LITTLE_R/RAWINS and related INTERPB modules are discarded. Horizontal interpolation for σ-coordinate data is ignored, and so the NESTDOWN program is also discarded. GRAPH/RIP is replaced by geographic information integrated visualization. At the same time, two new modules are added. In order to compile the air pollution source, a new module is created for preparing the pollution inventory, which is called the pollution source module. Another new module that is added is the CUGrid Computation Module, which is needed to run parallelized MM5. The working flow with the revised modules is shown in Figure 5.9.

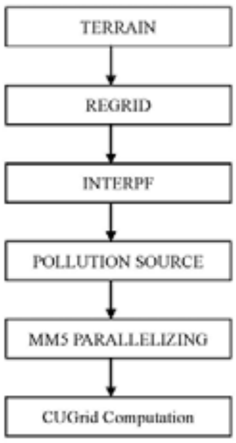


Figure 5.9: Working flow with revised modules for MM5 computation

In this research, MM5 is not run on one computer, but is distributed into three parts, which are the interface, preprocess and execution. The interface is installed on users' computers, and can be integrated through 2D or 3D visualization interfaces. The second part is preprocessing, which includes TERRAIN, INTERPF, REGRID, POLLUTION EDITOR and MM5 PARALLELIZING, and these are installed on the

CUGrid server. The third part is parallel computation, which is running on all of the selected computers on the CUGrid.

The overall methodology of MM5 integration is based on a C/S structure. The commands are generated by clients using a visualization operation interface, which will be discussed in the following sections, and these commands are sent to the server on the CUGrid (S-CUGrid). S-CUGrid executes the CSH file that contains the command lines, and replies with results and messages back to the clients.

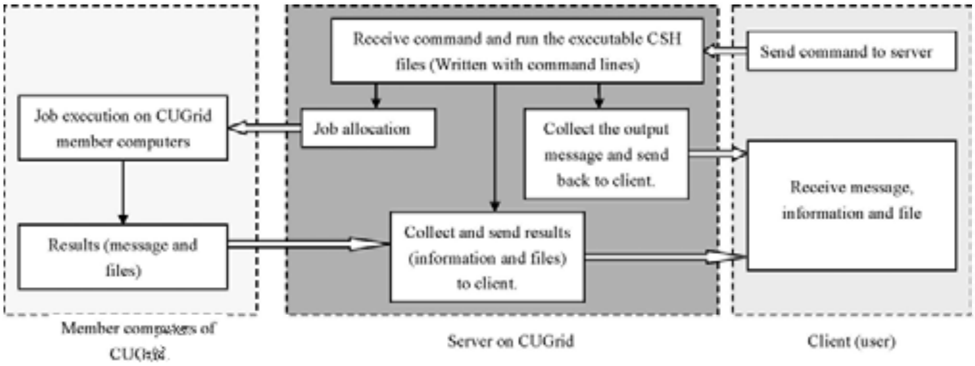


Figure 5.10: C/S structure of distributed operation on MM5

In the following sections, the various integrations of all the modules are each discussed in detail: TERRAIN integration, REGRID integration, INTERPF integration, pollution source compiling integration, MM5 preprocess integration and integration of MM5 computation on the CUGrid.

5.2.2.1 TERRAIN

Generally, there are several steps to run TERRAIN. The first step is to make some configurations which includes if NCAR Graphics is install, what is system compiler, setting system environment and so on. This configuration is written into one file named “Makefile”. After setting “Makefile”, compile it with command “make intel” and generate “terrain.deck”. Then user may run command “make terrain.deck” to compile generated “terrain.deck.X” (X is system compiler. For instance, if compiler is intel, X will be “intel”. So the generated file should be "terrain.deck.intel"). Following above step is to modify some parameters in "terrain.deck.X" according to users demanding. The final step is to run file "terrain.deck.X".

To divide and integrate functions to server(S-CUGrid) and client based on CUGrid, TERRAIN integration can be described as following steps.

- Client sends command (CollectFileFromLinux) to S-CUGrid to collect files. S-CUGrid transfers all initial files to clients (The "Makefile" is one of them).
- Client retrieves parameters from "Makefile" for TERRAIN configuration and displays them on operation interface.
- Client sets tag for NCAR state.
- Client determines system compiler.
- Client saves setting to a new created file which has name as "Makefile".
- Client sends new created "Makefile" to S-CUGrid. S-CUGrid receives new created "Makefile" and replaces the original one.
- Client sends command of compiling TERRAIN "Makefile" to S-CUGrid. S-CUGrid runs the CSH file with command lines. After compiling on S-CUGrid, a file named "terrain.deck" is generated and sent back to client.
- Client receives "terrain.deck", retrieves parameters from the file and fills in the form on user interface.
- Client fills in the form, saves the parameters into a new created file named same as the original one "terrain.deck".
- Client sends "terrain.deck" to S-CUGrid. S-CUGrid receives and replaces the original one.
- Client sends command (RunTerrainDeck) of running "terrain.deck" to S-CUGrid. S-CUGrid runs the file "terrain.deck" .

If all steps are correctly set and executed, files of "TERRAIN_DOMAIN(s)" are generated (s means number of domain). For example, if there are three domains, s should be 3, and there are three files generated, which are TERRAIN_DOMAIN1, TERRAIN_DOMAIN2 and TERRAIN_DOMAIN3. For each execution step on S-CUGrid, the processing output string will be sent to client. TERRAIN integration is shown in Figure 5.11.

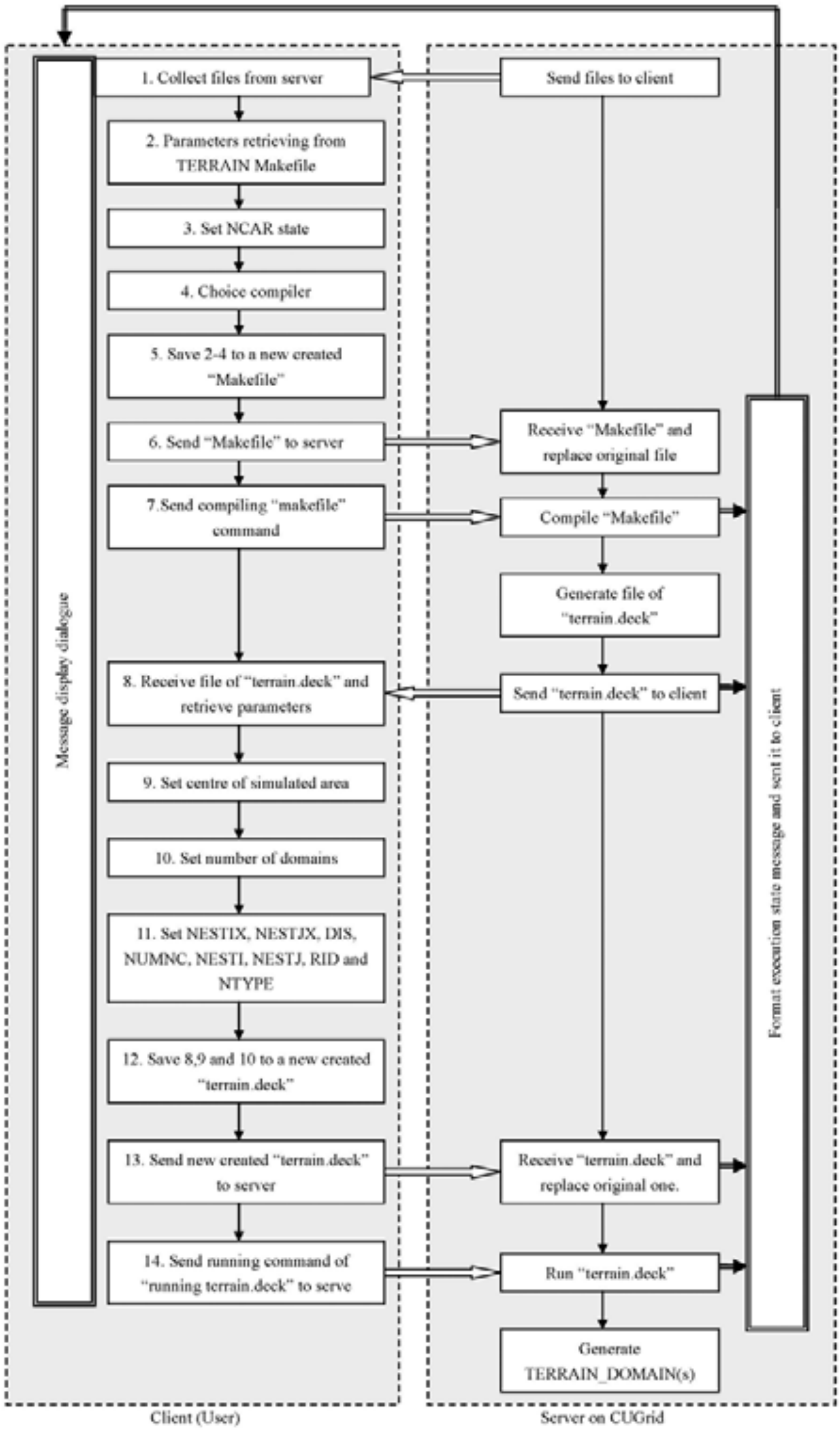


Figure 5.11: Integration of TERRAIN

There are many parameters should be set for TERRAIN. In this research, only some important parameters are opened to learning level users, which is listed in table 4.2.

Table 5.2 TERRAIN parameters opened to user

Parameter name	**Description**	**Value**
NCAR state	If NACR is installed on server of CUGrid.	Yes----NCAR is installed; No----NCAR is not installed;
Compiler	Which compiler is installed for server compiling	Intel; HP; IBM; SGI; SUN; Compaq; CRAY; Linux
Centre of simulated area	Centre point (Latitude, Longitude) of simulated area.	Latitude----Degree (-90°-90°); Longitude----Degree (-180°-180°);
MAXNES	Number of domain for process	Integer from 1 to 6.
NESTIX	Grid dimensions in Y direction	3n+1 (n is integer)
NESTJX	Grid dimensions in X direction	3n+1 (n is integer)
DIS	Grid distance.	3^n (n is integer) kilometers.
NUMNC	Mother domain ID	Integer
NESTI	Lower left I of nest in mother domain.	Integer
NESTJ	Lower left J of nest in mother domain.	Integer
RID	Radius of influence in grid units	
NTYPE	Input data resolution	1:1 deg (~111 km) global terrain and landuse 2:30 min (~56 km) global terrain and landuse 3:10 min (~19 km) global terrain and landuse 4:5 min (~9 km) global terrain and landuse 5:2 min (~4 km) global terrain and landuse 6:30 sec (~.9 km) global terrain and landuse

5.2.2.2 REGRID

There are about four obligatory steps for running REGRID. First is to compile REGRID with "make" command under folder of REGRID. The second step is to run "pregrid.csh". After input simulation start and end date and time on third step, the final step can be issued to run "regridder" to generate file named "REGRID_DOMAIN1".

The methodology for all steps' integration is shown as following.

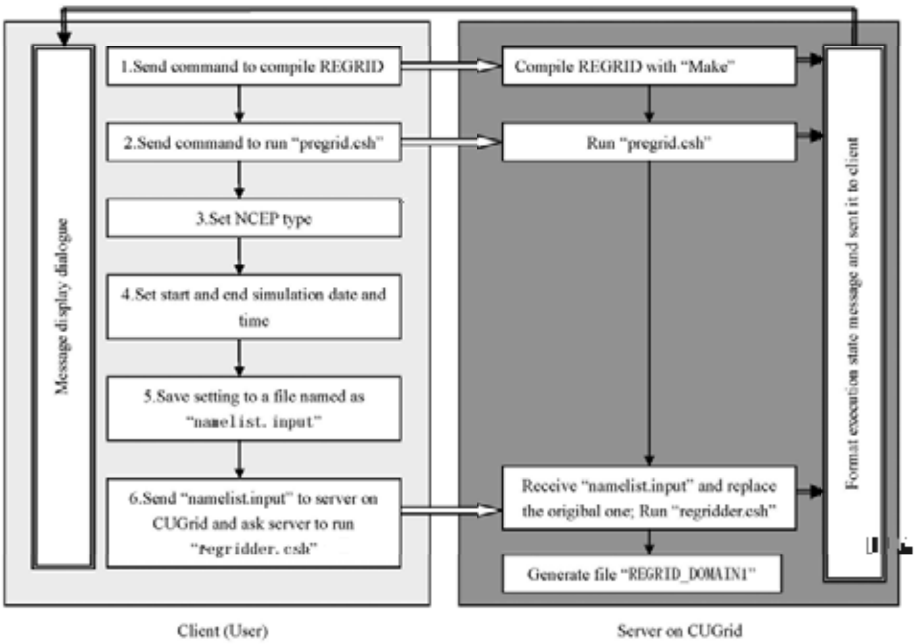


Figure 5.12: Integration of REGRID

- Client sends compiling REGRID command (REGRIDMake) to order S-CUGrid to compile configuration with command line "make";
- Client sends running REGRID command (RunPregridCSH) to S-CUGrid. S-CUGrid runs file of "pregrid.sch";
- Client sets NCEP data type;
- Client sets start and end date and time for simulation;
- Client saves NCEP, date and time into a new created file named "namelist.input";
- Client sends "namelist.input" to S-CUGrid;

- S-CUGrid receives "namelist.input" and replace the original one;
- Client sends command (RunPregridder) to run file of "regridder.csh"
- S-CUGrid runs file of "regridder.csh" to generate file(s) of REGRID_DOMAIN1.

For each execution on server, the processing output string is sent to client. REGRID integration can be explained as Figure 5.12.

5.2.2.3 INTERPF

INTERPF function is limited to take REGRID output data as input to generate a model initial, lateral boundary condition and a lower boundary condition. There are three key steps to run INTERPF, which are compiling INTERPF environment, reading data path and simulation period from REGRID setting, and finally running INTERPF.

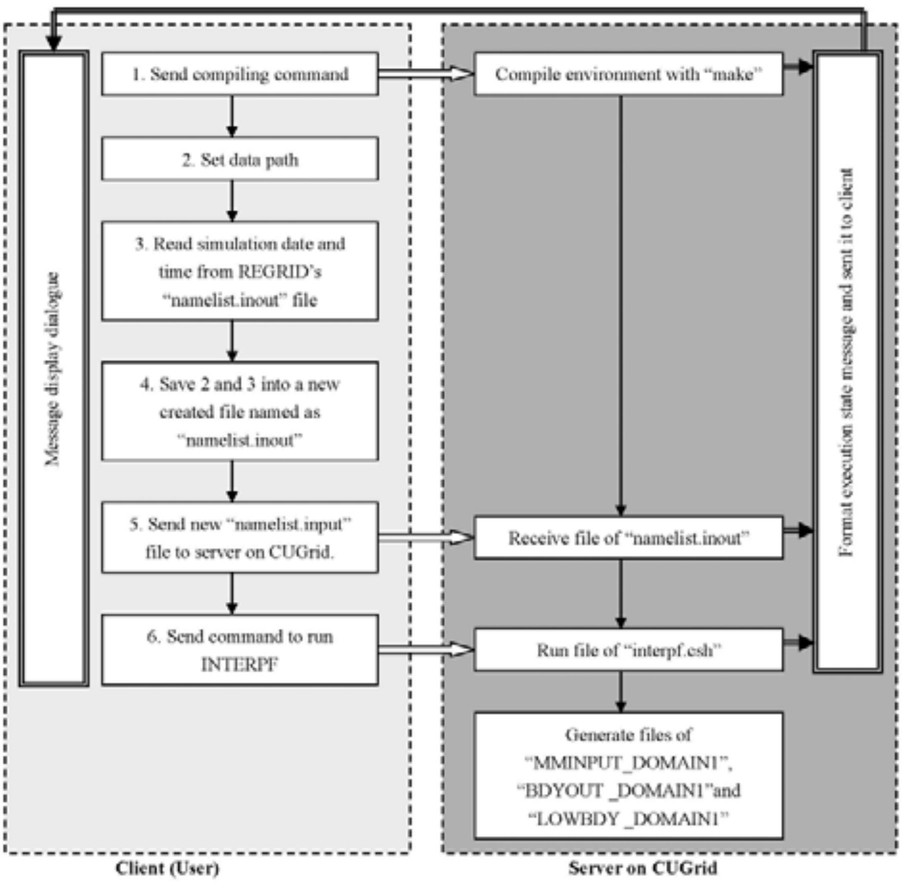


Figure 5.13: INTERPF integration

The steps for INTERPF integration can be described as followings.

- Client sends compiling command (INTERPFMakeIntel) to S-CUGrid. S-CUGrid compiles environment with the command line "make";
- Client reads data path and simulation date and time from "namelist.input" of REGRID and fill in a new created file named "namelist.input" of INTERPF;
- Client sends "namelist.input" of INTERPF to S-CUGrid. S-CUGrid receives the "namelist.input" file and replace original one.
- Client sends command (INTERPFRunInterpf) to S-CUGrid and orders it to run the command. S-CUGrid executes file "interpf" and generates "MMINPUT_ DOMAIN1", "BDYOUT_DOMAIN1" and "LOWBDY _DOMAIN1".

Like previous two steps, S-CUGrid will send back processing output strings back to clients. The integration of INTERPF is shown as Figure 5.13.

5.2.2.4 Pollution sources compiling

Pollution source edition is a new added module in this research. With this program, pollution source can be edited, such as adding or deleting a pollution source, closing or opening one or several cities' industries, closing or opening one kind of industries and so on. This module is added to generate some scenarios for simulation to analysis relationship between air pollution distribution and pollution sources.

In this research, pollution sources are not generated by simulation with some kind of tools, such as SMOKE, but survey records stored in database. Generally, pollution sources can be classified as point sources, line sources and area sources. But for MM5 based air pollution dispersion computation, all kinds of pollution sources are simplified into gridded point pollution sources, which can be described by Figure 5.14. In Figure 5.14 (A), the whole pollution sources are plotted into I*J grids. For each grid, there may be point sources, line sources and area sources are all included like Figure 5.14(B). But for MM5 computation of air pollution dispersion, these three kinds of sources are treated as point pollution source, whose synthetical emission is the addition of original three kinds of emission like Figure 5.14 (C). The location of synthetical emission is at the centre of the grid.

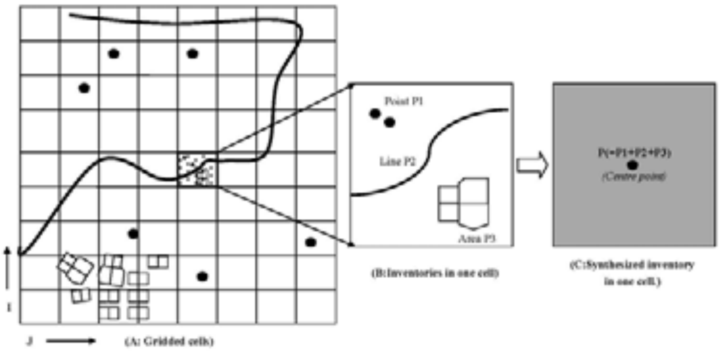
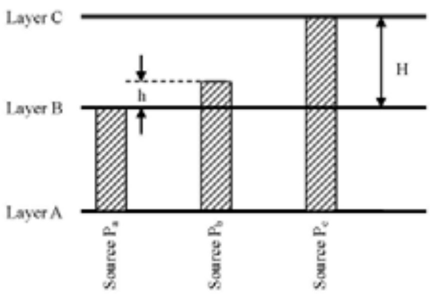


Figure 5.14: Pollution source horizontally plot

Pollution sources should also be allocated on vertical layers. If height of a pollution source is exactly match to a layer height, this source will be allocated to the matched layer. If a pollution source is between two layers, the source will be allocated to two neighbor layers by inverse proportion of distance. For example in Figure 4.15, three pollution sources are considered. The height of source P_a and source P_c are exactly same as heights of layer B and layer C respectively. So, P_a is totally allocated to layer B and P_c is totally allocated to layer C. For source P_b, because it is just between layer A and layer C, the P_b will be divided into the two layers with $(P_b*h)/H$ for layer B and $(P_b*(H-h))/H$ for layer C. So, the final sources for three layers are 0 for layer A, $(P_a+(P_b*h)/H)$ for layer B , and $(P_c+(P_b*(H-h))/H)$ for layer C.



$$\begin{cases} Source\ on\ Layer\ A = 0 \\ Source\ on\ Layer\ B = P_a + P_b \bullet h / H \\ Source\ on\ Layer\ C = P_c + P_b \bullet (H-h)/H \end{cases}$$

Figure 5.15: Pollution source allocation on vertical layers

To compile pollution sources for air pollution computation, several steps should be carried out, which include selecting the pollution sources covered by simulation area, synthesizing sources into grids, saving synthesized sources to formatted codes which are in line with "solve.f", and inserting the formatted codes into "solve.f".

The methodology of implementation pollution sources edition can be described by following steps.

- Client collects pollution source file "solve.f" from S-CUGrid.
- Client edits pollution source and saves it to text file with the format as "Latitude; Longitude; plume height; plume velocity; SO_2 concentration ".
- Client transforms formatted pollution text file to "solve.f" readable format. Pollution source horizontal plot and vertical allocation are included by this step.
- Client inserts formatted codes into "solve.f".
- Client sends "solve.f" file to S-CUGrid.
- S-CUGrid receives "solve.f" file and replaces original one.

The integration of pollution source edition is shown as Figure 5.16.

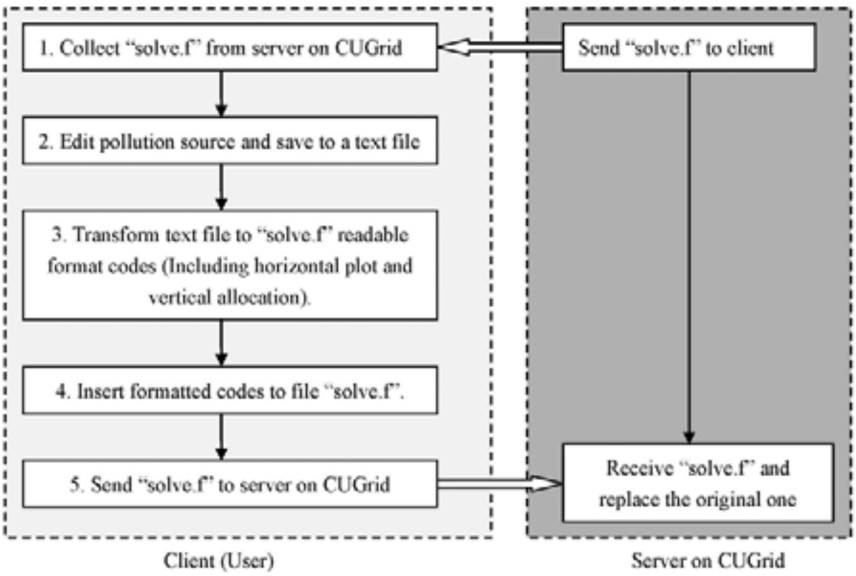


Figure 5.16: Integration of inventory compiling

5.2.2.5 MM5 preprocessing

This module is used to generate execution file of MM5. Before generating parallel executable file, some lines in the file of "configure.user" should be changed according to previous modules' settings, such as number and maximum dimensions of domain, system compiler, and so on. After that, running "make mpp" and "make mm5.deck" to generate two files of "mm5.mpp" and "mm5.deck". With "mm5.deck", some changes should be made to be consistent with TERRAIN domain parameters which are set in TERRAIN module. After run "mm5.deck", linking output data of INTERPF and TERRAIN, the parallelized executable MM5, which integrates not only atmosphere environment modeling and computation but also air pollution dispersion computation, is ready for calculating on CUGrid.

The integration of MM5 preprocessing is described as followings.

- Client saves number of domains, maximum dimensions of all domains, number of half sigma levels, and system compiler setting into file of "configure.user".
- Client sends edited file "configure.user" to S-CUGrid. S-CUGrid receives the file and replaces the original one.
- Client sends command (MM5MakeMPP) to S-CUGrid. S-CUGrid runs "make mpp" and "make mm5.deck" to generate "mm5.mpp" and "mm5.deck".
- S-CUGrid sends new generated "mm5.deck" back to client. Client opens the file and reads parameters for revision.
- Client revises and saves edited parameters in "mm5.deck". The parameters include forecast length, forecast time step, frequency of data outputting, frequency of saving output data, domains' attributes set by TERRAIN program, and so on.
- Client sends revised "mm5.deck" to S-CUGrid and asks it to run "mm5.deck" by command "MM5RunDeck". S-CUGrid receives "mm5.deck", replaces the original one, and runs "./mm5.deck".
- S-CUGrid links output data generated by TERRAIN and INTERPF into folder of

"Run".

The integration of MM5 preprocessing is shown as Figure 5.17.

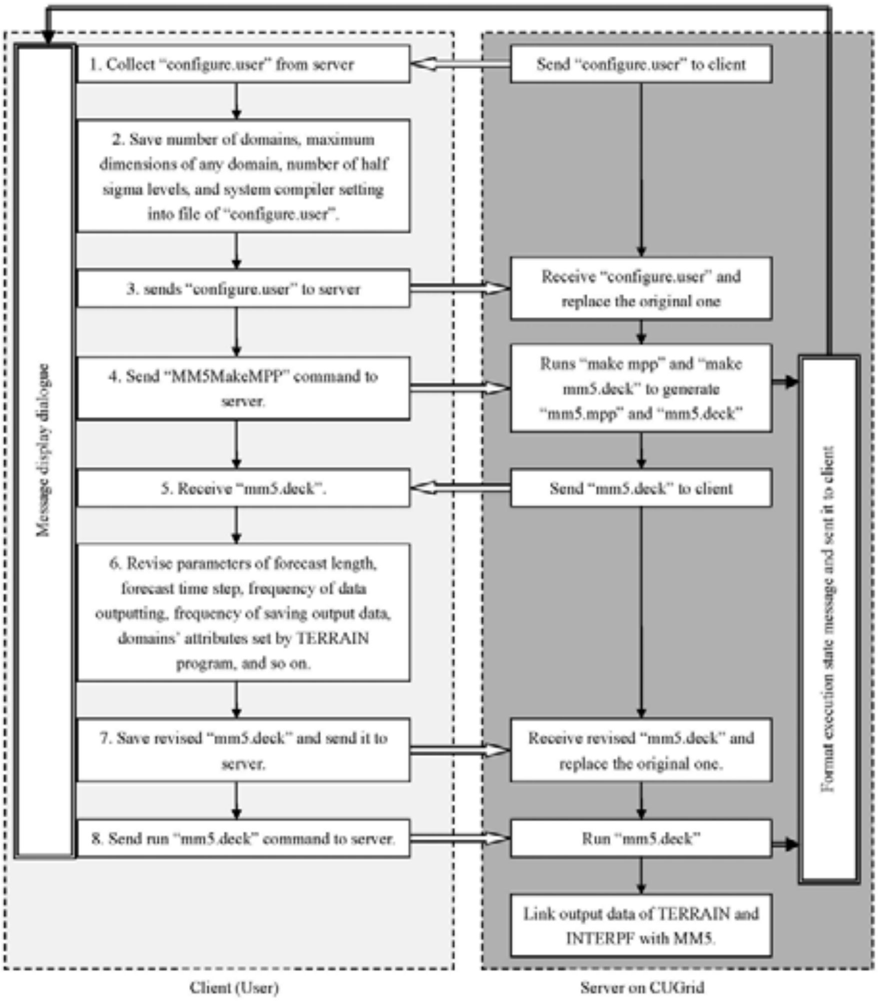


Figure 5.17: MM5 preprocessing integration

5.2.2.6 MM5 computation on CUGrid

This is the last step for CUGrid based MM5 computation. It finds the free computers on CUGrid and submits MM5 computation job to them. The methodology to implement CUGrid based MM5 computation can be shown as following.

- Client sends command (QueryFreeComputersOnCUGrid) to S-CUGrid to search free nodes. S-CUGrid execute the command, replies searched nodes and sends back to client.

- Client selects some or all of the free nodes for parallel MM5 computation.
- Client sends the select nodes tags to S-CUGrid and asks it to submit parallel job to selected nodes by command of "RunMM5JobOnCUGrid".
- S-CUGrid submits parallel job to CUGrid;
- Client sends query state command (QueryJobStateOnCUGrid) to S-CUGrid to query current computation state on CUGrid. S-CUGrid executes the query state command and reply results back to client.
- After parallel MM5 execution on S-CUGrid, client sends command of "TransMM5ResultByMM5ToGrads" to order S-CUGrid to do some basic data format transformation. The transformed data will be sent back to user for post processing.

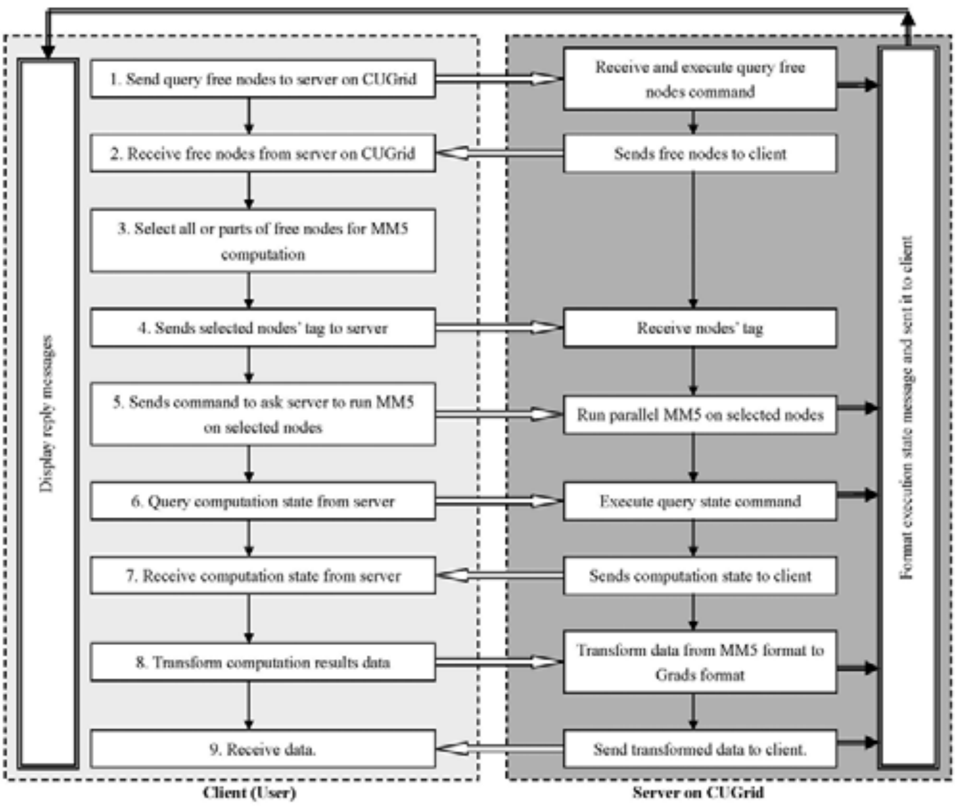


Figure 5.18: The Integration of CUGrid based MM5 computation

5.2.3 New interface for remote operation on MM5

To cooperate with above integration, a new interface is designed and implemented.

Generally, MM5 is run on Linux operation system. SSH secure shell (see www.ssh.com) is often used as interface for remote operation on MM5. This interface is not inconvenient for integration by CVGE and bed for geo-collaboration. At the same time, SSH only support command lines, which are almost impossible for layman to execute them.

However in this new interface, above shortcomings are solved by several steps. **First,** new interface converts many important command lines into windows buttons. Participant can execute command by simply clicking buttons. **Second**, the sequence of buttons is fixed according to commands sequences. For example, before running "terrain.deck" in TERRAIN module, the "makefile" should be compiled first. Thus, in the new interface, the "Run Terrain" button is imitated as "disable". After clicking "Make compiler", the "Run Terrain" button will be enabled for operation. This setting will decrease error operations. **Third**, new interface also supply professional operations for experts. The professional configure MM5 with SSH is by configure files. Thus, the new interface inherits this kind of operation and supply experts configure file either.

In the first chapter, one challenge of air pollution simulation is that to retrieve some information from geographic information to support air pollution modeling. Thus, for this new interface, the challenge has been tackled. The new model operation interface is coupled with geo-visualization interface where geo-data and geo-information are integrated. In this research, the geo-information which MM5 needs is the simulation area including location of low left corner, location of up right corner, and the numbers of grids at horizontal and vertical directions. The area can be gotten absolutely from geographic information, which will be discussed in following section.

Following sections will firstly discuss in detail how to select research area from geo-information integrated map and input the retrieved information into MM5. Then, the interfaces for operation modules are implemented.

5.2.3.1 Simulation area visualized selection implementation

To use MM5 for air pollution computation, several domains should be defined first. The outer domain is coarse domain, which is used to prepare coarse parameters. The coarse parameters can be input into inner domain to get precise results. In this research, three domains are adopted. The outer course domain, or domain Ⅰ (D01 as in Figure 5.19), is fixed in the rectangle from (108.469, 19.0721) at the bottom left corner and (117.622, 27.4653) at the top right corner. There are 118 grids on north-south direction and 127 grids on west-east direction with resolution of 27km on both directions. The second domain, or domain Ⅱ(D02 as in Figure 5.19) is child domain for domain Ⅰ and mother domain for domain Ⅲ (D03 as in Figure 4.19). The bottom left corner of domain Ⅱ is at (46, 42) in the coordinate of domain Ⅰ. For domain Ⅱ, it has 115 grids and resolution of 9km on both north-south direction and west-east direction. The third domain is domain Ⅲ, which is the most inner one. Three domains' regions are shown in Figure 5.19.

In fact, coarse and inner domains can be movable according to study area. But in this research, the research area is focused in PRD. So the domain Ⅰ and domain Ⅱ are fixed. Only domain Ⅲ is movable limitedly within domain Ⅱ and editable. In the following paragraphs, all the domain editions are limited on domain Ⅲ.

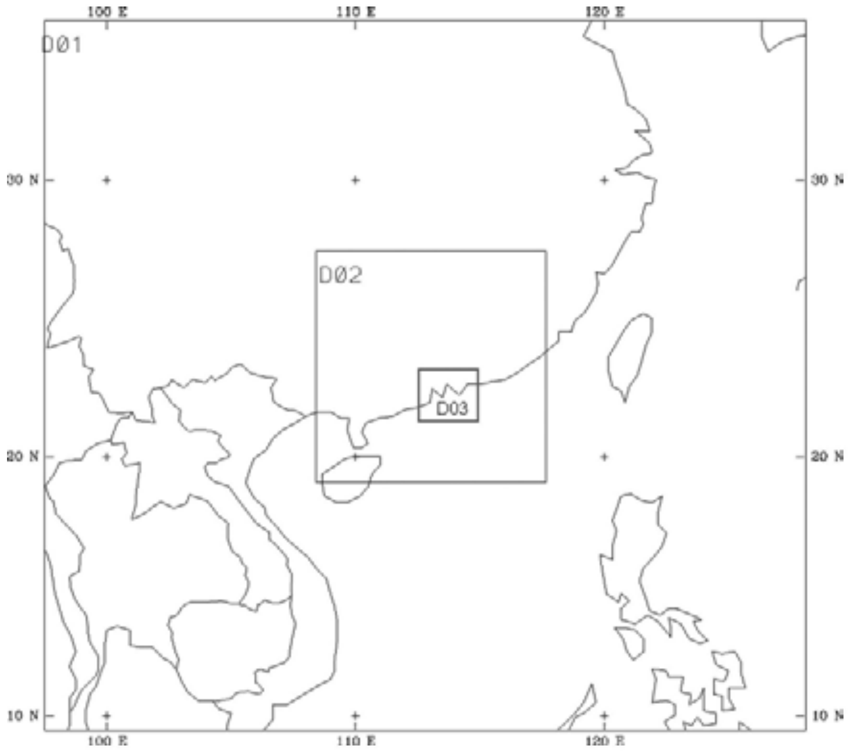
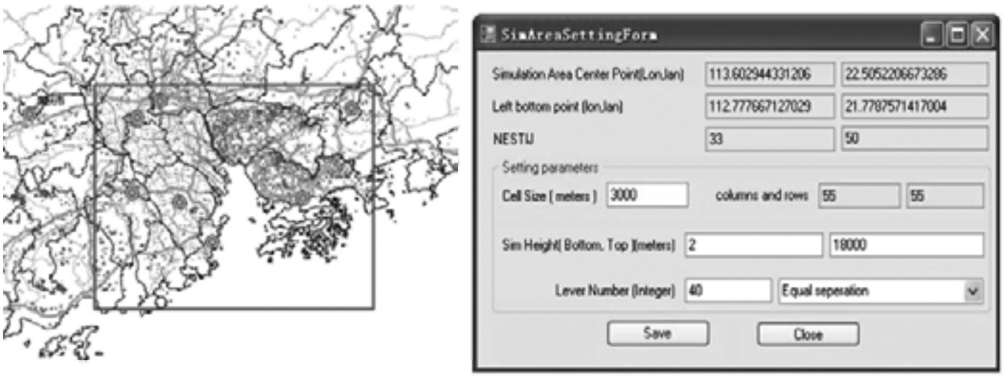


Figure 5.19: Three domains for PRD air pollution simulation (D01 and D02 are fixed. Only D03 is editable and opened to user.)

Unlike traditional method for simulation area setting by data set and visualization check in GrADS, this research does an intuitive way to select simulation area and set parameters for domain III, which is shown in Figure 4.20. Simulation area is visually "drag and drop" a rectangle on map integrated with geographical information like in Figure 5.20(A). After drawing simulation area rectangle, two corner points are retrieved from geographic information, which are lower left corner point and top right corner point. After two corners set, interface in Figure 5.20 (B) will pop-up to receive model computation resolution and calculate numbers of grids at horizontal and vertical directions, and coordinate value of domain III. All of above values are parameters which will be automatically input into one of MM5 modules----TERRAIN to support MM5 computation. The interface of TERRAIN will be discussed following.



(A) (B)

Figure 5.20: Simulation area selection ((A) Simulation area selection by drag and drop; (B) Parameters setting for selected area.)

5.2.3.2 Windows style interface for MM5 remote operation

(1) TERRAIN operation interface

TERRAIN operation is visualized as two parts. One is for computation environment setting, such as NCARG state, system compiler. This part can be set one time for all. The setting will be saved to a file named "Makefile" by clicking "Save Setting" button. Before compiling "Makefile" by "Make compiler" button, clicking "Professional Check" button to open "Makefile" for checking. The second part is domain setting. Some parameters are set by visualized simulation area setting automatically, such as central location and the third domain parameters. These parameters can also be revised according to user demanding.

At the bottom is message display dialogue, which is used to display replied message generated by server on CUGrid during execute command. TERRAIN visualization is shown as Figure 5.21.

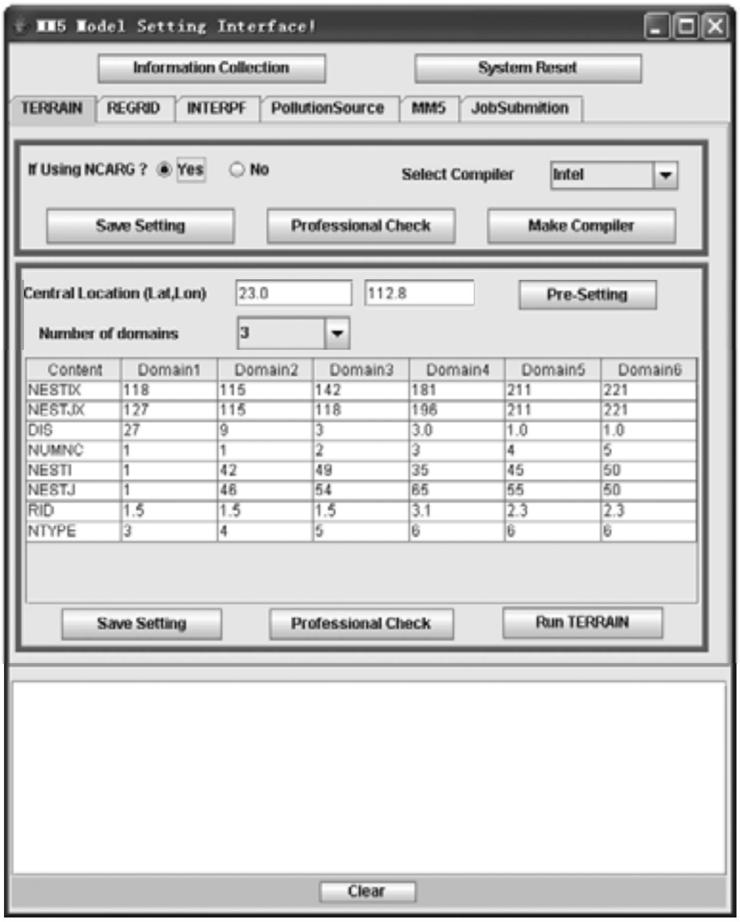


Figure 5.21: TERRAIN interface

(2) REGRID operation interface

Two steps for REGRID implementation. The first step is to compile REGRID environment. In the second step, simulation time period is set and saved to "Pregrid.csh" file by clicking "Save Setting" button. To check and run "Pregrid.csh", "Professional Check Pregrid.csh" button and "Run PREGRID" button are supplied. To run "Regridder", a button named "Run Regridder" is implemented. Button "Professional Check Namelist.input" is used to check if there is any wrong in file "Namelist.input" which is used to support run "Regridder". The lower dialogue is also used to display server reply message.

REGRID integration implementation is shown as Figure 5.22.

Figure 5.22: REGRID interface

(3) INTERPF operation interface

"Make intel" button is used to compile environment for INTERPF. After compiling, click "load setting" to read previous simulation time period setting by REGRID and fill them in text forms. Click "Save setting" button and "Professional Check" button to save and check setting file. After finishing all the setting and checking, click "Run InterPF" button to run INTERPF. Above description is shown as Figure 5.23.

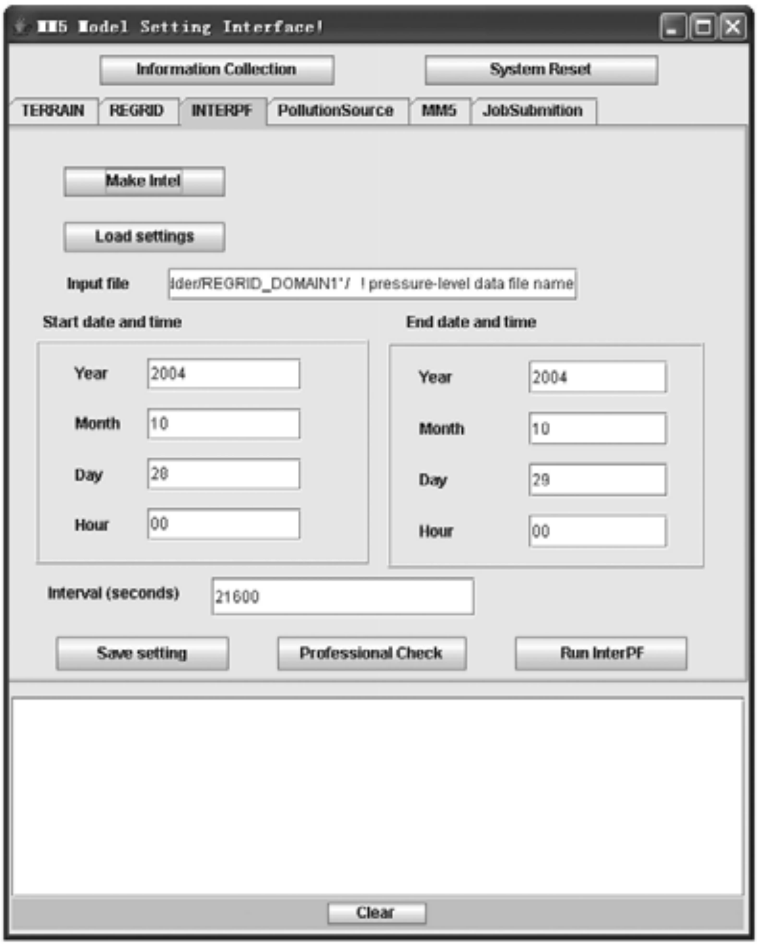


Figure 5.23: INTERPF interface

(4) Pollution source compiling interface

From the visualized interface, inventory is classified according to cities. Under city classification, pollution type is used as sub-classification. Pollution sources are structured as a tree. Pollution sources can be selected by click the check-box before each tree item. After checking item, click "Save Setting" button to save setting into a file, transform the file data into a "solve.f" readable format, insert transformed code into "solve.f", and send inserted "solve.f" file to server on CUGrid.

The pollution source edition implementation interface is shown in Figure 5.24.

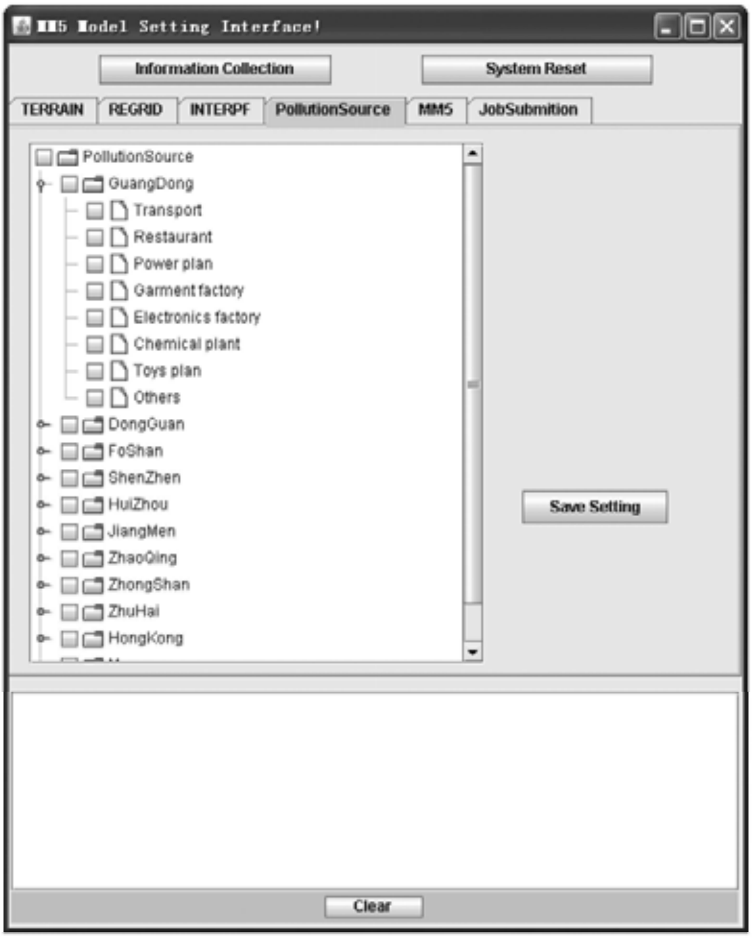


Figure 5.24: Pollution source compiling interface

(5) MM5 preprocessing operation interface

This visualized interface is used to generate parallel executable file which will be used to be submitted to CUGrid for computation. In the visualization interface, click "Get PreModule Setting" to get all the needed parameters, which are set by previous steps, to initiate this program's parameters. After parameter setting click the first "Professional Check" button (the top one) to check if there any wrong in "configure" file for MM5 program. If there is no wrong in "configure", click "MakeMPP" to generate parallel executable file and "mm5.deck". Click "Load Setting" button to read information from "mm5.deck" and fill the text forms. Keep default setting or do some revisions on text forms, then click "Save Setting" to save above text contents into "mm5.deck". After checking "mm5.deck" by clicking "Professional Check" button (the lower one), click "Run Deck" to generate final parallel executable file which will be submitted to CUGrid for computation.

Figure 5.25 shows the visualization interface of MM5 preprocessing.

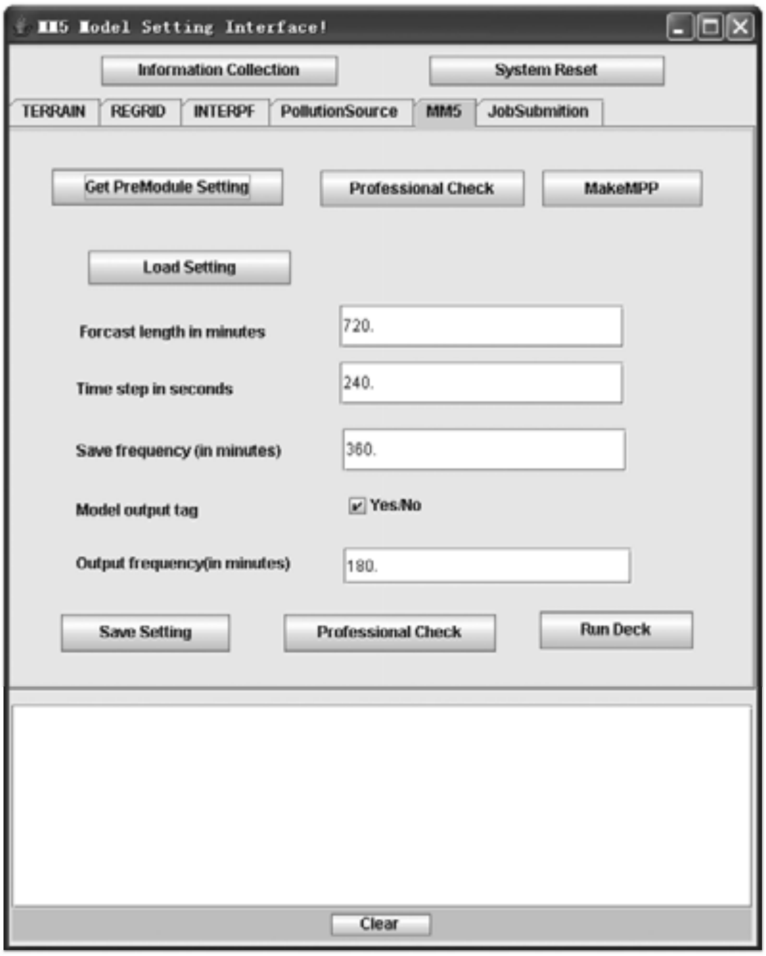


Figure 5.25: MM5 preprocessing interface

(6) Operation interface for MM5 computation on CUGrid

Figure 5.25 gives the visualization of CUGrid based job processing. On the left is the tree for free nodes on CUGrid. On the right is area for the operation items. Two types of free nodes are supported for model computation. One is monopolized type and the other one is shared type. For monopolized type, free nodes are only used for MM5 computation, which makes user to start computation immediately. For shared type, all the applications which are running on CUGrid share all the nodes. So, if a user wants to submit a job to some nodes, he must wait until all the selected nodes are free. In the interface of CUGrid based job processing visualization, the "Switch" is used to determine what type of free nodes will be used. Checking "Switch" means to use shared nodes for computation, otherwise to use monopolized nodes for computation. If using shared nodes for computation, the buttons of "Search nodes" and "Create FreeNodes Tree" will be enabled. Clicking "Search free nodes" is to send searching command to server and receive results. "Create FreeNodes Tree" is used to create free nodes tree according to searched free nodes. After selection of free nodes, click "Submit Job to Nodes" button to create script of submitting parallel MM5 job to selected nodes. "Run Job on CUGrid" button is to run parallel job on CUGrid. During job computation, computation states can be gotted from clicking button "Query job State". After model computation, the results can be transformed by two buttons, which are "Trans MM5Result" button and "TransToVGE Format" button.

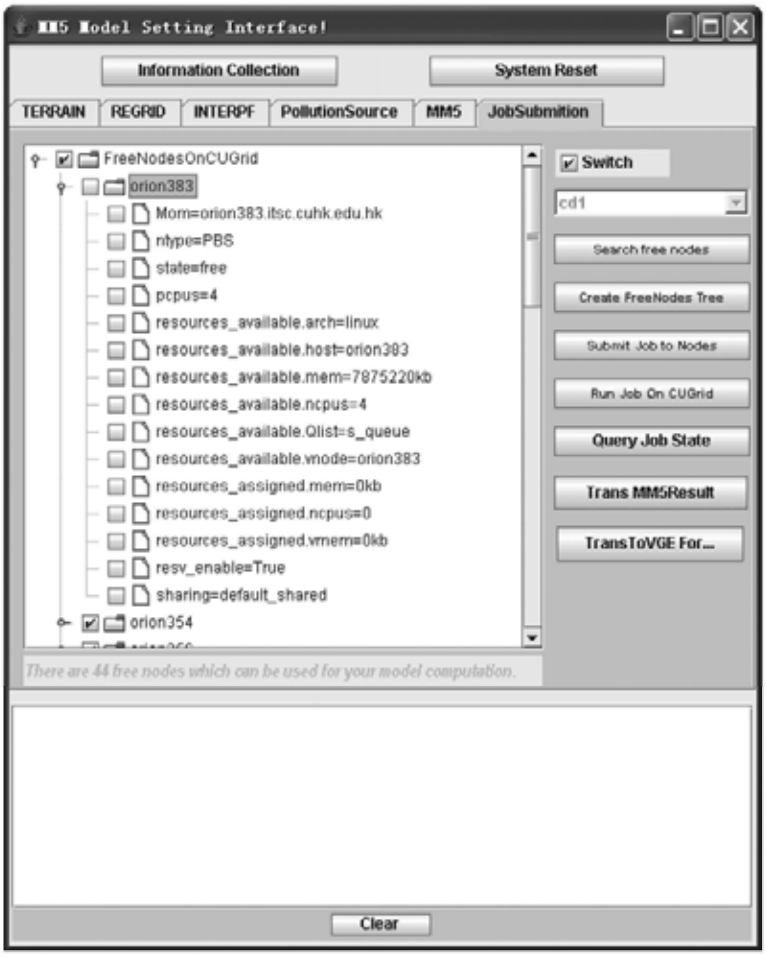


Figure 5.26: Interface of MM5 job computation on CUGrid

For above interfaces' implementation, study area selection is implemented with C# language on Visual Studio.net 2005. The other interfaces are implemented with java language. Study area selection is integrated into CVGE visualization platform, which will be discussed in detail in the following chapter. MM5 operation interfaces are portable and operation system independent. Thus the operation interface are easily be integrated as an embedded component. The open parameters and files of MM5 modules interfaces are not limited to those mentioned but can be extended.

5.3 Gaussian Plume Model Parallel Computation

In this research, the Gaussian plume model is used to study single or multiple point emission dispersion. In order to accommodate online discussion, the time consumption of model computation must be minimized to guarantee that the operation replies in real time or near real time. Unlike the MM5 computation on the CUGrid above, a different method is used to improve Gaussian dispersion model computation, which is parallel computation using cluster PCs.

5.3.1 Parallel arithmetic for Gaussian plume model computation

The number of point emissions is p. Each layer is plotted into regular grids that have m columns by n rows. Therefore, pollution sources and the research area can be regularly plotted, as shown in Figure 5.27. This forms the foundation for the following parallel arithmetic.

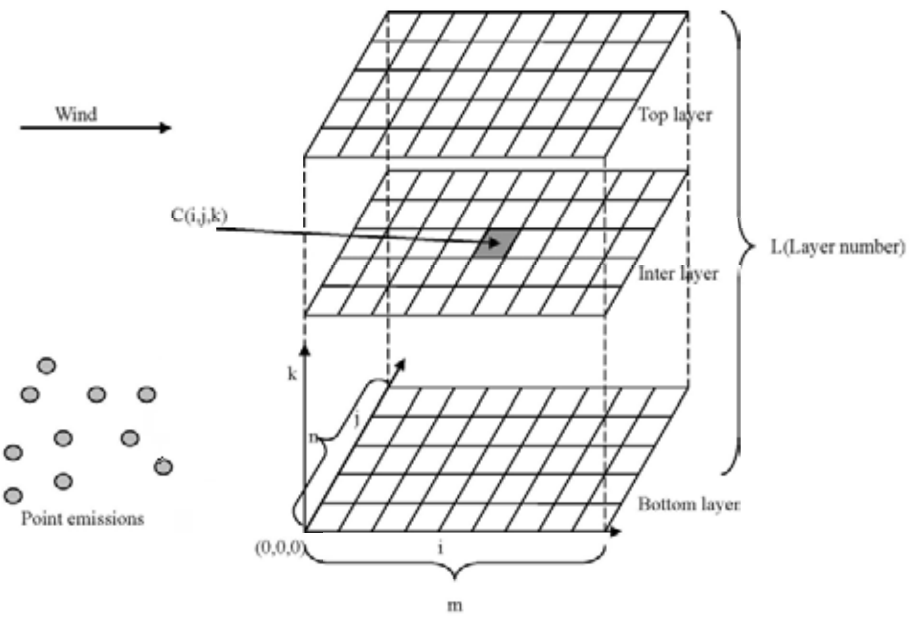


Figure 5.27: Plot of Gaussian Plume Model (C(i,j,k) is the pollution concentration on the grid with a location of (i,j,k))

The parallel computation framework is shown in Figure 5.28. This is based on three tiers of architecture, composed of users, task management and self-controlling computers. The user tier initiates the Gaussian plume model parameters. The task management processes parallel arithmetic to divide the total task into sub-tasks, and merge sub-results into the total result. The self-controlling computers execute the Gaussian plume model. Each of the self-controlling computers has one copy of Gaussian plume model. The computation process is as follows: after setting the model parameters, parallel arithmetic divides the computation task into sub-tasks and allocates these sub-tasks to the connected self-controlling computers. Each computer will then run the Gaussian plume model to generate a sub-result. When all of the sub-tasks computations are finished, the sub-results will be sent back to the task management for post sub-task synthesizing.

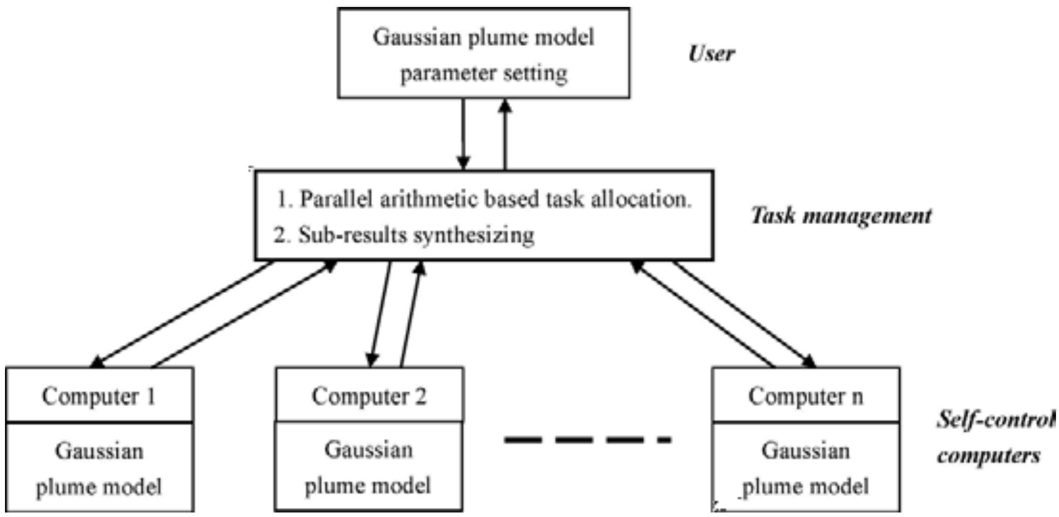


Figure 5.28: Framework of the Gaussian plume model parallel computation

The key issue in Figure 5.28 is the parallel arithmetic, which involves dividing the computation task into sub-tasks, allocating sub-tasks to computers and combining the sub-results into a synthetic result. In the following sections, the different types of parallel arithmetic design are outlined.

Before designing the parallel arithmetic, the parameters and operators should first be stated.

m,n,L: These have same meaning as shown in Figure 5.27;

NUMe: The number of point emissions;

NUMc: The number of self-controlling computers;

i: The grid column count for each layer ($1 \le i \le m$);

j: The grid row count for each layer ($1 \le j \le n$);

k: The layer count($1 \le k \le l$);

A: The computer index ($1 \le A \le NUMc$);

q: The point emission count ($1 \le q \le NUMe$).

C(i,j,k,A): The pollution concentration at location (i,j,k) calculated by computer A;

C(i,j,k,q,A): The pollution concentration at location (i,j,k) impacted by point emission q and calculated by computer A;

The operators "/" and "%" refer to getting an integer part and a remainder part of a division result, respectively. For instance, 5/3=1, while 5%3=2. Another operator is [ ], which is described in the equations below.

$$[Q/W] = \begin{cases} Q/W & (Q\%W = 0) \\ Q/W + 1 & (Q\%W \neq 0) \end{cases}$$

(1) Multiple point emissions based parallel arithmetic

For this arithmetic, multiple point emissions are sent to multiple computers. Each computer will calculate all of the grids on all layers.

Taking I as integer and R as remainder of NUM_e divided by NUM_c:

$$I = NUM_e / NUM_c$$
$$R = NUM_e \% NUM_c$$

According to the R value, two conditions can be made.

Condition I: R=0

In this condition, each self-controlling computer will process I point emissions for all of the grids. For computer A, the allocated point emissions range is:

[((A-1)*I+1), A*I]

For grid (i,j,k), the pollution concentration calculated by computer A can be calculated by using:

$$C\ (i,j,k,A) = \sum_{q=[(A-1)*I+1]}^{A*I} C(i,j,k,q,A) \qquad \text{..................(5. 11)}$$

Therefore, the final concentration for each grid C(i,j,k) can be synthesized by using:

$$C\ (i,j,k) = \sum_{A=1}^{NUM_c} C(i,j,k,A) = \sum_{A=1}^{NUM_c} \sum_{q=[(A-1)*I+1]}^{A*I} C(i,j,k,q,A) \qquad \text{..................(5. 12)}$$

Condition II: R≠0

In this condition, there are NUM_c–R computers processing I point emissions, while R computers process I+1 point emissions. There are two methods that can be used to determine how the reminder R should be allocated. One is head allocation, and the other tail allocation, both of which are shown in Figure 5.29. For both head allocation and tail allocation, the number of I is allocated using averages to NUM_c computers first, and the remainder R pollution sources are then allocated from head or tail one computer at a time.

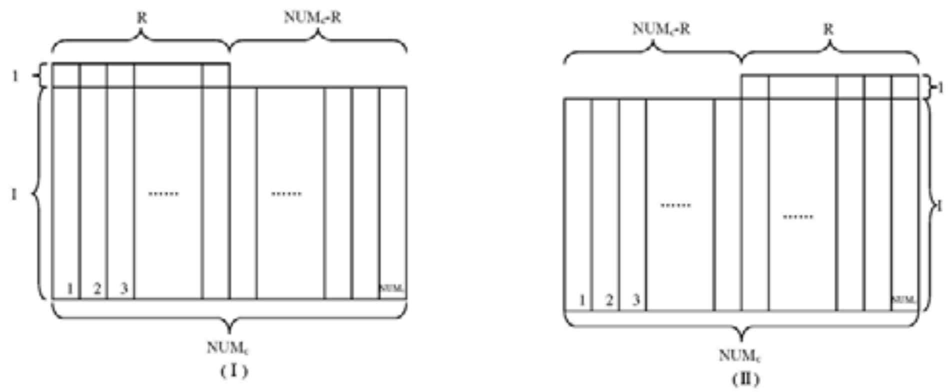


Figure 5.29: Point emissions allocation ((Ⅰ) concerns head allocation, which means that the remainder emissions are equally allocated to head R computers. (Ⅱ) concerns tail allocation, which means that the reminder emissions are equally allocated to tail R computers.)

For head allocation arithmetic, computer A will allocate with a point emission range calculated from:

$$\begin{cases}[(A-1)*(I+1)+1,\ A*(I+1)] & (A \le R)\\ [(I+1)*R+(A-R-1)*I+1,(I+1)*R+(A-R)*I] & (A > R)\end{cases}$$

Computer A calculates pollution concentration in grid (i, j, k,A) from:

$$\begin{cases}C\ (i,j,k,A)=\sum_{q=(A-1)*(I+1)}^{A*(I+1)} C(i,j,k,q,A) & (A \le R)\\ C\ (i,j,k,A)=\sum_{q=(I+1)*R+(A-R-1)*I+1}^{(I+1)*R+(A-R)*I} C(i,j,k,q,A) & (A > R)\end{cases} \quad \text{......................(5. 13)}$$

The synthesized grid concentration at location (i,j,k) is:

$$C\ (i,j,k)=\sum_{A=1}^{NUM_c} C(i,j,k,A) \quad \text{..........................(5.14)}$$

For tail allocation arithmetic, computer A will allocate with a point emission range calculated using:

$$\begin{cases}[(A-1)*I+1,\ A*I)] & (A \le NUM_c - R)\\ [I*(NUM_c-R)+(A-NUM_c+R-1)*(I+1)+1,\ I*(NUM_c-R)+(A-NUM_c+R)*(I+1)] & \\ & (A > NUM_c - R)\end{cases}$$

The pollution concentration calculated by computer A in grid (i, j, k,A) is:

$$\begin{cases}C\ (i,j,k,A)=\sum_{q=(A-1)*I+1}^{A*I} C(i,j,k,q,A) & (A \le NUM_c - R)\\ C\ (i,j,k,A)=\sum_{q=I*(NUM_c-R)+(A-NUM_c+R-1)*(I+1)+1}^{I*(NUM_c-R)+(A-NUM_c+R)*(I+1)} C(i,j,k,q,A) & (A > NUM_c - R)\end{cases} \quad \text{......(5.15)}$$

The synthesized grid concentration at location (i,j,k) is :

$$C\ (i,j,k)=\sum_{A=1}^{NUM_c} C(i,j,k,A) \quad \text{..............................(5.16)}$$

(2) Multiple layers based parallel arithmetic

This arithmetic is to allocate the average number of layers to self-controlling computers.

$$I = L / NUM_c$$
$$R = L \% NUM_c$$

According to the R value, two conditions can be made.

Condition I: R=0

In this condition, all self-controlling computers process the same number of layers, which is the average I. Computer A will be allocated with layers range as:

[(A-1)*I+1, A*I]

The pollution concentration at the location (i,j,k) is:

$$C\ (i,j,k) = \sum_{q=1}^{NUMe} C(i,j,k,q,A) \qquad (A = \left[\frac{k \bullet NUM_c}{L}\right]) \qquad \text{……..……….(5.17)}$$

The sub-results do not need to be synthesized to get the final result.

Condition Ⅱ: R≠0

Using head allocation, layers allocated to computer A arithmetic is:

$$\begin{cases} [(A-1)*(I+1)+1,\ A*(I+1)] & (A \le R) \\ [R*(I+1)+(A-R-1)*I+1, R*(I+1)+(A-R)*I & (A > R) \end{cases}$$

The pollution concentration at location (i,j,k) is:

$$C\ (i,j,k) = \sum_{q=1}^{NUMe} C(i,j,k,q,A) \quad (\begin{cases} A = \left[\dfrac{k}{I+1}\right] & (k \le R) \\ A = R + \left[\dfrac{k-R}{I}\right] & (k > R) \end{cases}) \text{………….(5.18)}$$

As condition R=0, in this condition, there does not need synthesizing sub-results.

(3) Grids based parallel arithmetic

All layers are allocated in the same way. For each layer, there are m*n grids. These grids can be divided and allocated to different computers. Below Figure 5.30 gives the description about grids based parallel arithmetic with three computers.

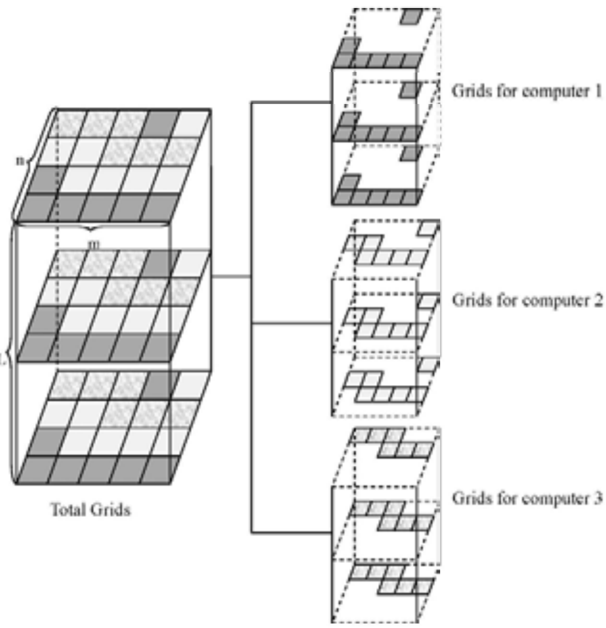


Figure 5.30: Grids based allocation for parallel computation

Assumptions:

I = (m*n) / NUM_c
R = (m*n)% NUM_c

Condition Ⅰ: R=0

The computer A which is allocated grid (i,j,k) can be get with:

$$A = \left\lceil \frac{(j-1)*m+i}{I} \right\rceil$$

Pollution concentration calculated by computer A at (i,j,k) is:

$$C(i,j,k,A) = \sum_{p=1}^{NUM_e} C(i,j,k,p,A) \quad A = \left[\frac{(j-1)*m+i}{I}\right] \qquad \text{......................(5.19)}$$

The final combined pollution result should be:

$$C(i,j,k) = \sum_{A=1}^{NUM_c} C(i,j,k,A) \qquad \text{......................(5.20)}$$

Condition Ⅱ: R≠0

Pollution concentration at (i,j,k) is:

$$C(i,j,k,A) = \sum_{p=1}^{NUM_e} C(i,j,k,p,A) \qquad \text{......................(5.21)}$$

Where

$$\begin{cases} A = \left[\dfrac{(j-1)*m+i}{I}\right] & (((j-1)*m+i) \le (m*n-R)) \\ A = ((j-1)*m+i)-(m*n-R) & (((j-1)*m+i) > (m*n-R)) \end{cases}$$

The final result is combination as:

$$C(i,j,k) = \sum_{A=1}^{NUM_c} C(i,j,k,A) \qquad \text{......................(5.22)}$$

(4) Hybrid parallel arithmetic

This arithmetic is combination of above three kinds of arithmetic, which means to divide grids, layers and point emissions at the same time. According to combinations of layers, point emissions and grids, there are at least four types of hybrids, which are multiple layers and grids based parallel arithmetic, multiple layers and point emissions based arithmetic, multiple point emissions and grids based arithmetic, and multiple layers, point emissions and grids based arithmetic.

- **Hybrid of multiple layers and grids based parallel arithmetic**

This arithmetic regards grids and layers as a whole, which means the allocated number will be L*m*n. The allocation arithmetic is shown in Figure 5.31.

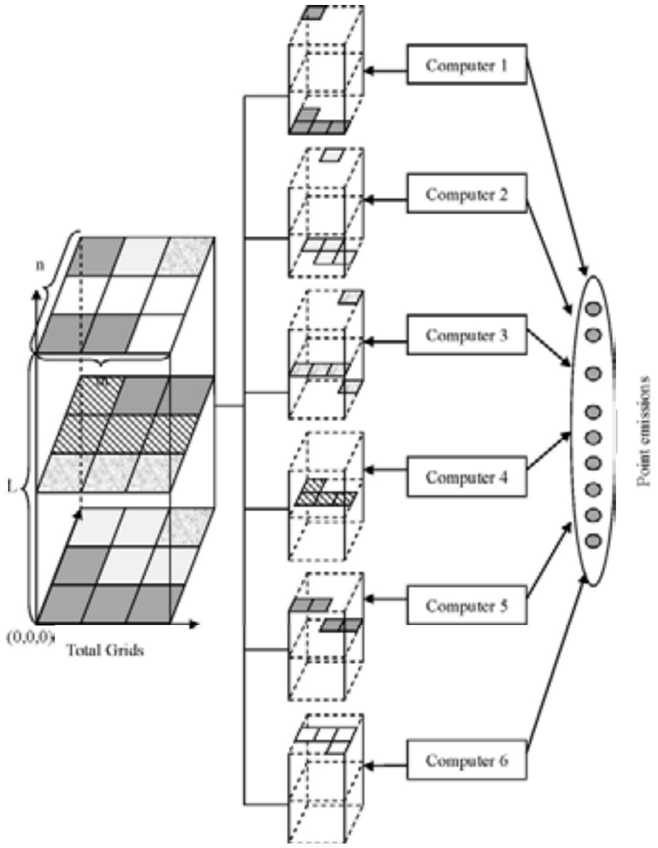


Figure 5.31: Hybrid of multiple layers and grids based parallel arithmetic

Assumptions:

$I = (L*m*n) / NUM_c$
$R = (L*m*n) \% NUM_c$

In this condition, concentration at location (i,j,k) is calculated by computer A based on all point emissions, which is described as below equation.

$$C(i,j,k,A) = \sum_{p=1}^{NUM_e} C(i,j,k,p,A) \qquad \text{......................(5.23)}$$

The whole concentration is:

$$C(i,j,k) = \sum_{A=1}^{NUMc} C(i,j,k,A) \qquad \text{.......................(5.24)}$$

Where

$$A = \begin{cases} \left\lceil \dfrac{(k-1)*m*n+(j-1)*m+i}{I} \right\rceil & (R=0) \\ \left\lceil \dfrac{(k-1)*m*n+(j-1)*m+i}{I} \right\rceil & \left(R \neq 0 \;\; and \;\; \dfrac{(k-1)*m*n+(j-1)*m+i}{NUM_c} \leq I\right) \\ ((k-1)*m*n+(j-1)*m+i) - I*NUM_c & \\ & \left(R \neq 0 \;\; and \;\; \dfrac{(k-1)*m*n+(j-1)*m+i)}{NUM_c} > I\right) \end{cases}$$

- **Hybrid of multiple layers and point emissions based parallel arithmetic**

Each layer is calculated by several computers which form a group. At the same time, emissions are also divided to group computers. The final grids pollution concentrations are addition by all the group computers.

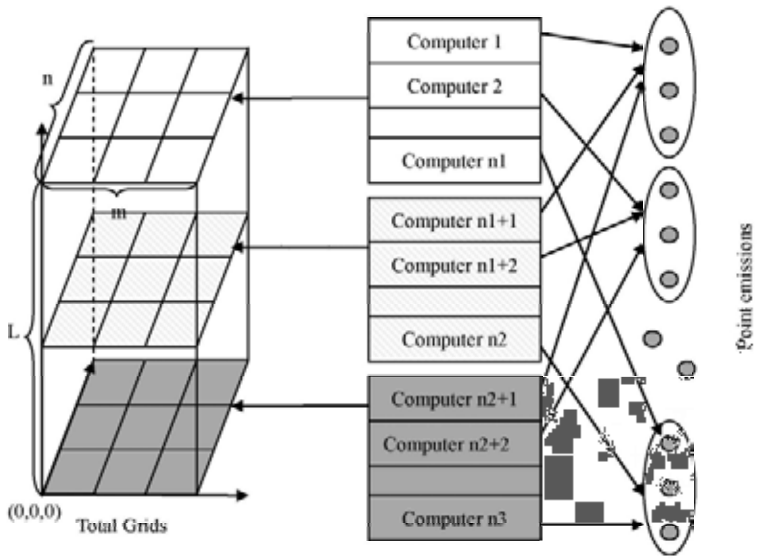


Figure 5.32: Parallel arithmetic of hybrid of multiple layers and point emissions

$$C\ (i,j,k) = \sum_{A=1}^{n_k} C(i,j,k,A) = \sum_{A=1}^{n_k} \ \sum_{p=1}^{p_A} C(i,j,k,A,p) \qquad \text{.......................(5.25)}$$

Where n_k is number of computers in a group to calculate grids pollution concentrations on layer k; p_A is number of point emissions to be calculated by computer A. The arithmetic of getting n_k and p_A can be combined from previous layer allocation and point emission allocation.

- **Hybrid of multiple point emissions and grids based parallel arithmetic**

All layers are divided by the same mode which is like grids based parallel arithmetic. At the same time, point emissions are also divided into groups. The grouped grids and point emissions are calculated by grouped computers to get pollution concentrations on grouped grids which can be shown by Figure 5.33. The arithmetic of grids division, point emissions division and computers division can see previous grids based parallel arithmetic, point emissions based parallel arithmetic and layers based parallel arithmetic respectively.

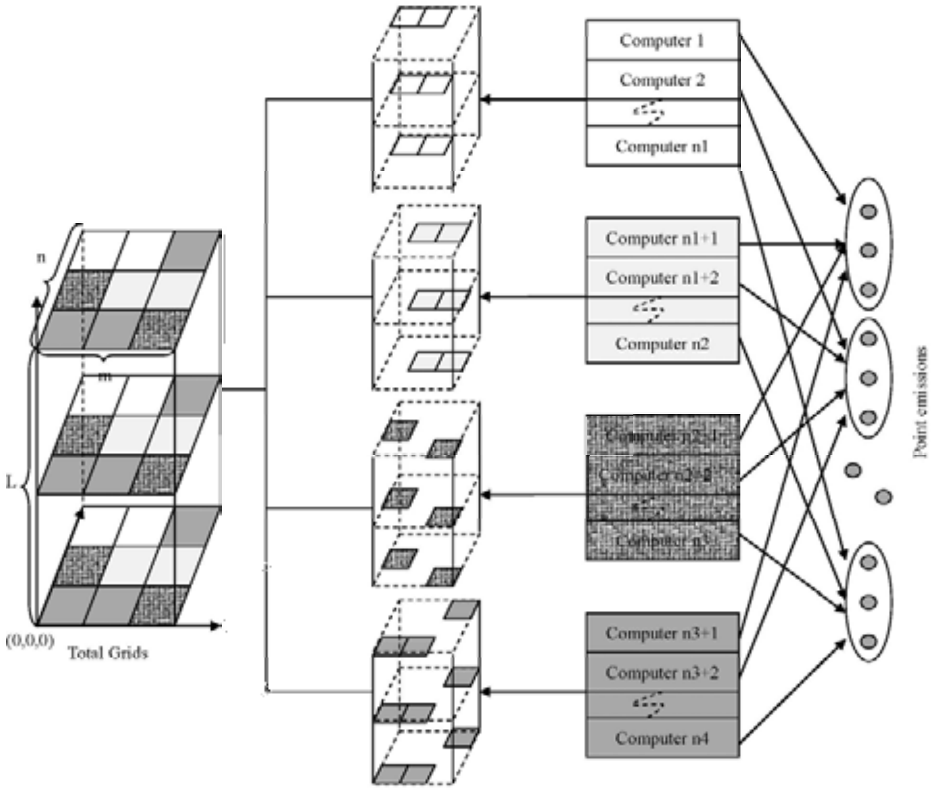


Figure 5.33: Hybrid of multiple point emissions and grids based parallel arithmetic

- **Hybrid of multiple layers, point emissions and grids based parallel arithmetic**

This arithmetic is the most complex but effective one for multi-layer and multi-point-emission style of Gaussian model parallel calculation. For the arithmetic, the grids are divided in both horizontal and vertical directions. At the same time, point emissions are divided to match grouped computers for each grouped grids pollution concentrations calculation. This arithmetic can be combined from previous arithmetic includes multiple point emissions based parallel arithmetic, multiple layers based parallel arithmetic, and hybrid of multiple layers and grids based parallel arithmetic.

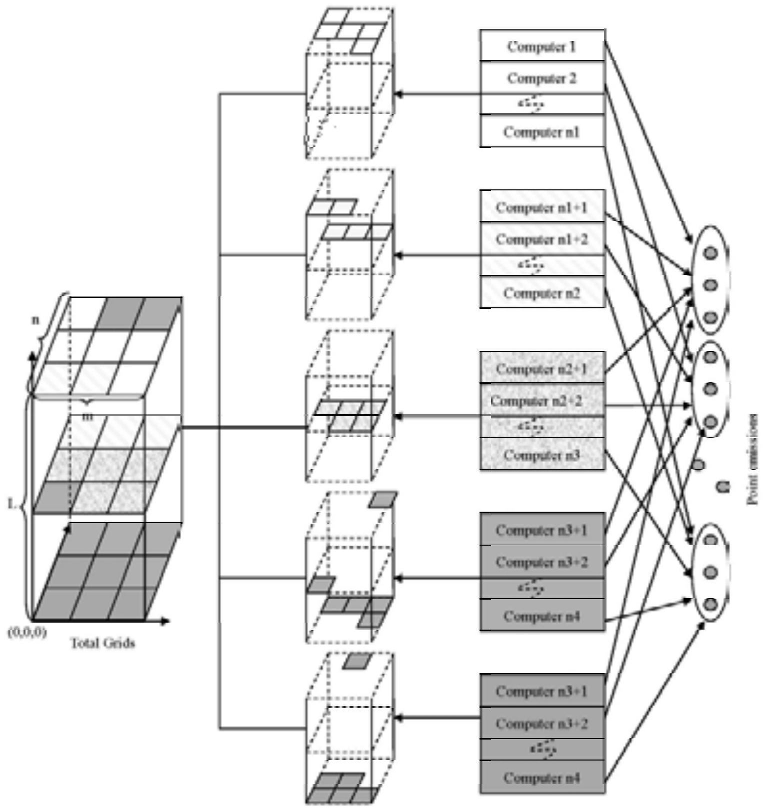


Figure 5.34: Hybrid of multiple layers, point emissions and grids based parallel arithmetic

5.3.2 Arithmetic compare

For above seven arithmetic, computation times decrease with number of computers increasing. The minimum computation times are different with different arithmetic, but none of all the minimum computation times can be less than time which costs one computer to calculate one grid with one emission. The minimum computation time together with other three features are listed in table 4.3 for comparing.

Table 5.3: Seven parallel arithmetic compare

Arithmetic	Minimum computation time	Number of computers to get minimum computation time	Complexity	How to get result by sub-results
Multiple emissions based parallel arithmetic	When number of computers equal that of emissions points	p	low	addition
Multiple layers based parallel arithmetic	When number of computers equal that of layers	L	low	Combination
Multiple grids based parallel arithmetic	When number of computers equal that of Grids	m*n	middle	Combination
Hybrid of multiple layers and grids based parallel arithmetic	When number of computers equal that of m*n*L	m*n*L	middle	Combination
Hybrid of multiple layers and point emissions based parallel arithmetic	When number of computers equal that of number of layer multiplies number of emissions	P*L	middle	Addition and Combination
Hybrid of multiple point emissions and grids based parallel arithmetic	When number of computers equal m*n*p	m*n*p	High	Addition and Combination
Hybrid of multiple layers, point emissions and grids based parallel arithmetic	When number of computers equal m*n*L*p	m*n*L*p	High	Addition and Combination

CHAPTER 6: GEO-VISUALIZATION OF THE OUTPUT OF AIR POLLUTION MODELLING

This chapter presents the geographic information, air pollution model computation results and analysis in 2D and 3D environments. The two environments are both studied because the system attempts to meet the different operating habits in order to achieve geo-collaboration. For instance, when displaying air pollution distribution/dispersion, experts from the atmospheric and environmental sciences prefer a 2D environment, while laymen would prefer using high dimension visualization, such as 3D or even high environments.

2D visualization of air pollution dispersion is not a novel concept. However, there has been integration of geographical visualization with geographic information and air pollution dispersion together. A widely used tool to display air pollution today is GrADS, and yet the interaction in this is very low. In addition, GrADS does not support spatial and temporal analysis. Unlike GrADS, a 2D environment not only inherits the main function of GrADS, but also improves interaction and extends the function of spatial and temporal analysis. The diagram, which is one method of geo-visualization, is also integrated into a 2D environment.

A 3D environment also has obvious merits that compare with the existing high dimension visualization tools that can be used to display air pollution distribution and dispersion, such as Vis5D. The literature review indicated that Vis5D and other high dimension visualization tools for air pollution presentation are all scale fixed; they also do not support information querying. Spatial and temporal analysis is not yet supported. The 3D environment developed by this research has tackled these problems and improved the situation for most of them. Besides these requirements, the toughest work for 3D geographical visualization is visualizing air pollution as a volume object with fuzzy boundary and dynamic behaviours. Generally in a virtual environment, objects are modelled using model creation tools, such as 3D Max, Creator or SketchUp, and are then driven by 3D engines, for instance OpenGL, in order to get the virtual scene. But visualization is different for air pollution mass. An Air pollution mass cannot be created using model creation tools because its shape and density are changing as time passes. Therefore, the pollution mass must be generated during the active running time. Meanwhile, air pollution masses do not have clear boundaries. It is almost impossible to tell where the boundary of a pollution mass is. In fact, a pollution mass boundary is fuzzy. Finally, a pollution mass is a volume body. Viewed externally, a pollution mass looks like a cloud. When moving close to or even into the mass, it seems to turn into fog.

2D and 3D geo-visualizations are based on the output of air pollution dispersion models. However, the output cannot be read directly. Therefore, the conversion and transformation of the output of models' computation must first be studied.

This chapter is organized as follows: the first section transforms the air pollution models computation results as calculated by MM5 and Gaussian plume model into a geographically located format; sections two and three then study geographic visualizations for air pollution distribution/dispersion in 2D and 3D environments respectively.

6.1 Data Transformation

Data transformation is the basis for visualization. In this part, there is a discussion of how to transform the model (MM5 and Gaussian) computation results into a CVGE readable format. The overall data transformation flowchart is shown in Figure 6.1, which includes model result conversion, creating 2D visualization data and creating 3D visualization data.

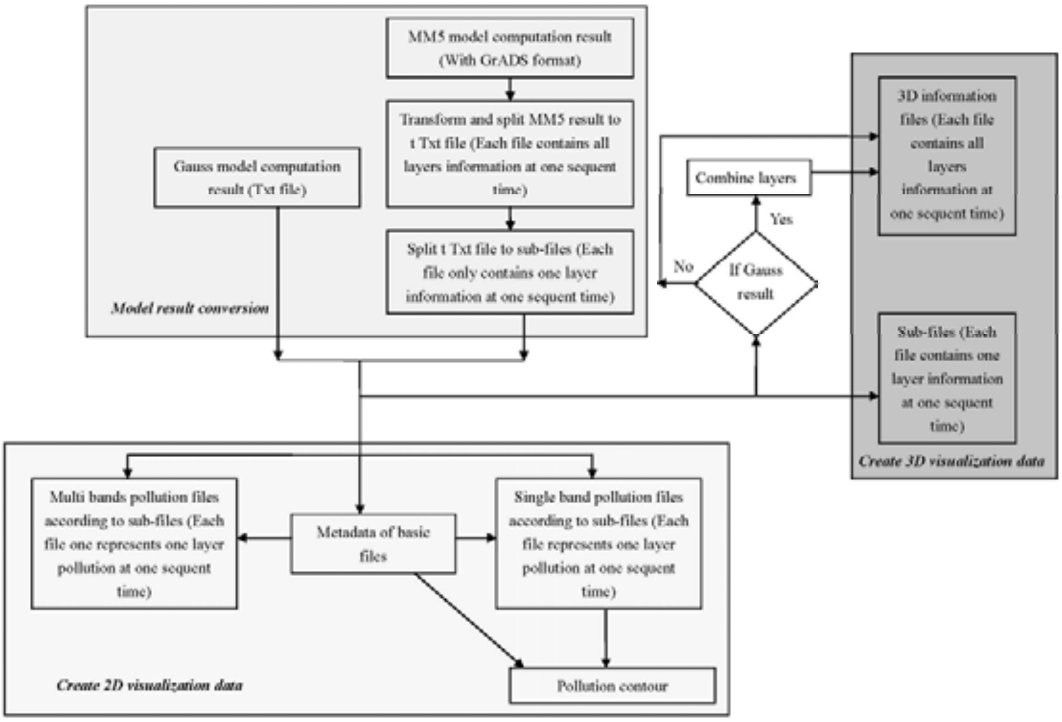


Figure 6.1: Data transformation path

MM5 outputs pollution concentrations on defined sigma height layers and at defined sequential times. The result calculated by MM5 is a combination file containing information on all layers at all sequential times. The output file is binary and can be read and converted using a data conversion tool — MM5ToGrADS (www.mmm.ucar.edu /mm5/ mm5v3/ tutorial/mm5tograds/mm5tograds.html, access date 2009.4.13). After conversion, the result is GrADS data and maintains the combination. The transformation of the MM5 result by this thesis starts from this point. The first transformation step is to transform and split the GrADS data into Txt files according to the sequential times. After this step, the contents of each Txt file are the pollution concentrations of all layers at one point in time. If the total time sequence is t, after the first step the MM5 result will become t Txt files, which are called sequence files. These files will be used to generate the 3D pollutant mass for 3D visualization. The second step is to split the sequence files by layer. The result is that the split files contain only one layer's pollution concentrations at one point in time. These files are called layer and sequence files, and are used to generate pollution planes which can be used in 2D and 3D visualizations.

For the Gaussian plume model, the output has been formatted directly as layer files.

In this research, the 2D platform adopts the popular GIS methodology to represent the pollution data. GIS marks all of the data and information with geographic locations, which serve as an index to organize and manage the geographic information. GIS introduces raster and vector as two popular formats for organizing geo-data. The raster data is spatial data, which is divided into rows and columns in order to form a regular grid structure. This regular grid structure is very similar to the output of MM5. This is why the GIS raster format is used to present air pollution distribution. GIS vector data is helpful in presenting the isoline of pollution concentration. The isoline can easily be retrieved from the raster data using GIS tools.

3D visualization uses two ways to represent pollution distribution and dispersion. The first is to represent the volume pollution mass, which contains pollution concentrations of all layers generated from the MM5 sequence files or from the combined Gaussian layer files. The second way is to represent pollution only at a fixed sigma height layer in order to make plane distribution or dispersion animation.

6.1.1 Transformation of models' computation results

6.1.1.1 MM5 computation result transformation

The MM5 output is a binary file. This file can be read and converted by a tool named MM5ToGrADS. After conversion, the file provides GrADS readable data in two files. One is the content file with the extension name ".dat", and the other is a metadata file with the extension name ".ctl". The contents of the GrADS format file are a sub-set of the original MM5 output, and are determined by setting opening or closing tags in the initiation file for MM5ToGrADS. For instance, if wind vector needs to be converted to GrADS data, the wind vector tag should be set to 1. In this research, five parameters –latitude, longitude, wind speed in a west-east direction, wind speed in a north-south direction and pollution concentration – are set to 1. This means that these parameters will be exported.

MM5 output data, GrADS format data and pollution raster data share grid coordinates, which are presented in Figure 6.2. The grids resolution and the distances between the two neighbouring layers are fixed once the simulation area is initiated before model computation.

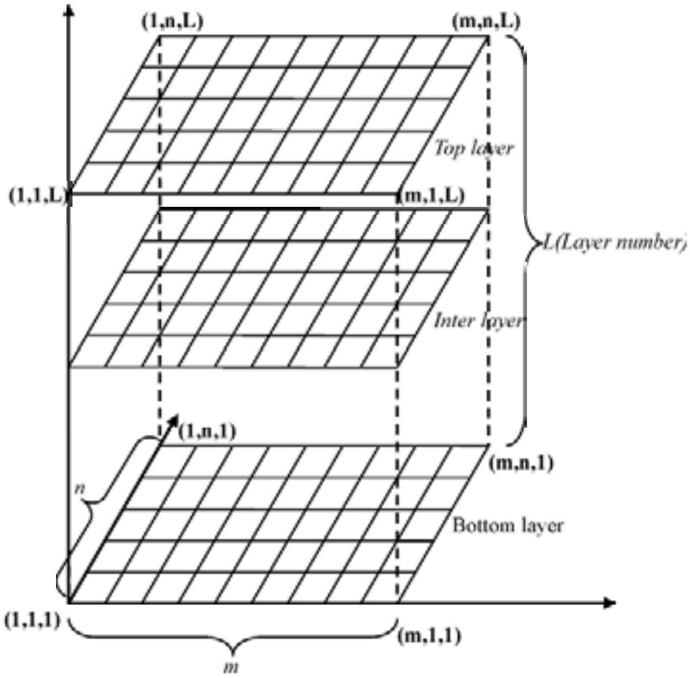


Figure 6.2: Grid coordinates for air pollution result

The data structure of GrADS data is shown in Figure 6.3. Each grid is recorded in one line with five attributes attached: latitude (lat), longitude (lon), wind speed in a west-east direction (u), wind speed in a north-south direction (v) and SO2 concentration (C). The lower layer will be stored first, and then those above, one by one, until all the layers are stored. If there is more than one element of time information, then the file will store the data from time one first, then time two, and so on until the last time.

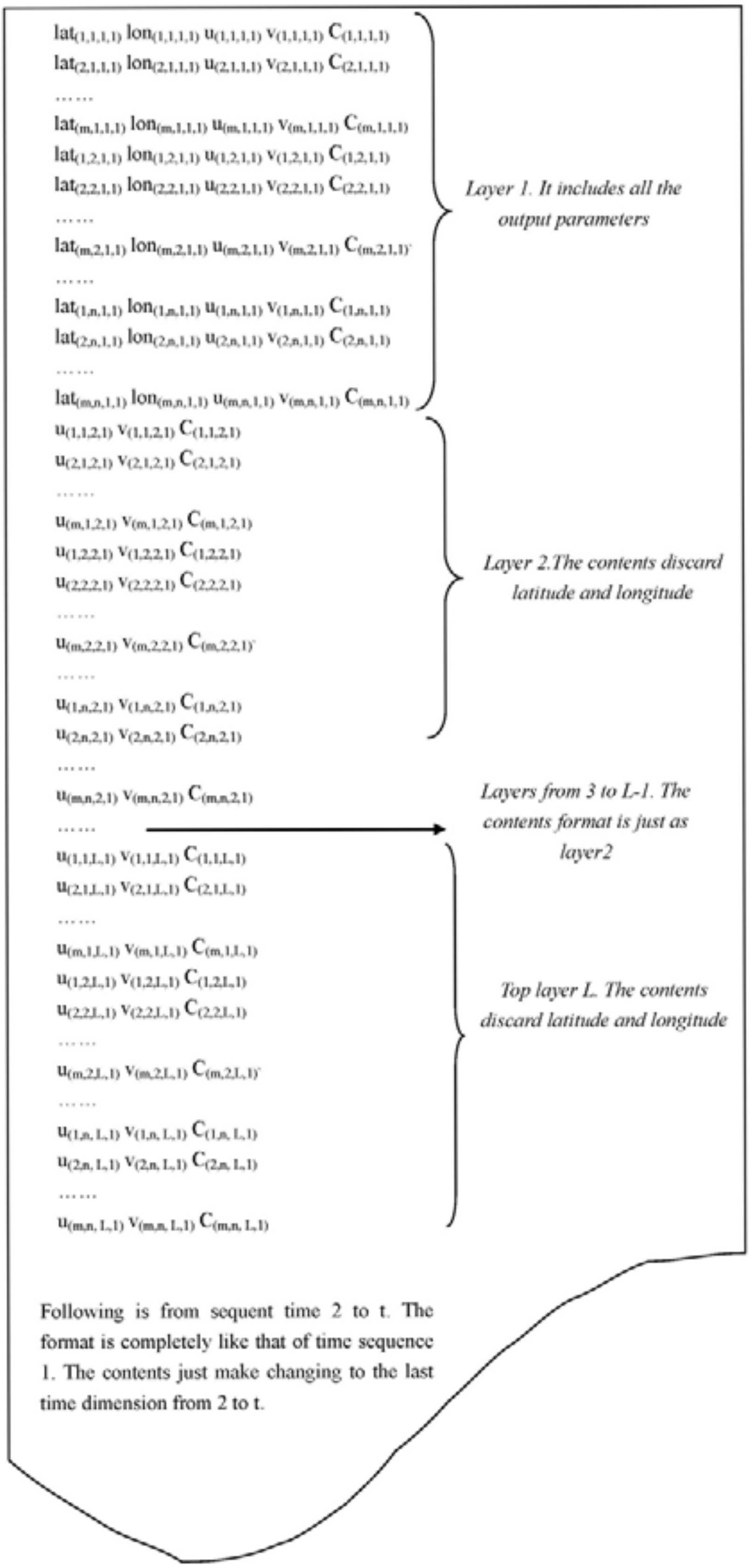


Figure 6.3: Format of GrADS data ((i,j,k,t) means a grid location with time)

GrADS data is a combination of all layers and all times. The next step is to split GrADS data into time based data, called Time Sequence Data (TSD), which means that each section of data contains pollution information on all layers at one time. If there are n times of pollution output, there will be n TSD files. TSD are Txt files, which can be read by the naked eye. A TSD file is filled with line records. All of these lines have the same format, shown in Table 6.1, which includes latitude, longitude, column, row, layer, height, wind speed at X, wind speed at Y and pollution concentration.

If there is one more step is taken to split the TSD files by layer, the Layer and Time Sequence Data (LTSD) will be generated. An LTSD file contains information on only one layer at one time. Therefore, if there are n times and m layers, there will be n*m LTSD files. LTSD files are also filled with line records. All lines in LTSD are formatted as shown in Table 6.2, which only omits the field of layer from Table 6.1.

Table 6.1: Sequence files format

Latitude	Longitude	Column	Layer	Row	Height	Wind speed X	Wind speed Y	pollution concentration

Table 6.2: Layer (and sequence) files format

Latitude	Longitude	Column	Row	Height	Wind speed X	Wind speed Y	pollution concentration

6.1.1.2 Gaussian computation result transformation

The Gaussian plume model computation result does not need any conversion before generating 2D and 3D visualization data. The output is directly formatted as shown in Table 6.2, which is the final format for visualization data generation.

6.1.2 Data generation for 2D visualization

With TSD and LTSD files, raster and vector data can be generated for 2D visualization.

(1) Pollution raster generation

One clear feature of MM5 and Gaussian outputs is that they are regular grids attached with pollution concentration attribute. This feature is very similar with that of GIS raster. So, it is easy and reasonable to apply GIS raster as air pollution format. To generate raster of pollution, below several key steps should be taken.

Step one: Set projection of raster

The output data projection of MM5 and Gaussian plume model is Mercator, and the location is formatted as (latitude, longitude) with unit of decimal degree. If this projection does not match the visualization environments, the output data of air pollution dispersion modeling should be re-projected. For instance, if the research area is PRD region, and the visualization environment is projected at WGS84UTM_49N, the modeling output data should also re-projected to WGS84UTM_49N.

Step two: Create pollution grids

According to location of left and bottom corner of simulated area, the numbers of column and row and grid distance (or grid resolution), an empty raster is created. During creating empty raster, number of raster bands is specified. In this research, the number can be set with 3 and 1 respectively. Three bands raster is a pseudo color raster which is vivid to represent pollution distribution. The single band raster is a gray color raster, which is convenient for information extraction for analysis. Single band raster is also the base to generate air pollution contour.

Step three: Grids value filling with pollution color

After empty grids creating, each grid should be filled with pollution color which is calculated by interpolation between maximum and minimum pollution concentrations.

For three bands raster, the pollution color is composed of three components, which are red, green and blue as C(R,G,B). It is user specified color range. Following example set 510 colors range as pollution color. The values of three components are calculated by below equations.

$$ColorValue = (\text{int})\,(P_x - P_{\min}) \bullet 2 \bullet 255 / (P_{\max} - P_{\min}) \quad \text{……………………} \quad (6.1)$$

Where ColorValue is an integer between 0 and 510; P_x is pollution concentration calculated by air pollution models; P_{max} and P_{min} is maximum and minimum pollution concentrations of layers.

If ColorValur≥255

$$\begin{cases} R = 510 - ColorVale \\ G = 0 \\ B = 0 \end{cases} \quad \text{……………………} \quad (6.2)$$

If ColorValur<255

$$\begin{cases} R = 255 \\ G = 255 - ColorValue \\ B = 255 - ColorValue \end{cases} \quad \text{……………………} \quad (6.3)$$

For single band raster, the pollution color is interpolated between 0 and 255. The equations are:

$$Color = (\text{int})\,(P_x - P_{\min}) \bullet 255 / (P_{\max} - P_{\min}) \quad \text{……………………} \quad (6.4)$$

Above equations for color determination are only examples. Users can define whatever color they prefer.

(2) Pollution contour generation

Based on single band gray raster of pollution, pollution contour can be generated by finding equal neighbor grids color values and drawing line across the centers of these grids. This step can be easily realized by invoking ArcEngine API of Surface Operation Interface.

(3) 2D wind field generation

The output of MM5 content contains information of wind vector with three direction components (x,y,z). For 2D visualization, only two-direction (x,y) components are used. Unlike original wind presentation in Cartesian coordinate system, in 2D visualization, wind filed will be presented into polar Coordinates. Wind field files are GIS vector files, which are generated from LTSD. The steps to generate wind field are listed below.

Step one: Create an empty GIS point vector file;

Step two: Open a LTSD file, read it one line by one line, and extract two components of wind (w_x, w_y);

Step three: Make synthesis of w_x and w_y to get wind plane vector. Extract wind speed and direction from synthesis vector and fill them into wind GIS point file.

Step four: Loop from step one to step three until all LTSD files are processed.

6.1.3 Data generation for 3D visualization

Unlike data for 2D visualization, in 3D environment, all data should be located with not only horizontal coordinates but also vertical coordinates. Pollution data for 3D visualization have two types. One is 3D layer pollution data (3DLPD) to represent pollution distribution on a layer at a time. The other one is 3D time pollution data (3DTPD) to show all layers' pollution at a time.

(1) Layer pollution data generation

Triangulated Irregular Network (TIN) is adopted to represent 3D layer pollution data. With LTSD txt files, 3DLPD files can be generated by below method.

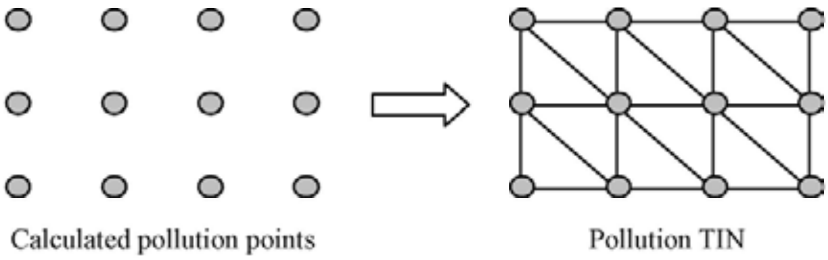


Figure 6.4: Method of pollution TIN generation for one layer at one time

Pollution points calculated by air pollution models are regularly distributed like Figure 6.4(left). The pollution TIN is formed by connecting three neighbor points like Figure 6.4(right). For each triangle, the vertexes colors are set according to their pollution concentrations with equation 6.2 or 6.3. At the same time, vertexes transparences are also set according to pollution concentrations. For example, if there is no polluted at a vertex, the concentration should be 0. Thus the transparence at this vertex will be set with 100%, which means this vertex will not be colored.

The pollution TIN is not required as plane. In fact, each vertex is set with sigma height which is the height above terrain. Thus, the pollution TIN has the same undulation as its below terrain.

After setting three neighbor vertexes' attributes (color and transparence), the triangle can be rendered by invoking OpenGL API.

(2) Volume pollution data generation

Comparing with 2D and 2.5D data, 3D volume data try to fill all occupied space with information. For pollution representation, a volume pollution data (VPD) presents all levels pollution information at a time. VPD does not simple overlay one level by one level, but unifies them into a synthetical volume mass. This thesis tries to use pollution boxes to represent VPD.

Before showing how to represent VPD, some terms should be given first. **A pollution box** is a 3D enwrapping box with the center at the polluted point (Figure 6.5(A)). **Pollution region box** is the whole 3D space which enwraps all pollution points and is composed of a lot of sides by sides **sub-region pollution boxes** (Figure 6.5(B)). **A sub-region pollution box** enwraps a pollution box like Figure 6.5(C).

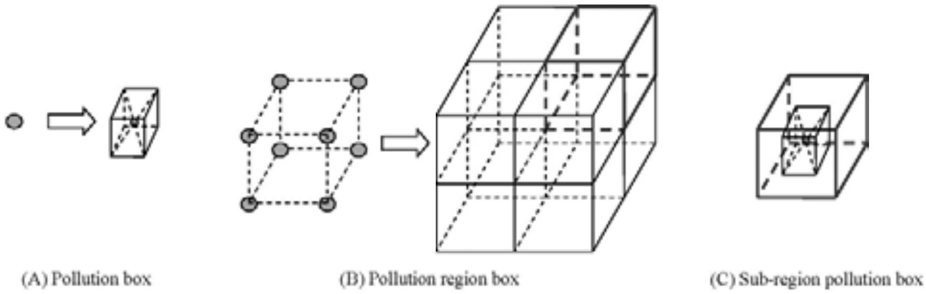


Figure 6.5: Pollution box to form 3D volume of pollution mass

The methodology of creating 3D volume pollution mass data is listed below.

Step one: According to TSD Txt files, make an equal distance interpolation with all layer.

Step two: With interpolated layers, enwrapping pollution points with sub-region boxes and getting pollution region box.

Step three: In each sub-region box, drawing a pollution box, whose center is the interpolated pollution point. The box side lengths are random, but can not go out of the outer sub-region box.

Step four: Fill the pollution box with particles or paste pollution box sides with textures.

For three kinds of boxes, only the pollution box is rendered. The other two boxes, sub-region box and region box, are virtual box and not rendered at all.

6.1.4 Data transformation implementation

The data transformation for 2D environment is implemented as Figure 6.6 which supports MM5 and Gaussian plume model outputs transformations. For MM5 result, the transformation includes splitting time sequence txt files into layer files, creating raster data and vector data. For Gaussian plume model, because the computation result is already split, the transformations only take place on generating raster data and vector data.

Figure 6.6: Data transformation interface

In Figure 6.6, two grouped radio boxes (MM5 and Gaussian) are used to determine other buttons action. Functions of other buttons of Figure 6.6 are listed in table 6.3.

Table 6.3: Function of data transformation button

Item	Function
TransMM5 to GrADS	Read MM5 output data, transform it into GrADS readable file, and split it into time sequent files.
MM5Result PreProcess	To retrieve MM5 result information, such as number of domain, column number, row number, layer number, cell distance, and so on. It is only worked for MM5.
Pollution separate	To split sequence files into layer sequence files. It is only worked for MM5.
FileInfor extraction	Get the layer sequence files information, such as file name, the minimum and maximum pollution concentrations, and so on. It is only worked for MM5.
GaussianResult Process	Retrieve and write Gaussian result information, which includes layer number, model resolution, column number, row number. It is only enabled for Gaussian data processing.
Small region	It is used for pollution color stretch. It is only enable for Gaussian. If calculated area is no larger than 5km*5km, check this item.
MultiBand raster	To generate multiple bands raster from layer sequence data. It is worked for both MM5 and Gaussian.
SingleBand raster	To generate single bands raster from layer sequence data. It is worked for both MM5 and Gaussian.
Isoline	To generate pollution contour vector from layer sequence data. It is worked for both MM5 and Gaussian.
WindField	Generate GIS vector files for MM5 wind field.
Standardize	If checked, the pollution color is interpolated between total files maximum and minimum pollution concentrations. It is worked for both MM5 and Gaussian.
Stretch	If checked, each file's pollution color is interpolated between

	only its maximum and minimum pollution concentrations. It is worked for both MM5 and Gaussian.
Information	Display processing information. It is worked for both MM5 and Gaussian.
Source file	The source file of current processing. It is worked for both MM5 and Gaussian.
Destination file	The destination file of current processing. It is worked for both MM5 and Gaussian.
CurrentFile progress	The progress bar of progress state of current progress file. It is worked for both MM5 and Gaussian.
TotalFile progress	The progress bar of progress state of total file progress file. It is worked for both MM5 and Gaussian.

6.2 2D Visualization

6.2.1 Architecture

The architecture is shown in Figure 6.7. This not only visualizes air pollution distribution and dispersion, but also supplies visual settings and operations of the air pollution models (MM5 and Gaussian). A data transformation function is also integrated, with the data for visualization being GIS raster and vector. The 2D visualization only implements the relevant functions for the atmosphere experts, which include pollution overlay and animation on a background of geographic information. 2D analysis is based upon 2D visualization, and this is point based vertical profile analysis.

A 2D environment supports almost all of the steps of air pollution simulation, including pollution source compiling, model setting, model computation, data transformation, 2D layer visualization and 2D analysis. However, this work flow is different from the way it is generally conducted by scientists from the atmospheric and environmental sciences in that: ① the model is initiated using geo-data and is directly input through a visualized interface; ② model computation is based on high performance computation; ③ the visualization is geographically referenced; ④ the interface is highly interactive; and ⑤ the analysis includes spatial and temporal analysis.

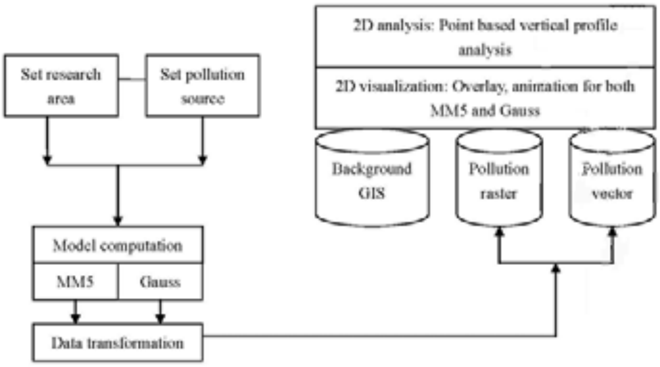


Figure 6.7: Architecture of a 2D visualization environment

6.2.2 Implementation

2D implementation is based on study area of Pearl River Delta (PRD).

The main interface of 2D visualization is displayed as Figure 6.8. On the top is menu including system controlling, pollution source edition, simulation area setting, model computation launcher, model output data transformation, pollution visualization, analysis visualization, and so on. The toolbar under menu is used for operation on visualized layers. Toolbar has many items such as room in/out, map pan, information query, measure, layer editor and so on. The other part of 2D visualization platform is layer operation including control part on the left and display part on the right. The bottom most is status bar.

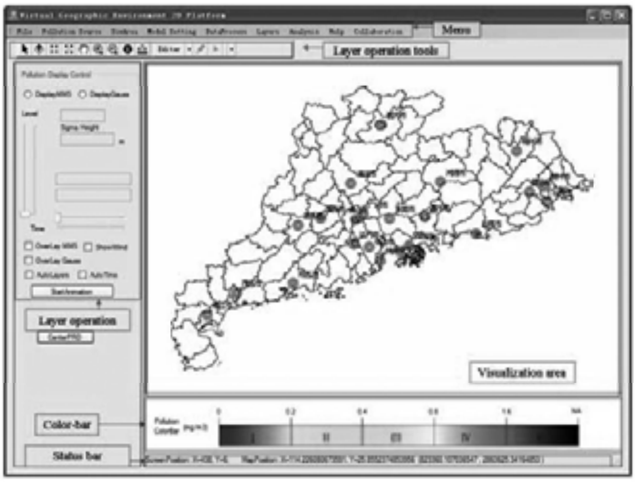


Figure 6.8: Interface of 2D visualization

(1) Pollution source visualized compiling

This module includes functions of pollution source display, source edit (source reset, add source, and delete source), select pollution source, and show closed pollution source. When click “Display pollution source” item, the pollution source will be loaded and overlaid on background map. “Add Source” item in Source Edit group is used to add a new pollution source at clicked point on control map. The “Delete Source” is used to delete a pollution source at the clicked point. If resetting pollution source as the original state, just click “Source Reset” item. The item of “Select Pollution Source” tries to open or close pollution sources like Figure 6.10 (A). In Figure 6.10(A), pollution sources are grouped by county administration boundary. If a county is checked, the pollution in this county will be opened for air pollution model computation. Otherwise, the source will be closed and has no impacts on air pollution. Pollution source open and close states will influence pollution display on map. If a pollution source is closed, it will not be shown or marked as another symbol on map. The Figure 6.10 shows how to set pollution source states (Open/Close) and how to visualize them.

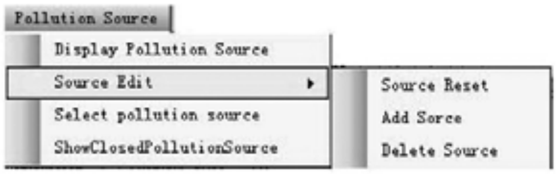


Figure 6.9: Pollution Source module

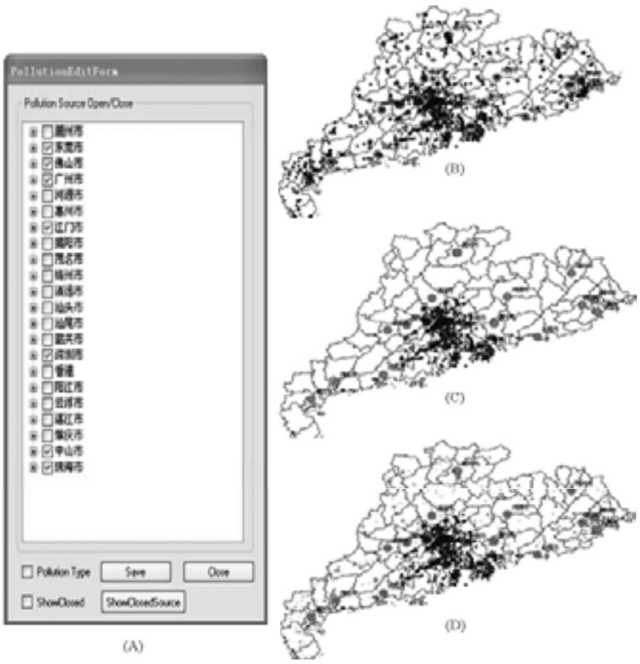


Figure 6.10: Pollution source selection and visualization ((A) is to set pollution source state. (B) All pollution sources distribution in PRD. (C) Only opened source is displayed according to (A) set. (C) Opened pollution sources are displayed in bright color and symbol, while closed source shown with dark symbol.)

(2) Geographic information visualization

Geographic information is integrated as background. The background layers include administrations' points, administration boundary, road, railway, lakes, grass lands, and buildings, which are listed in table 6.4. All the background layers and pollution layers are re-projected into same project (WGS84_UTM_N49). The basic operations on 2D visualization are implemented with embedded modules such as add data, room in/out, pan, query information, measure, and map edition which are components of ArcGIS which have been implemented by Figure 6.11.

Figure 6.11: Basic operation functions on 2D visualization

Table 6.4: Integrated geographic information

Layer name	Type	Content	File number
Village	Point vector	Villages' location and name in Guangdong province.	1
Township	Point vector	Towns' location and name in Guangdong province.	1
County	Point vector	Counties' location and name in Guangdong province.	1
City	Point vector	Cities' location and name in Guangdong province.	1
Guangdong boundary	Polyline vector	County polyline of Guangdong province.	1
Road	Polyline vector	Multiple level roads in Guangdong province.	8
Railway	Polyline vector	Main railway in Guangdong province.	1
Lake	Polygon vector	Lakes in Guangdong province.	1
Grass	Polygon vector	Grass lands in Guangdong province.	1
HangKong boundary	Polyline vector	Districts boundary of Hong Kong	1
Building	Polygon vector	All building in Hong Kong	1

(3) 2D pollution distribution visualization

After simulation area selection and configuration and air pollution models (MM5 and Gaussian plume model) computations, which are given in previous chapter, data transformation will be launched to generate 2D visualization raster and vector. Raster is for regular grids and vector for pollution concentration contours. When doing visualization, the two kinds of pollution results (by MM5 Gaussian plume model) are processed in the same way.

Below Figures (Figure 6.12 and Figure 6.13) show air pollution distributions at a time point calculated by MM5 and Gaussian plume model respectively. In Figure 6.12, pollution raster generated from MM5 computation result is overlaid on GIS data. In Figure 6.13, Gaussian plume model calculated pollution raster and vector are all displayed on the background GIS data.

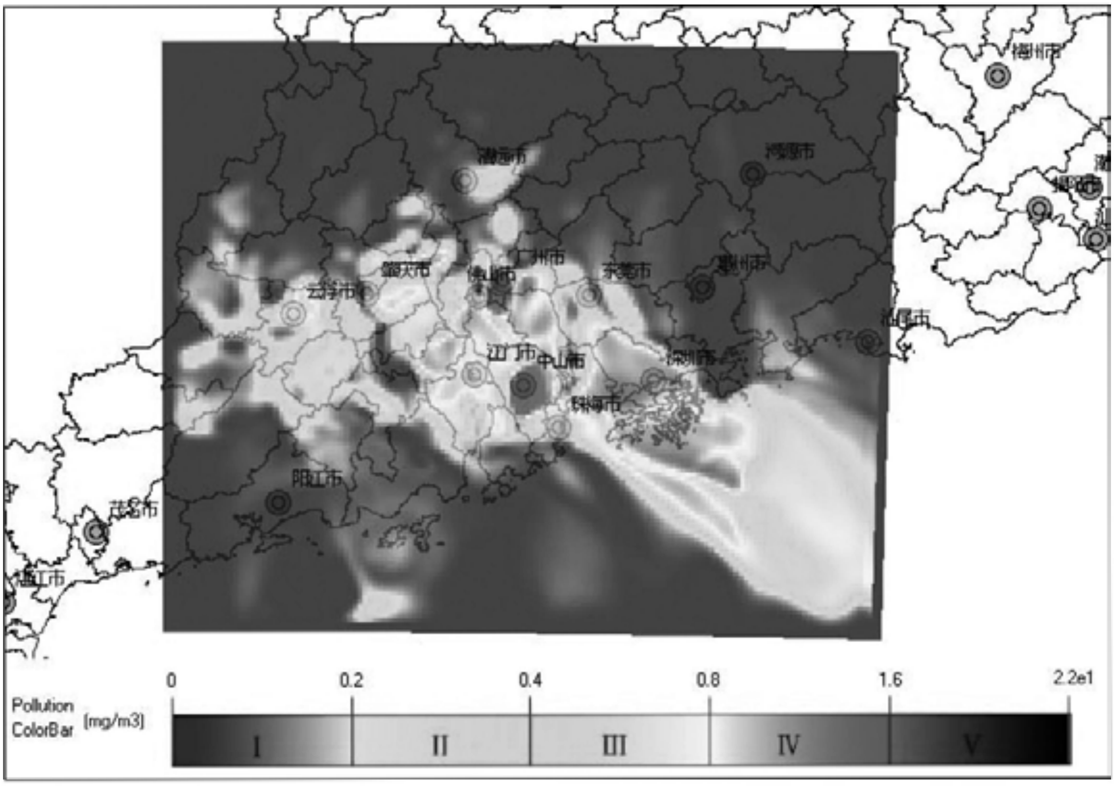


Figure 6.12: MM5 based air pollution distribution visualization

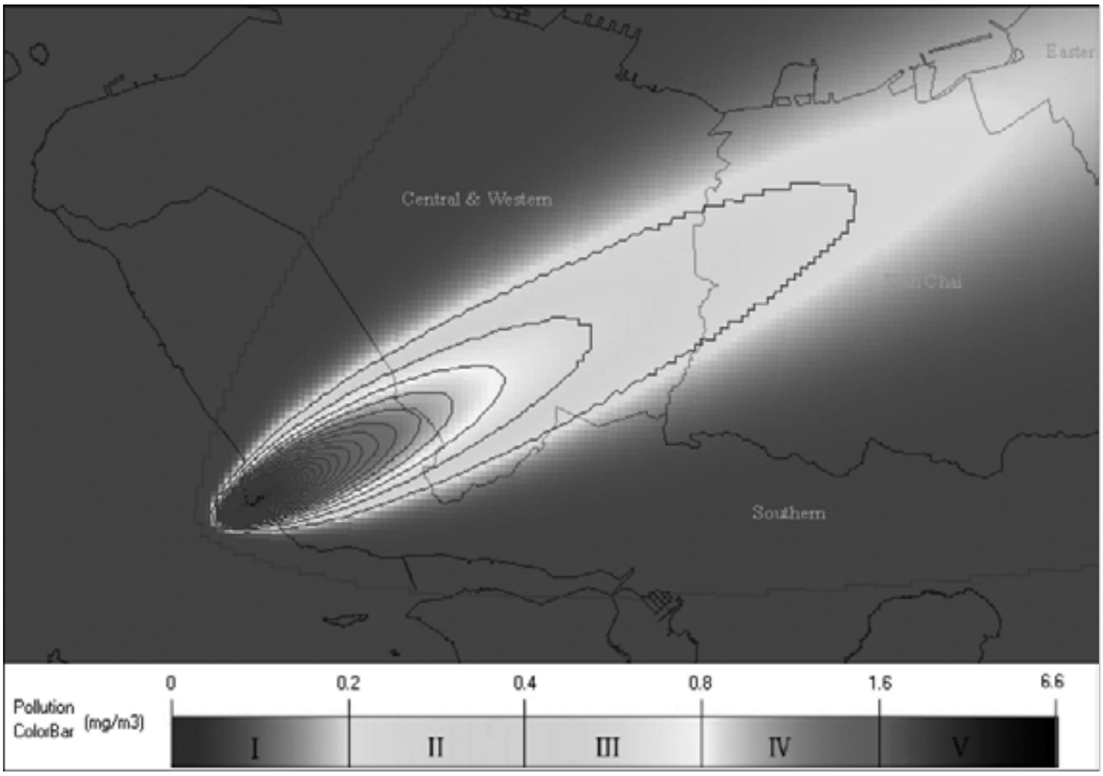


Figure 6.13: Gaussian plume model based pollution visualization

(4) Pollution layers control

Pollution layers are controlled by interface of Figure 6.14. It is used to control air pollution distribution/dispersion calculated by both MM5 and Gaussian plume model. The functions of items on layer control interface are listed in table 6.5.

Figure 6.14: Pollution layer controlling interface

Table 6.5: Items' functions of pollution layers control

Item	Type	Function	Other
DisplayMM5	Radio button	Checked to make two TrackBars only effect on MM5 result layers.	
DisplayGaussian	Radio button	Checked to make two TrackBars only effect on Gaussian result layers.	
LayerNumber	Text	Show number of the layer which determined by layerTrackBar	
layerHeight	Text	Show height of a layer.	
LayerTrackBar	TrackBar	Drag to set display layer count.	
TimeTrackBar	TrackBar	Drag to set display layer time.	Only enabled for MM5
OverlayMM5	CheckBox	Checked to shown MM5 result layer set by LevelTrackBar and TimeTrackBar	
ShowWind	CheckBox	Checked to display wind field calculated by MM5. The wind layer is determined by both LevelTrackBar and TimeTrackBar.	
AutoTime	CheckBox	Make time animation with level fixed.	
OverlayGaussian	CheckBox	Checked to shown Gaussian result layer set by LevelTrackBar and TimeTrackBar	
Autolayer	CheckBox	Make layer animation with time fixed.	

The layer control module can be used to, on one hand, only display an air pollution layer at a time to get the stationary presentation, on the other hand, animate air pollution dispersion. Setting layer TrackBar to a step, checking AutoTime CheckBox can make a time based animation, which means pollution dispersion at a layer set by LayerTrackBar. If fixing time TrackBar, checking AutoLayer, the layer based animation can be made, which shows all layers' pollution distribution at a set time.

Layer control module can also be used to display wind field for MM5 result.

(5) Wind filed visualization

Wind fields closely influence air pollution dispersion. Thus, wind fields always adopted by atmospheric and environmental scientists to analysis air pollution dispersion.

According to previous data transformation, the generated vector files of wind fields have two new attributes. One is wind speed(S), and the other one is wind direction (θ).

When loading wind vector point files for visualization, wind points are marked with arrows. Arrows' lengths are corresponding to wind speed, and angles are equal to wind directions. Wind field visualization in 2D is shown as Figure 6.15.

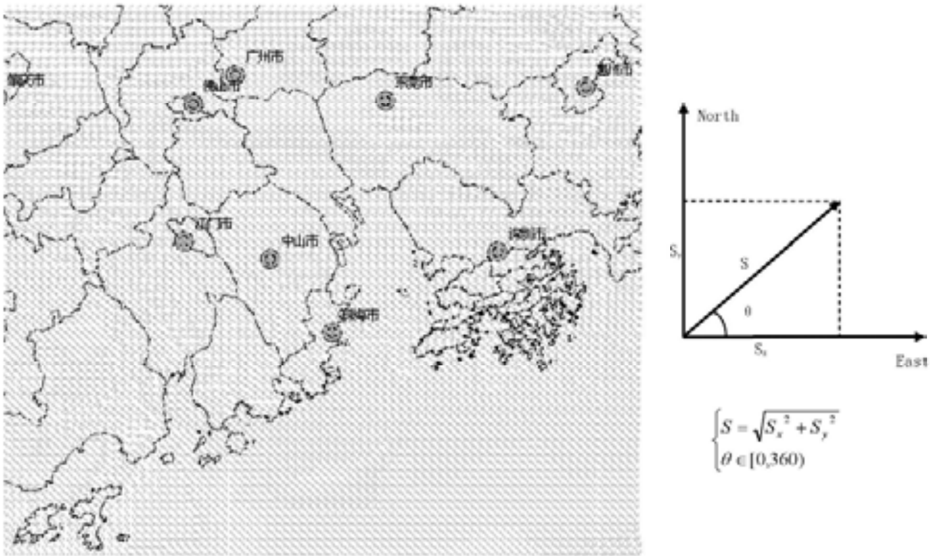


Figure 6.15: Wind filed visualization

(6) Diagram of point concentration profile

This diagram visualization is used to support spatial analysis.

On 2D map, click mouse button to get the analysis point. With the clicked point location (Latitude, Longitude), pollution concentrations on this point are queried. The concentrations cover all layers and all time sequences at the click point. Figure 6.16 shows a vertical profile at a time. Other time profile can be gotten by changing TrackBar value.

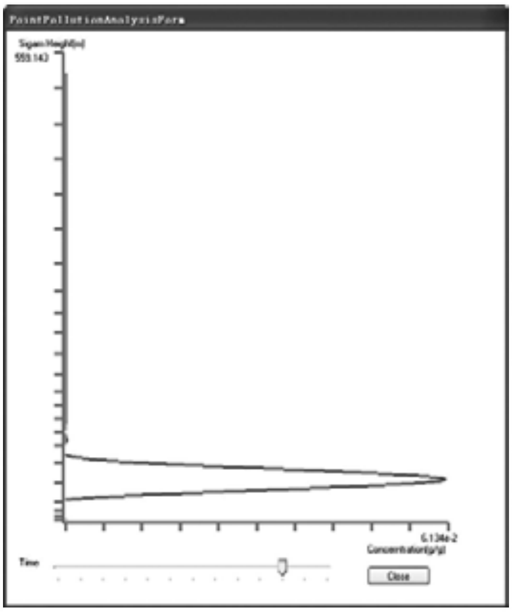


Figure 6.16: Diagram of pollution concentration profile at a clicked point

6.3 Geo-visualization of a Fuzzy Boundary Volume Air Pollution Mass in a 3D Environment

The toughest work for 3D visualization in CVGE is visualizing an air pollution mass in a natural way. In order to emphasize this key aspect, this section will mainly focus on air pollution mass visualization. Although they are also visualized in a 3D environment, geo-data and geo-information are both easier, and thus receive less attention.

6.3.1 Dynamic fuzzy boundary volume object

In terms of visualization, objects in virtual environments can be classified according to their features as being, for example, stationary/dynamic, clear/fuzzy boundary, enwrapped by surfaces or volume.

Clear boundary object —— An object with clear boundary.

Fuzzy boundary object —— An object whose boundary is fuzzy.

Stationary object —— An object that does not change its shape, contents and location.

Dynamic object —— An object that can change its shape, contents or location.

Surface object —— An object that is represented with surfaces.

Volume object —— An object that is full of visualized information in its occupied space.

Objects with clear boundaries that are enwrapped by surfaces have been widely studied, whether they are steady or dynamic. However, large-scale fuzzy boundary volume objects have received little attention.

In this research, the air pollution mass has three complex features that are hard to deal with: it has a fuzzy boundary, volume and is dynamic. In the physical world, the air pollution mass looks like a cloud with a boundary from a distance. But close to, the mass boundary becomes blurred, which is why the mass is called a fuzzy boundary object. At the same time, it is possible to go inside the pollution mass, which makes the surroundings look like fog. Unlike surface enwrapped objects, the space which is filled with a fogged mass is all occupied with pollution information, and it is therefore called a volume object. Finally, a pollution mass is not steady. It changes its shape and location and it is driven by the wind. To summarize, a pollution mass is a dynamic fuzzy boundary volume object (DFBVO).

The method of presenting a pollution mass differs from current means used to visualize a clear boundary object. An air pollution mass cannot be created before visualization using model creation tools. It must be generated during the running time, which is called dynamic modelling. A particle system is one suitable technology that can be implemented for an air pollution mass. A particle system has the merit of shape flexibility, but does consume computer resources. Particle system based air pollution is used for study in this research.

6.3.2 Particle system based air pollution mass modeling

6.3.2.1 Particle system introduction (Encyclopedia 2009)

Particle system is a computer graphics technique to simulate certain fuzzy phenomena, which are otherwise very hard to reproduce with conventional rendering techniques. Examples of such phenomena which are commonly replicated using particle systems include fire, explosions, smoke, flowing water, sparks, falling leaves, clouds, fog, snow, dust, meteor tails, hair, fur, grass, or abstract visual effects like glowing trails, magic spells, etc.

Typically a particle system's position and motion are controlled by an emitter, which acts as the source of the particles. Particle system's location in 3D space determines where they are generated and whence they proceed. A regular 3D mesh object, such as a cube or a plane, can be used as an emitter. The emitter has attached to it a set of particle behavior parameters. These parameters can include the spawning rate (how many particles are generated per unit of time), the particles' initial velocity vector (the direction they are emitted upon creation), particle lifetime (the length of time each individual particle exists before disappearing), particle color, and many more.

It is common for all or most of these parameters to be "fuzzy" — instead of a precise numeric value, the artist specifies a central value and the degree of randomness allowable on either side of the center (i.e. the average particle's lifetime might be 50 frames ±20%)..

A typical particle system's update loop (which is performed for each frame of animation) can be separated into two distinct stages, the parameter update/simulation stage and the rendering stage. During the simulation stage, the number of new particles that must be created is calculated based on spawning rates and the interval between updates, and each of them is spawned in a specific position in 3D space based on the emitter's position and the spawning area specified. Each of the particle's parameters (i.e. velocity, color, etc.) is initialized according to the emitter's parameters. At each update, all existing particles are checked to see if they have exceeded their lifetime, in which case they are removed from the simulation. Otherwise, the particles' position and other characteristics are advanced based on some sort of physical simulation, which can be as simple as translating their current position, or as complicated as performing physically-accurate trajectory calculations which take into account external forces (gravity, friction, wind, etc.). After the update is complete, each particle is rendered, usually in the form of a textured billboarded quad (i.e. a quadrilateral that is always facing the viewer). However, this is not necessary; a particle may be rendered as a single pixel in small resolution/limited processing power environments. Particles can be rendered as Metaballs in off-line rendering; isosurfaces computed from particle-metaballs make quite convincing liquids. Finally, 3D mesh objects can "stand in" for the particles — a snowstorm might consist of a single 3D snowflake mesh being duplicated and rotated to match the positions of thousands or millions of particles.

6.3.2.2 Particles design for air pollution mass

Before the generation of pollution particles, the location, shape, color, dynamics and lifetime of the particles should be determined.

It was shown in chapter 5 that the output of MM5 calculation is divided into multiple layers, and all of these layers are regularly plotted on a grid. For each grid, the centre is the point at which air pollution concentration is calculated with MM5 or the Gaussian plume model. Using these layers on the grid, virtual cubes are formed, as shown in Figure 6.17 (A) by linking neighbouring vertexes. In the original output calculated by MM5 and the Gaussian plume model, the distances between neighbouring layers are not equal. The original cubes thus have different volumes. In order to simplify the process of locating particle emitters, pollution layers are relocated with equal distances using interpolation. The pollution concentrations on the new layers are interpolated from the two neighbouring layers above and below (Figure 6.17(B)). With the relocated layers, the cubes are equal in volume. These cubes are then used to determine the location of emitters and the boundaries of particles moving through space. For each cube, the particle emitter is located at the centre (see Figure 6.17(C)). The following steps are taken to model particles.

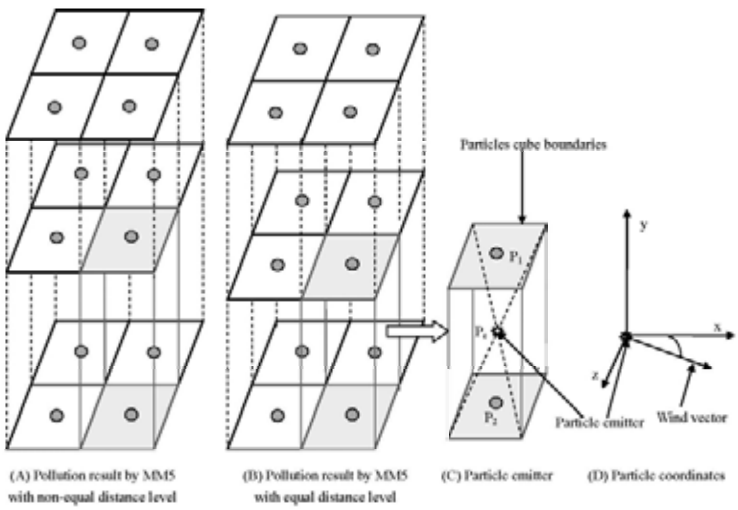


Figure 6.17: Pollution cubes and particle emitter ((A) The output layers of MM5 with non-equal distance between two neighbor layers. (B) The layers are relocated with equal distance between neighbor layers. (C) Particle emitter is located at the centre of virtual cube. (D) Coordinates for particles)

Step one: Emitters are attached using pollution concentrations, which are interpolated according the two neighbouring points above and below. For instance, in Figure 6.17(C), an emitter is attached with a pollution concentration of P_e. The value of $P_e=(P_1+P_2)/2$.

Step two: Over a period of time, the emitter will eject a number of particles. Each particle has the attribute of a pollution concentration of P_e.

Step three: A white color and various levels of transparency are set for all ejected particles according to the pollution concentration attached to the emitter. If the pollution concentration is 0, the transparency is set to 1. If the pollution concentration is at the maximum, the transparency is set to 0. If the pollution concentration is between 1 and 0, the transparency will be set to an interpolated value between 0 and 1.

Step four: The random size of the particles is set. This is one method that can be used to obtain a fuzzy effect of the view. Particles can be spheres, cubes or other small objects. The sizes of the particles are diameters of spheres or length of sides of cubes. The arithmetic for random size setting is presented in equation (6.5)

$$size = avSize + Rand() \times sDisturb \qquad \text{.........................(6.5)}$$

Where:

Size = Particle size;

avSize = Average size of particles;

Rand() = Random value range from 0 to 1;

sDisturb = Maximum of disturbance, which is restricted in a pollution cube.

Step five: Particle behaviours are set. Particles are not stationary, but move here and there. Thus, this step determines the velocities and the locations of particles.

$$\begin{cases} velocity.x = windSpeed \times \cos(windAngle) + Rand() \times vDisturb.x \\ velocity.y = ((mg - \rho_A V) \div m) \times t + Rand() \times vDisturb.y \\ velocity.z = windSpeed \times \sin(windAngle) + Rand() \times vDisturb.z \end{cases} \qquad \text{..........(6.6)}$$

Where:

Velocity.x = Particle velocity component in the x direction;

Velocity.y = Particle velocity component in the y direction;

Velocity.z = Particle velocity component in the z direction;

windSpeed = Wind speed at emitter;

windAngle = The angle between x direction and the wind vector;

vDisturb = Maximum disturbance of particle velocity;

m = Particle mass;

g = Acceleration of gravity;

ρ_A = Density of air;

V = Volume of particle;

T = Lifetime of particle;

The particle position can be calculated using the following group of equations.

$$\begin{cases} P_x(t+1) = P_x(t) + (velocity.x + Rand() \times vDisturd.x) \times \Delta t \\ P_y(t+1) = P_y(t) + (velocity.y + Rand() \times vDisturd.y) \times \Delta t \\ P_z(t+1) = P_z(t) + (velocity.z + Rand() \times vDisturd.z) \times \Delta t \end{cases} \quad \text{………….(6.7)}$$

Where:

$P_x(t+1)$ = Component x of particle position at time t+1;

$P_y(t+1)$ = Component y of particle position at time t+1;

$P_z(t+1)$ = Component z of particle position at time t+1;

$\triangle t$ = Time step for computation.

In order to make particles more random, their moving directions are also given disturbances.

$$\begin{cases} newAngle.x = oldAngle.x + Rand() \times rDisturb.x \\ newAngle.y = oldAngle.y + Rand() \times rDisturb.y \\ newAngle.z = oldAngle.z + Rand() \times rDisturb.z \end{cases} \quad \text{.................(6.8)}$$

6.2.3.3 Air pollution particles optimization

A particle system is good for simulating fuzzy objects. However, it does have a major problem, which is the consumption of computer resources, especially if is makes a precision simulation of a large-scale area. It is evident that there is more precision when more particles are used for simulation, yet this requires more computer resources which will result in rendering stagnancy. If computer resources are limited, the only thing that can improve particle system efficiency is optimization of particle generation, killing and rendering.

For example, in order to visualize air pollution over the whole PRD region requires vertical cover from the terrain surface to 1000 meters. This would involve the generation of a huge number of particles. At the same time, particles should be added into a CVGE, which renders virtual terrain, virtual weathers (rain, snow, lighting, cloud) and so on at the same time. The combination of the virtual geographic environment and the particles increases the consumption of computation resources and prevents more fluent rendering. To overcome this problem, three steps are taken to decrease particles as much as possible from particles emitting through rendering to killing.

(1) View frustum culling based optimization

In a virtual environment, a virtual camera is the vision channel for presenting the virtual scene onto a screen, like eyes in the real world. The virtual camera's position and view angle determine the screen graphic. The view frustum (Figure 6.18(A)), which uses a perspective project and takes the shape of a truncated pyramid, is the volume that holds visible objects. Only those objects that fall into the view frustum will be presented on the screen. For example, in Figure 6.18 (B), all of the green elements (which are totally inside the view frustum) and all the yellow elements (which are partially inside) would be rendered, whereas the red elements would not be rendered.

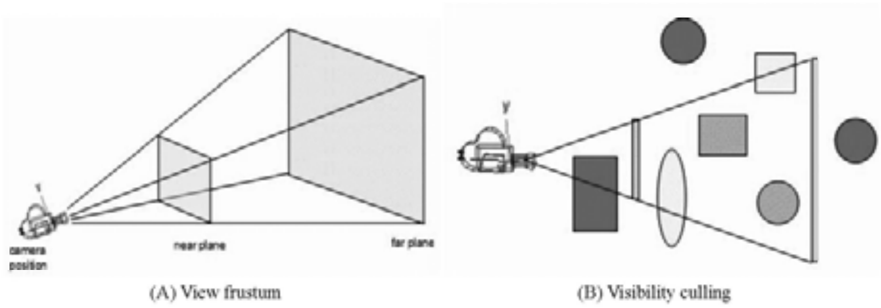


Figure 6.18: View frustum culling

The view frustum can be used to decrease the number of particles in the following three ways:

If a particle is out of the view frustum, it will be deleted from the scene;

If a particle is sheltered, it will be deleted from the scene;

A cube that is far from the viewpoint emits a small number of particles, or may not even emit any particles. A cube near the viewpoint will certainly emit more particles than a more distant one.

(2) Particle rendering

The shape of particles can be not only a small volume object, like a sphere or a cube, but also a billboard pasted with textures. A cube or a sphere need more time for rendering than a billboard because they have more surfaces. Therefore, to save rendering time, a billboard pasted with textures may be a better method for rendering particles.

(3) Particle death optimization

Particles have their own life span. If a particle gets to the end of its life span, it will die and should be deleted from the virtual environments. The dead particles do not consume computer resources for rendering.

There are three conditions that can be used to check if a particle is dead. The first is to check whether the particle has fallen out of the view frustum. The second is to check whether the particle perspective size on the screen is smaller than pixel one. The last is to check whether the particle is sheltered. If any one of these conditions is satisfied, the particle will be set as dead and deleted from the scene.

6.3.3 3D implementation

Pearl River Delta is selected as the study area.

3D environment is implemented based on OpenSceneGraph(OSG), which is open source, with C++ language under platform of Visual Studio.Net 2003. There are several key modules in 3D environment, which are multiple scale terrain data integration and visualization, typical weather visualization, air pollution mass visualization, and 3D analysis visualization. The collaboration, which is saved for next chapter, is also implemented and integrated in 3D platform. The overall interface of 3D environment is shown in Figure 6. 19.

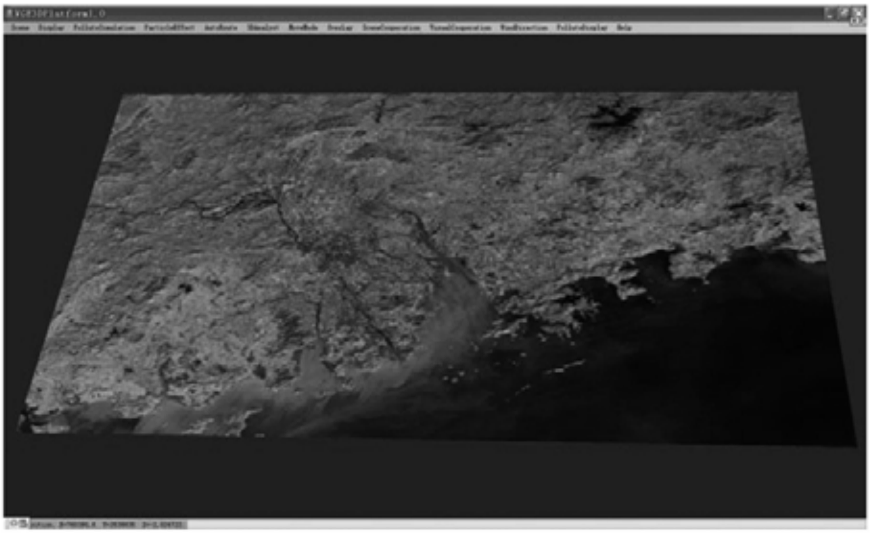


Figure 6.19: 3D environment interface

6.3.3.1 Terrain visualization

Multiple scales terrain data are used to visualize terrain. The data model is Digital Elevation Model (DEM). For entire PRD DEM, the resolution is 90 meters, which is the smallest scale. The middle scale is 30 meters for Hong Kong region, while the largest scale is 1 meter for the campus of the Chinese University of Hong Kong (CUHK).

The textures to be pasted on terrain also have multiple resolutions. The texture covering whole PRD has the resolution of 30 meters which generated from remote sensing TM data. Hong Kong region texture resolution is 5 meters and CUHK texture is 1 meter.

Table 6.6: Data for terrain visualization

Name	Resolution	Type	Size
PRD DEM	90 meters	Raster (.img)	20MB
PRD texture	30 meters	TM	1.4GB
Hong Kong DEM	30 meters	Raster (.img)	730MB
Hong Kong texture	5 meters	Spot5	480MB
CUHK DEM	1 meter	Raster (.img)	78MB
CUHK texture	1 meter	Raster (.img)	58MB

Following Figure 6.20 shows multiple scales terrain visualization.

Figure 6.20: Multiple scales terrain integration and visualization

The terrain data contains three scales is very big. For example in this research, the PRD DEM is about 20MB and the texture that covers PRD is 1.4GB. The 30 meters resolution DEM covering Hong Kong has the size of more than 700MB. To render such big size terrain data is not an easy thing. In this research, two technologies adopted to make terrain rendering smoothly. One is level of detail (LOD), and the other one is Triangulated Irregular Network (TIN). LOD is used to decrease terrain detail when view from far. When moving view point closer to terrain, the detail rendering should be made. But the detail will not slow down rendering because the frustum culling is also active at the same time. The other way to smooth big size of terrain data rendering is TIN, which try to use a few triangles to represent flat surface, while more to create bumpiness areas.

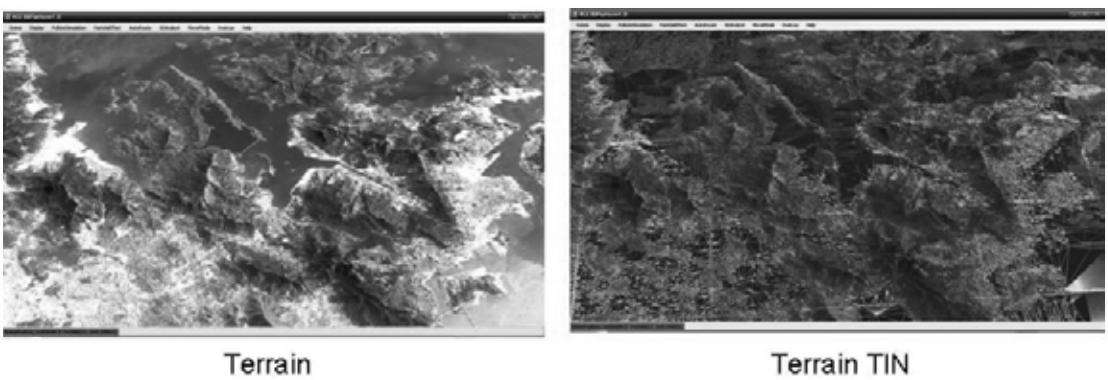


Figure 6.21: TIN to represent terrain

6.3.3.2 Weather simulation

Intuitive weathers simulation can improve immersion. In this research, some typical weathers are implemented which include rain, snow, cloud, lighting, fog and so on. In CVGE, weathers can be control by open or close tags. Intensities for each kind of weather can also be set. Pollution mass, cloud, rain, and snow are all simulated with particles system.

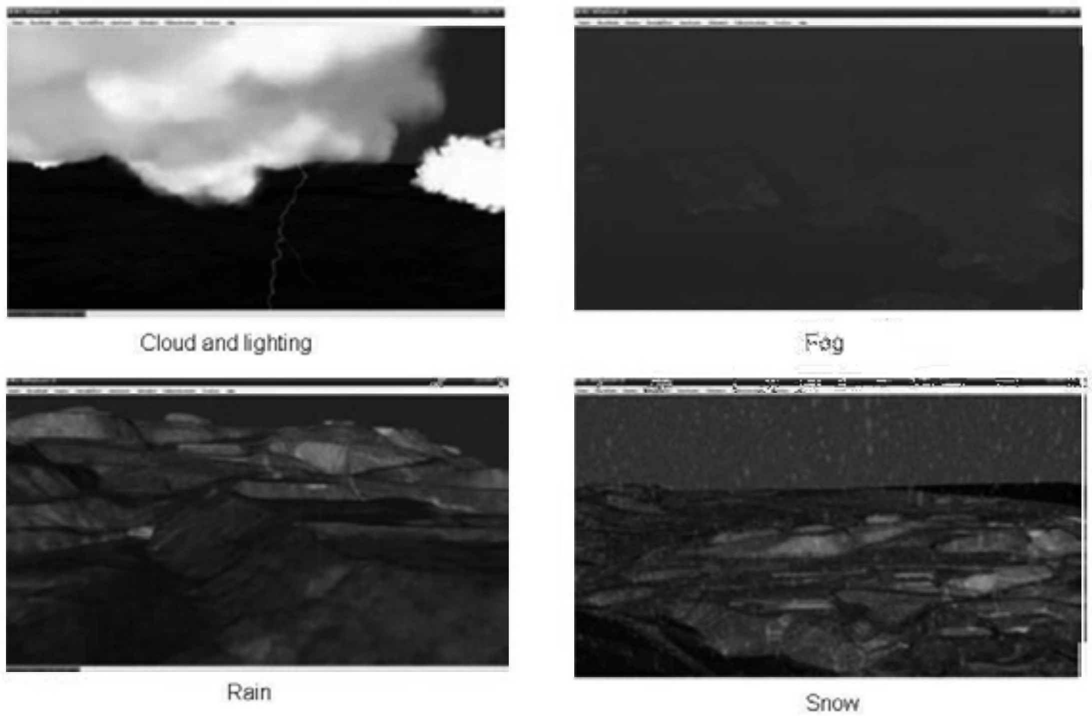


Figure 6.22: Typical weathers in CVGE

6.3.3.3 Virtual environment edition

Unlike physical world, virtual environment should be editable in CVGE, such as digging a hole on terrain, adding a factory at some place and so on. In this research, some basic functions of virtual environment edition, such as terrain elevation changing, pollution factory adding and deleting, typical plants adding and deleting are implemented.

The factory adding and deleting is the primary function for scenarios simulation. At some time, we want to put a new factory at a place to see what is its impact on soundings from the air pollution emission, or remove a factory to see if the sounding air quality can be improved or not.

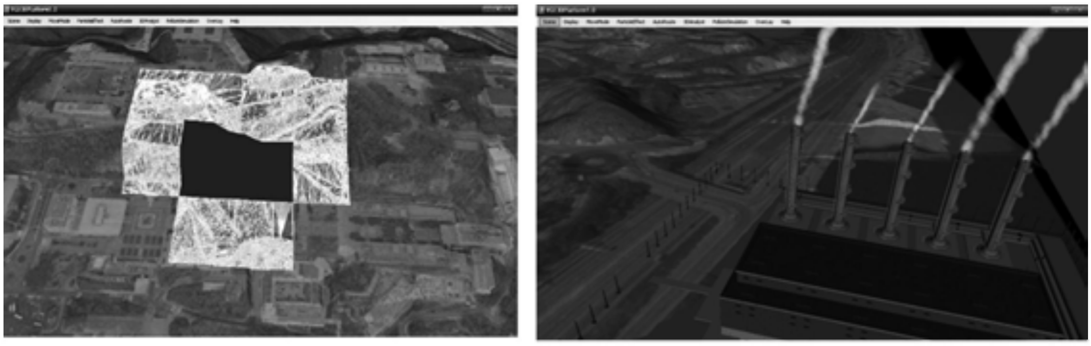
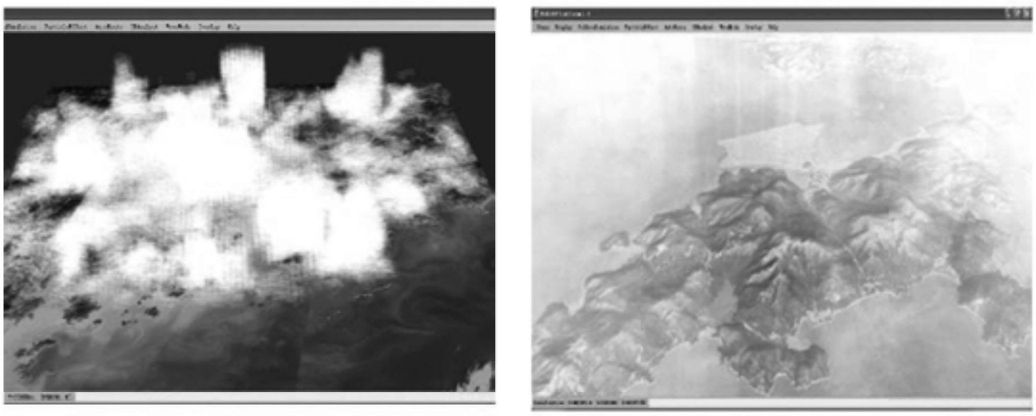


Figure 6.23: Virtual environment edition

6.3.3.4 Pollution mass visualization

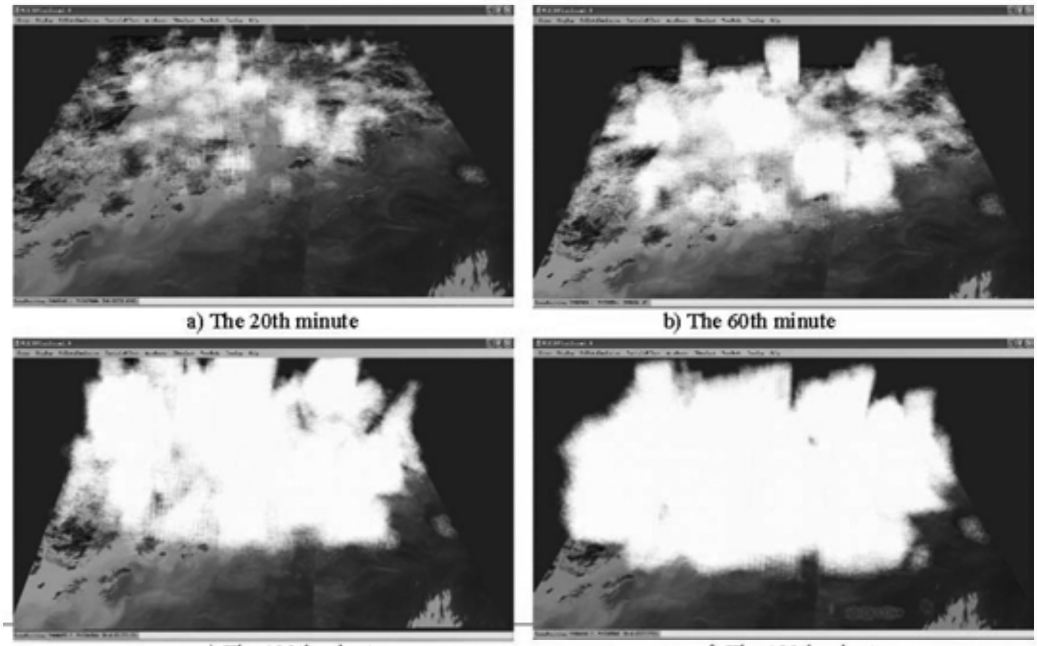
Based on particle system and billboard texture rendering, air pollution mass is visualized. Like real world, air pollution mass looks like a cloud from out side (Figure 6.24 left). But when moving into it, sounding looks like fog (Figure 6.24 right).



View from out of pollution mass | View from inside of pollution mass

Figure 6.24: Pollution mass visualization

Above pollution mass is based on all layers data at a time. If make time go step by step, an animation of air dispersion can be made. Figure 6.25 shows the time shots at four time points. In Figure 6.25, pollution mass expands from simulation time 20 to time 180.



a) The 20th minute

b) The 60th minute

c) The 120th minute

d) The 180th minute

Figure 6.25: 3D air pollution mass animation

In 3D environment, air pollution distribution and dispersion for a layer are also implemented. Figure 6.26 gives an example of a layer pollution distribution. Making simulation time auto going, layer pollution dispersion can be gotten.

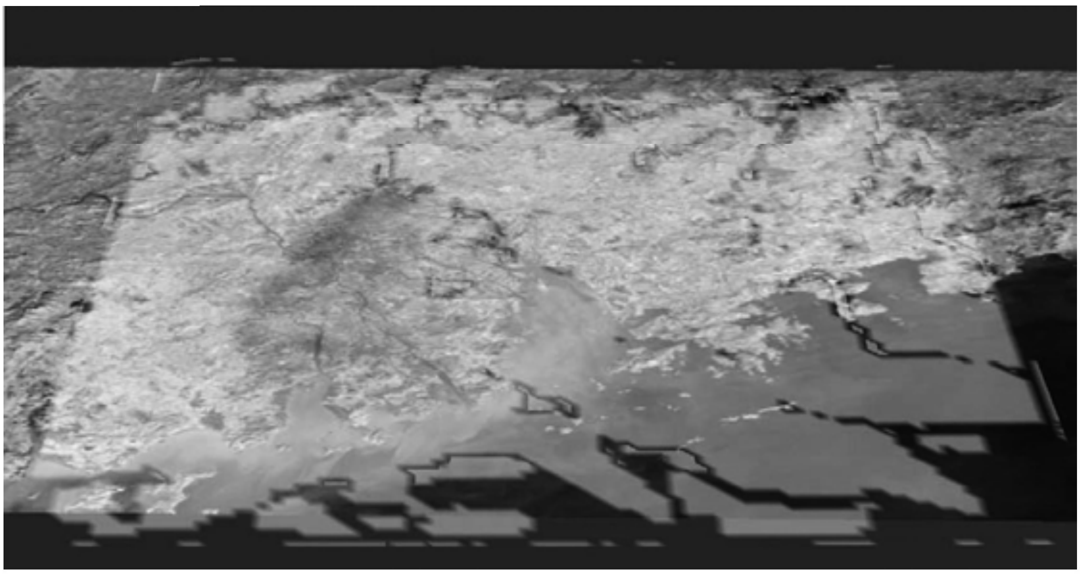


Figure 6.26: One level pollution distribution in 3D environment.

6.3.3.5 Analysis visualization in 3D

Air pollution dispersion is closely related with wind field. So, to get to know the behaviors of pollution mass, wind field should be visualized. Arrows are used to represent wind (Figure 6.27) speed and direction. An arrow length and angle show wind speed and direction at the grid. The direction of wind is in 3D.

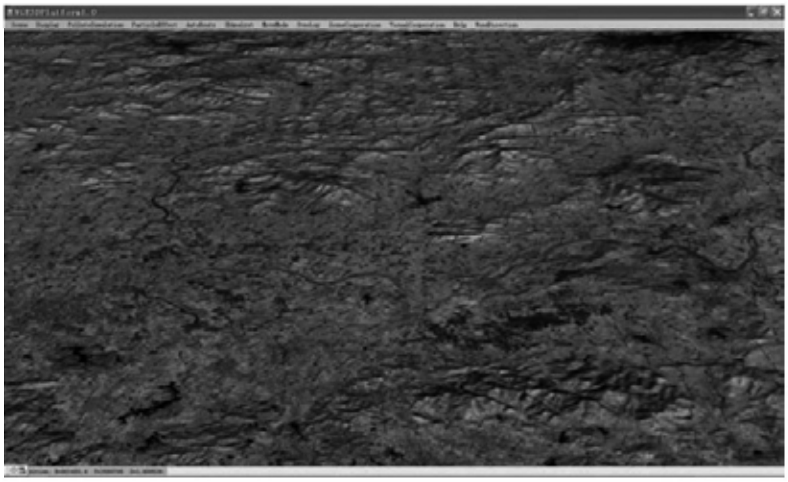


Figure 6.27: Wind field visualization

3D virtual environment supports two kinds of pollution queries. One is point profile query, and the other one is line transection query. Point query is used to display the pollution profile at a clicked point, which is given as Figure 6.28 (left). After drawing a line on terrain surface and opening query operation, the line transection of pollution distribution is displayed like Figure 6.28(right).

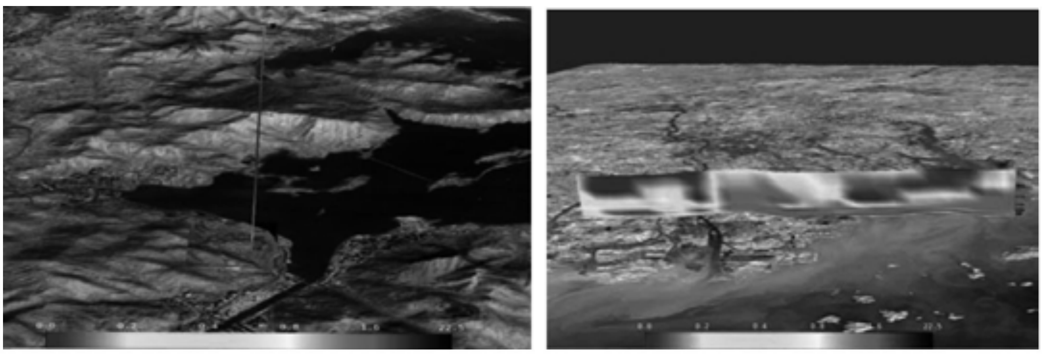


Figure 6.28: Point based pollution analysis line and line based pollution analysis face (Left: Point pollution profile; Right: Line pollution transection)

Some other analysis functions are also implemented. In Figure 6.29 (left), a distance from two points is calculated. In Figure 6.29 (middle), theme layers are overlay on terrain. For the right picture of Figure 6.29, the Visualization ToolKit (VTK) is applied to created isosurface of pollution concentration.

Figure 6.29: Measure, overlay and equal value surface visualization in 3D virtual environment.

CHAPTER 7: GEO-COLLABORATION FOR AIR POLLUTION SIMULATION

Geo-collaboration was defined by MacEachren as, "group work about geographic scale problems facilitated by geospatial technologies" (MacEachren et al. 2003). In this chapter, the collaboration for air pollution simulation is prefixed with "geo" because: ① the data (including air pollution sources) is geographically referenced; ② the models have spatial and temporal parameters, and are coupled with geo-data and geo-information; ③ the visualization is geographically located; ④ the analysis is spatial and temporal; and ⑤ the tool used to facilitate this simulation is a CVGE, which is developed using geo-spatial information technologies.

This chapter will first identify the contents that can be targeted for geo-collaboration in the lifecycle of air pollution simulation. The chapter will then discuss the modes which are fit for collaborative air pollution simulation. In the third part, the architecture is designed to organize geo-collaboration on air pollution simulation, and each part of the collaboration will be studied in detail. The final part concerns geo-collaboration implementation.

7.1 Contents of Geo-collaboration on Air Pollution Simulation

In order to achieve geo-collaboration on air pollution simulation, the targets for collaboration must first be identified. Air pollution simulation has many components: data preparation, modelling and computation, visualization, analysis and decision making. Thus, the geo-collaboration which supports this simulation should tackle each of the above operation components. This research identifies the specific contents of geo-collaboration for air pollution simulation as: air pollution sources, setting and computing geo-models, visualization and analysis. The details of these contents are discussed below.

(1) Shared pollution sources

Shared pollution sources are the basis for air pollution model computation. In the real world, pollution sources are dynamic in that they change emission size, emission velocity and so on. Meanwhile, some pollution sources may be closed, while others may be newly established. In order to achieve a reasonable simulation, the above information regarding changes should be input into the system as early as possible. Unfortunately, the pollution sources for this research are scattered around the research area. It is impossible to accomplish this work using centred workers for a large region. The optimum method is to open the information updating priority to all of the distributed experts, officers and even the public scattered throughout the research region. The participants, in spite of their specific roles, have a priority to update the pollution sources in cities where they live. An update by any one participant can be detected automatically by all of the others. This constitutes the collaboration on pollution sources maintenance.

Another form of collaboration on existing pollution sources is the collaborative compiling of air pollution sources to get deferent scenarios. Let us take the air pollution sources in PRD as an example. Supposing Hong Kong governmental officials and Guangdong governmental officials work collaboratively to simulate the impact of industry emissions on the regional environment. They agree to close all food processing factories. When the compilation of pollution sources, governmental officials from Hong Kong will deal with the sources that belong to Hong Kong SAR., while the officials from Guangdong deal with those sources that belong to Guangdong province.

(2) Geo-models setting and computation

The second element of collaborative air pollution simulation concerns the setting and computation of geo-models, engaging experts from multiple disciplines. The input parameters of geo-models are various. As in previous discussion, it is more reasonable that some parameters, such as land use and land cover (LULC), should be input by geographers, while other parameters are more rationally input by atmospheric scientists, such as the atmospheric boundary layer. Thus, in order to achieve the optimal setting for models, multiple participants should work collaboratively, taking care of their own professional setting while leaving those that are beyond their knowledge.

The above set-up is cooperative operation, which means that all the participants must fully trust each other. At some time, there is a competitive setting, which means that participants are only confident in their own setting. In this condition, each participant attempts to control all settings, and ignores others' work. This condition would cause conflicts when different participants set shared parameters with different values.

(3) Visualization and analysis

The third and fourth elements of geo-collaboration concern visualization and analysis. Collaboration on the visualization of virtual environments includes virtual scene navigation linkage and object consistency in virtual environments. Under collaborative conditions, all participants share the same virtual camera to capture virtual scenes, and all participants will see the shared virtual environments from the same viewpoint. The consistency of objects in virtual environments is another collaborative aspect of the visualization of virtual environments. If a participant changes his virtual objects in their positions, shapes or any feature, the changed results can be detected by other participants automatically.

Analysis collaboration means that any analysis conducted by a participant can be viewed by the other participants.

7.2 Modes and Mediator of Geo-collaboration for Air Pollution Simulation

7.2.1 Modes

According to the dimensions of time and place, two modes of geo-collaboration for air pollution simulation are supported in this research: one is different place, same time geo-collaboration; and the other one is different place, different time geo-collaboration. In real application, these two modes can be mixed to form a hybrid mode.

(1) Different place, same time geo-collaboration

Multiple participants are located at geographically dispersed locations. During the process of air pollution simulation, all of the participants operate each step synchronously. They start with data preparation, then geo-modelling and computation, geo-visualization and finally analysis. This mode is shown in Figure 7.1. The same time here does not mean that all the steps will be conducted synchronously. In fact, the air pollution simulation step must take place in the required order. Same time here means that multiple participants will conduct same operations on the same steps synchronously.

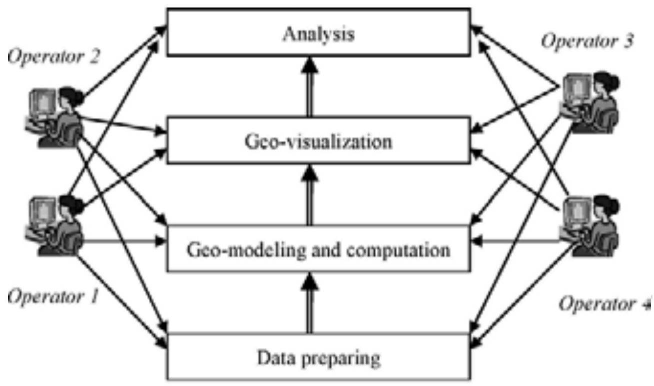


Figure 7.1: Different place, same time geo-collaboration for air pollution simulation

In this mode, conflicts exist in each step. In order to avoid confusion in the simulation, the conflicts must first be solved before participants move on to the next step. The methods used to detect and solve conflicts are discussed in the next section.

(2) Different place, different time geo-collaboration

Multiple participants working from geographically dispersed locations join CVGE to form a working group for operating air pollution simulation. The steps for air pollution simulation still keep the required order. However, unlike the first mode, each step in the air pollution simulation is only controlled by one operator. This mode is shown in Figure 7.2.

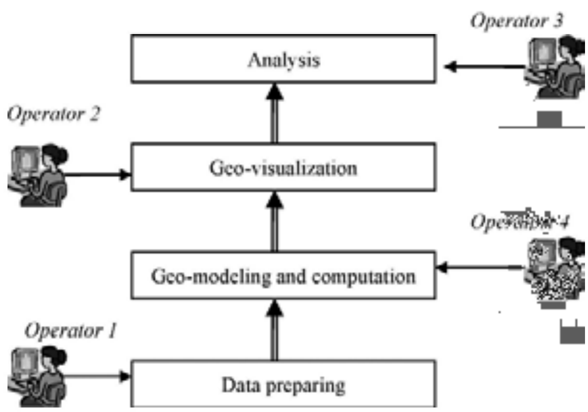


Figure 7.2: Different place, different time geo-collaboration for air pollution simulation

This mode in fact forms a collaborative workflow. Because there is only one participant operating each step, there is no conflict during the geo-collaboration.

7.2.2 Mediator

The mediator is a kind of shared intermedium, which can be accepted and understood by all participants. In both real and virtual environments, there are many mediators used for communication, discussion and collaboration. For example, in daily life, a blackboard is a mediator for communication between teachers and students, and a map is a mediator for discussion on urban planning between governmental officials and designers. In virtual environments, email, online chat room (such as Windows Messenger) and whiteboards are all used to mediate information between senders and receivers.

Computer support collaboration depends upon the mediator connecting with multiple participants for knowledge sharing. For geo-collaboration, the mediator should be created using geo-spatial technologies. A map is regarded as one form of mediator for geo-knowledge communication (Goodchild 2000). For example, Alan and Cai adopted a map to mediate between multiple participants to achieve geo-collaboration (MacEachren et al. 2006, Cai et al. 2005, MacEachren 2001, MacEachren 2000).

In this research, a shared collaborative virtual geographic environment is adopted as the mediator for multiple participants to collaborate on air pollution simulation. Unlike a map, CVGE has many advantages, being distributed, multi-dimension geo-visualization, supporting multi-channel interaction between operators and computer. A CVGE is an integrated mediator, combining map, diagrams, tables, high dimension visualization and chat room. Thus, it can also be regarded as a hybrid mediator, which means that participants may use different mediators in their geo-collaboration. For instance, in order to collaborative in compiling air pollution sources in CVGE, participant A may use a 2D environment, while participant B uses a 3D environment. Even though the mediators are different, the geo-collaboration is still achieved. In order to ensure efficient geo-collaboration between multiple types of mediator in CVGE, the following points are recommended.

(1) Share contents between multiple mediators

As was stated in the discussion above, participants may use different mediators to conduct geo-collaboration in CVGE. The differences are only limited to geographic visualization. However, the contents behind the visualizations must be shared. For example, to collaborate in identifying a building as an air pollution source, a mediator with a 2D map represents it as a point or a rectangle, while another mediator with a 3D environment represents it with a 3D virtual building. Although the representations of the visualizations are different, the building's location, height and emission parameters are all shared.

(2) Consistency of projects between multiple mediators

2D and 3D environments are based on projects, such as WGS84. The same location values in different projected environments will result in different locations. Thus, in order to support the exchange of geographic locations, multiple mediators must be consistent in the project.

(3) Awareness of the other multiple mediators

Supposing a participant conducts some operations for geo-collaboration, other participants should be aware of these operations, and should consequently update their environments to maintain consistency. This awareness should be automatically done by the mediators.

7.3 Developing Geo-collaborations for Air Pollution Simulation

7.3.1 Architecture for geo-collaboration for air pollution simulation

The architecture for geo-collaboration for air pollution simulation is presented in Figure 7.3. Geo-collaboration occurs in data sharing, model setting and computation, visualization and analysis. Geo-collaboration is a kind of group working environment. If a participant wants to join the geo-collaboration, he must firstly pass system authentication and be given a role. In this research, five roles are identified: computer scientists, geographers, atmospheric scientists, governmental officials and the public. Each role is limited to the authorized operations listed in Table 7.1.

Table 7.1: Roles and responsibilities for participants conducting geo-collaboration for air pollution simulation

Roles	Responsibilities
Geographers	• To create and maintain geo-data • To provide geo-visualization, spatial and temporal analysis
Atmospheric and environmental scientists	• To initiate air pollution dispersion models • To conduct air circulation modelling and air pollution dispersion modelling • To conduct analysis on air pollution visualization, analysis and model evaluation
Computer scientists	• To maintain a healthy CUGrid • To perform model computation on the CUGrid
Governmental officials	• To update and compile air pollution sources • To show the requirements the each step • To evaluate results
Public	• To serve as monitors on air pollution emission • To vote on decisions

The public can receive system outputs and vote on decisions. Governmental officials supply information on updates of pollution sources and requirements on visualization and analysis. Atmospheric and environmental scientists operate on model setting and computation, visualization and analysis. Geographers work in air pollution result data transformation, geographical visualization and spatial and temporal analysis. Computer scientists only do operations relating to high performance model computation.

In Figure 7.3, the centre workflow from the top down shows the steps involved in air pollution simulation. Around the workflow are the multiple participants who conduct geo-collaboration on air pollution source compiling, circulation and environmental modelling, data transformation, geo-visualization and analysis. In the following sections, each of these collaborations is studied in detail.

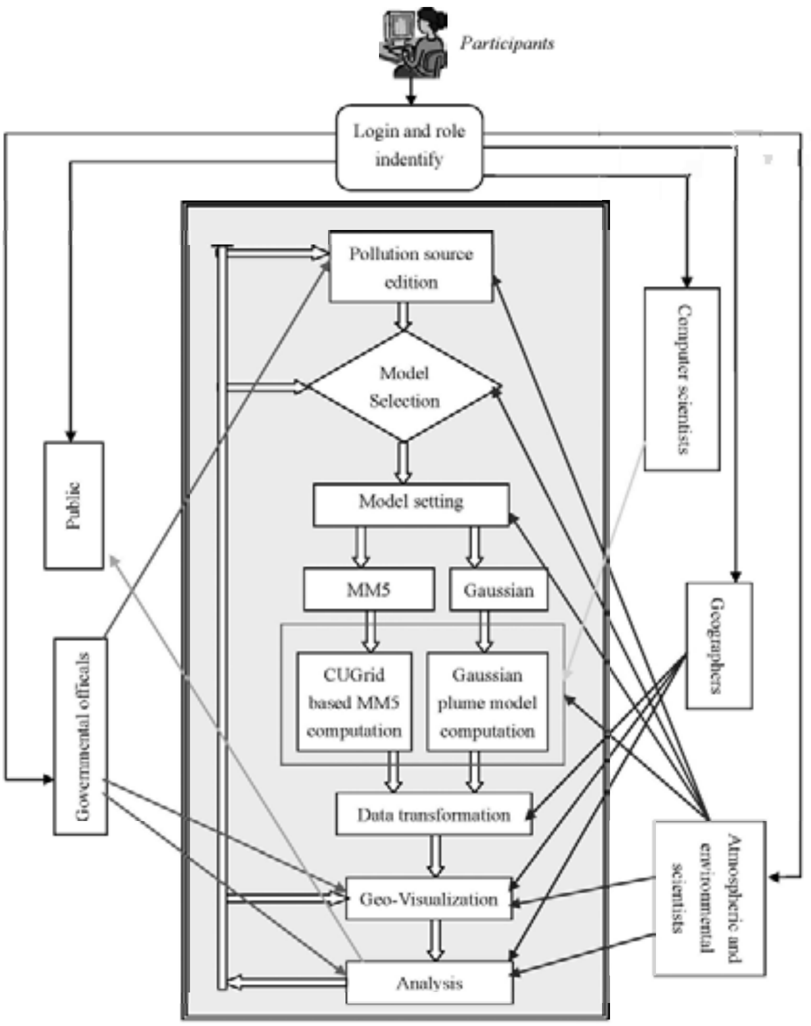


Figure 7.3: Architecture for geo-collaboration for air pollution simulation

The mode of geo-collaboration in Figure 7.3 is hybrid. Some steps are controlled by one participant, while some steps are conducted by multiple participants.

7.3.2 Collaboration in pollution source compiling

Two methodologies are designed to implement the collaborative compilation of pollution data, as shown in Figure 7.4. One of these is a centralized style of collaborative compiling based on shared pollution sources, as in Figure 7.4 (A); the other is distributed collaborative compiling based on copied pollution sources, as in Figure 7.4 (B).

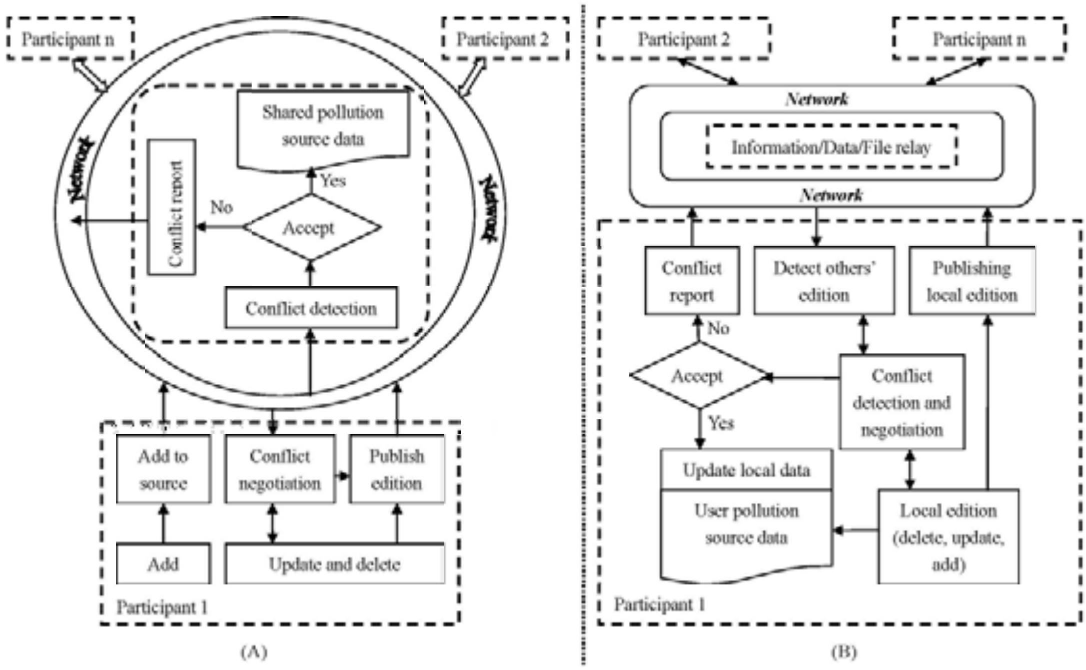


Figure 7.4: Two types of collaborative data compiling (In (A), there is only one copy for the whole system installed in the centralized management node. In (B), each user has a copy of the pollution sources.)

The centralized and distributed styles are defined according to the storage location of the pollution source files. For the centralized style, there is only one copy of the pollution sources installed in the central management node for the whole system. The collaborative operations commands from the participants, including updating and deleting information, will be processed at the management node. On the central node, conflicts in operations will be detected automatically. If there is no conflict, the operation will be accepted and the pollution sources will be compiled. If a conflict does exist, a report will be sent to the participants. The conflict resolution, such as negotiation (which will be studied in the next section), will then be launched to solve any conflicts.

The distributed pollution source style achieves collaboration in another way. In this style, each participant has a copy of the sources in his own node. On each participant's node, four routines are running: ① compiling local sources; ② publishing newly compiled records; ③ receiving remote participants' newly compiled records; and ④ carrying out conflict detection and sending out reports. Therefore, compared with the previous methodology, the conflict detection and negotiation process occur on each of the distributed participants' nodes. In this style, the management node only serves as an information/data/file relay station.

7.3.3 Collaboration on air pollution dispersion modelling

This part studies the conduct of geo-collaboration on air pollution dispersion modelling. The discussion in the previous chapter shows that two kinds of air pollution dispersion model are integrated into CVGE. One is MM5 coupled with SYSUM for regional scale modelling, and the other is the Gaussian plume model for local modelling. The methodology for conducting geo-collaboration on the two models is almost the same. Therefore, in this part, only the Gaussian plume model is examined.

Gaussian plume model collaboration concerns how to set model parameters collaboratively. In this research, ten input parameters are open to users. Supposing three participants join to form a collaboration group. There are two ways of setting the Gaussian plume model. Firstly, the ten parameters can be divided into three groups. Each group of parameters is then assigned to one participant. After the three grouped parameters are set, the results are combined together to form the final input parameters to initialize the Gaussian plume model. In this situation, there is no conflict and so negotiation is not needed. This type of collaboration is in fact a kind of parallel work. For the second method, the three participants set the ten parameters at the same time. In this situation, a comparison will take place to find conflicts for the next steps of negotiation.

The methods for Gaussian plume model collaboration, including the two methods, are presented in Figure 7.5. Before starting collaboration, the style of collaboration will be chosen. If the parallel parameter setting style is selected, participants will be allocated grouped parameters according to some defined rules. After all the participants upload their group of parameters, the final input initiation parameters are generated, which will then be used to drive Gaussian plume model computation. If the second style of parameter setting is chosen, each participant will set all parameters and submit these to the management node for conflict detection. If conflicts are detected, a report will be sent back to the participants and a negotiation will be launched. After all the conflicts are solved, the parameters will be input into the Gaussian plume model. The Gaussian plume model is then calculated by a stand-alone computer or by computers in parallel. The results from the model computation will be sent back to the participants for checking to see if there are any problems. If there are, the parameters can be changed and a new process of setting, conflict detection and negotiation will begin. These two kinds of collaboration on the Gaussian plume model are shown in Figure 7.5.

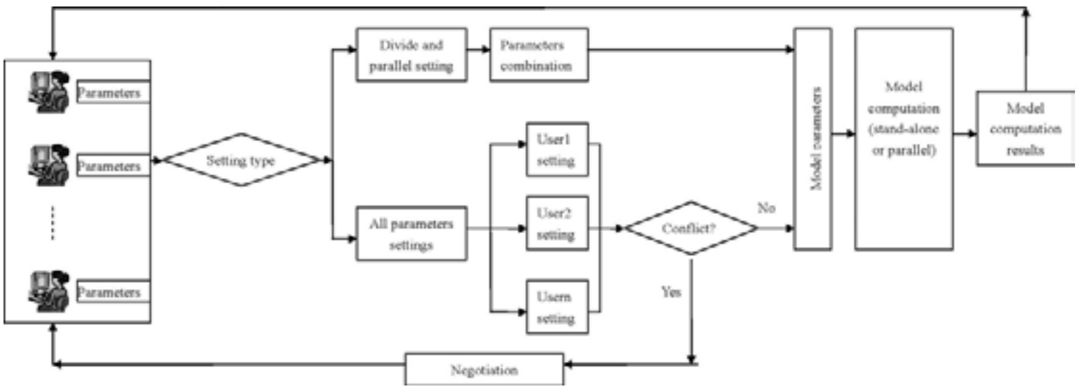


Figure 7.5: Flow diagram of collaboration for Gaussian plume model computation

7.3.4 Visualization collaboration

Visualization collaborations include viewpoint collaboration and object/process collaboration. The visualized scene on screen for a virtual environment is determined by a virtual camera. As Figure 7.6 shows, a virtual camera can be fixed using four parameters, which are height and the three rotation angles around the X, Y and Z axes, as (α, β, γ). The parameters of the virtual camera, together with the angle of the vision field (θ), form the final picture on screen. Therefore, to approach visualization collaboration in virtual environments, two aspects must be realized at the same time. The first is to share the five parameters of the virtual camera, which are the camera height (H), the three rotation angles around the X, Y and Z axes (α, β, γ) and the angle of the vision field (θ). The second is to share the contents of the field of vision. If the above two kinds of sharing are achieved, visualization collaboration can be realized.

Returning to air pollution, the visualization collaborations exist in both 2D and 3D environments. In a 2D environment, the three rotating parameters are fixed to set the virtual camera upright facing the terrain. The height increasing and decreasing will induce a change to the map scale. The position of the virtual camera will impact the map location. Therefore, collaboration in a 2D environment can be achieved by sharing the contents of a map and the two parameters of the height and location of the virtual camera. In a 3D environment, all the parameters of a virtual camera are free. The collaboration is thus more complex than in a 2D environment. To approach visualization collaboration in a 3D environment, in addition to the contents of the field of vision, the height, the angle of the vision field and the three rotation angles of the virtual camera should all be shared. Sharing the parameters of a virtual camera can be achieved using virtual scene navigation, while the shared contents can be realized through virtual scene collaborative editing.

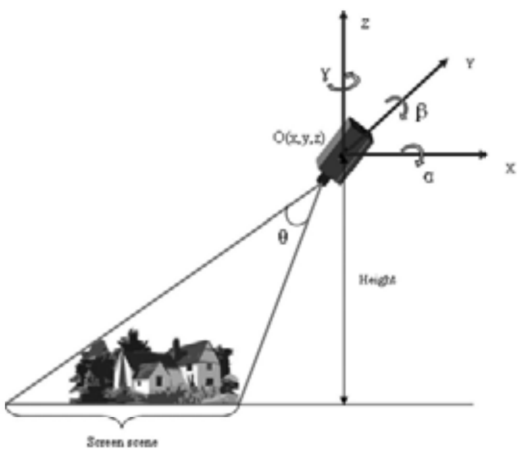


Figure 7.6: Virtual camera and screen scene relation

The process of visualization collaboration is shown in Figure 7.7. When a participant wants to conduct collaboration visualization, he should check whether there are any opened collaborations. If there are none, he may create a new one which can be joined by other participants. According to whether they create or join the visualization collaboration, participants can be classified into two groups: the visualization collaboration creators (VCC) and followers (VCF). Only one participant can be VCC, while there can be any number of VCF. VCC has the right to control the virtual camera, while a VCF can only follow it. However, for collaboration on the content of a virtual environment, VCC and VCF have the same priorities to add objects, delete objects, update objects and process control. After each participant edit (including adding, deleting, updating and controlling) on any object or process, the results will be published. The published contents will be checked by a conflict detection module. If a conflict exists, negotiation will be launched to solve this. The solved contents are then shared as the visualization contents, which can be used, together with the virtual camera, to update the computer screens of all participants.

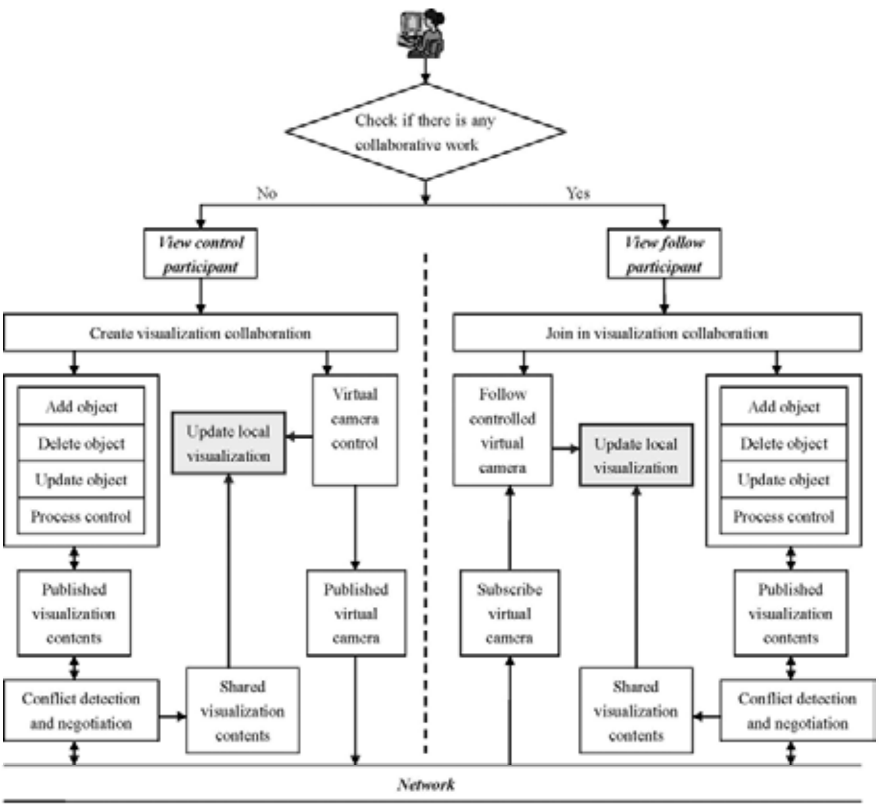


Figure 7.7: Visualization collaboration

7.3.5 Analysis collaboration

Analysis collaboration is combines and shares the analysis of all participants. Each participant conducts his analysis alone, and these are then published for sharing automatically. Participants can also subscribe to others' analyses. Therefore, the final contents of analysis collaboration are the sum of each participant's analysis.

7.3.6 Process collaboration

Process collaboration can be viewed as a combination of all the above collaborations, including data collaboration, modeling collaboration, visualization collaboration and analysis collaboration. It is a loop, and is also the basis for decision making. This collaboration can stop at any step, go back to a previous step or speed onto the next step. If an extended function of recording is enabled, the playback will be valuable for training learners, which will make the system a training platform.

7.4 Conflict Detection and Resolution in Collaborative Air Pollution Simulation

Definition of the term conflict depends on the perspective and issues under consideration. Therefore, there is no one definition satisfactory for the physical domain of collaborative engineering (Lara and Nof 2003). Conflicts can be caused by task, role, responsibility ambiguity and lack of trust, and these can be resolved using methods such as negotiation, mediation, facilitation, arbitration and litigation (Shin 2005). A conflict can be categorized as competitive conflict or collaborative conflict, and the resolutions of conflict are specific to each individual system (Lara and Nof 2003).

Negotiation is a process wherein conflicting participants attempt to reach an agreement about the issues on which they presently or potentially might disagree

(Shin 2005). Although a variety of electronic technologies can be employed as means of conflict resolution in virtual settings, such as e-mail, intranet and on-line chat, the last is ideal for virtual workers because it allows multiple parties (including the mediator) to communicate synchronously (Shin 2005).

Two stages comprise this part: one is conflict detection in the collaboration for air pollution simulation; the other is how to solve the conflict. The conflict detection and resolution discussed in this section is not a general discussion, but is specific to problem solving.

7.4.1 Conflict detection

Conflict arises when concurrent operations are conducted on shared objects. Conflict here means that there are different operations on shared contents in an air pollution simulation system. Although participants may have different opinions in mind, as long as they are not input into a system, no conflict will arise. Therefore, conflicts here are restricted to those that occur in a system.

Conflicts are caused by different value settings or operations on shared targets. Different value settings can include several conditions. The first is giving a parameter a different value. For example, one participant sets a wind speed of 2m/s, while another sets it at 5m/s. Another condition is setting a parameter to different units. For instance, one participant sets wind speed using m/s, while another uses km/h to measure it. An operation conflict is caused by multiple participants performing different operations on a shared object at the same time. There are three factors that work together to induce operation conflicts: *different operations* on *shared targets* at the *same time*. Only if all three of these factors are fulfilled do conflicts arise.

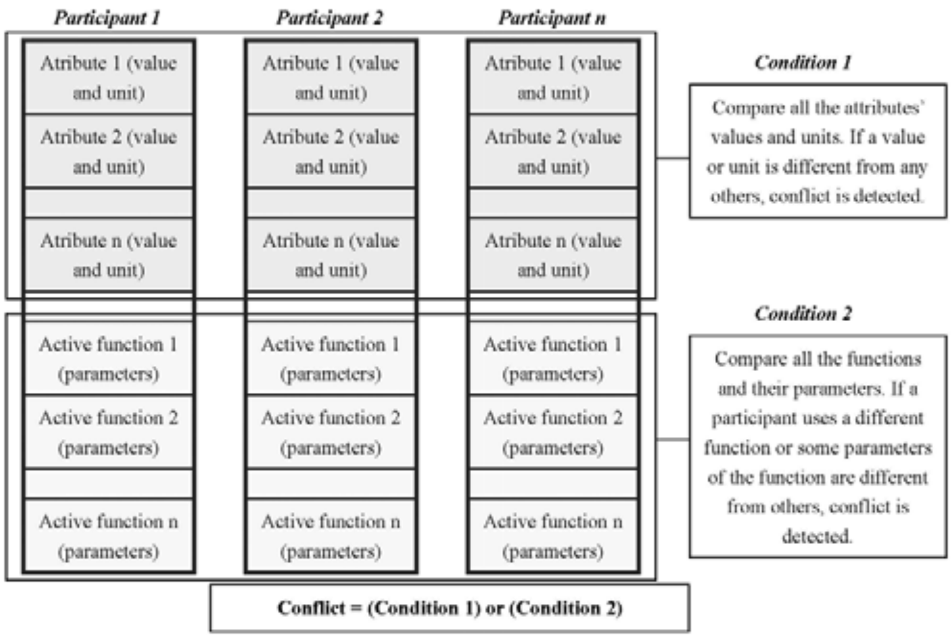


Figure 7.8: Conflict detection

Conflict detection can be simply achieved by operation comparison, which is the approach adopted by this thesis. The target (such as an air pollution source) for collaboration can be defined as an object with attributes and functions. During the collaboration, each participant will have a copy of the target. Any operation on the copied target will be recorded as an object. When all the copied targets have been collected by the conflict detection module, a comparison will be made to check whether there are any conflicts among the attributes and functions. Attribute conflict detection is achieved by comparing all of the attributes' values and units. If one or several of the values or units are different from others, conflicts are detected. In order to detect function conflicts, all of the functions of the copied targets will be compared. The functions that will be compared are restricted to active functions. Therefore, if the active function invoked by a participant is function A, while another participant invokes function B, a conflict arises. In another situation, if both participant A and participant B invoke the same function, but give different parameters, a conflict is also generated. These methods of conflict detection are shown in Figure 7.8.

7.4.2 Conflict resolution

In this part, two forms to solve conflicts will be discussed: rules based reasoning and communication based negotiation.

(1) Rules-based reasoning

In this research, the conflicts happen in a system, but they are reflections of conflicts involving real participants. The rules used to solve air pollution simulation conflicts attempt to harmonize the different opinions of participants. In order to establish rules, two factors are considered, which are the roles and the number of participants who hold the same opinion.

Five roles are considered in this research: geographer, atmospheric experts, computer experts, officers and the public. The content of the research, which is also the target for the collaboration, includes editing pollution sources, establishing air pollution models, computing air pollution models, transforming the results of model computations, visualizing the results of model computations and analysis of the results of model computation. For operations on shared targets, different roles are granted different weights, and these are listed in Table 7.2.

Table 7.2: Table of weights for roles on collaboration contents

Contents	Weight for geographer (GW)	Weight for atmospheric experts (AW)	Weight for computer experts (CW)	Weight for officers (OW)	Weight for public (PW)
Pollution source edition (Cont1)	1	1	0	1	0
Air pollution model setting (Cont2)	1	2	1	1	0
Air pollution model computation (Cont3)	1	3	2	1	0
Computation result transformation (Cont4)	1	0	0	0	0
Pollution result visualization (Cont5)	1	1	1	1	0
Pollution result analysis (Cont6)	1	1	1	1	0

In Table 7.2, there are no correlations on the rows. If the role has no right to operate on the content, the weight is set to 0. If the role has a higher priority for operating on the content, it will be given a higher weight. If several roles are given same weight, this means that they have the same priority for the relevant operation. Table 7.2 is scalable, and the weights for all of the content (the rows in Table 7.2) are continuous. The weight arithmetic is explained in Figure 7.9.

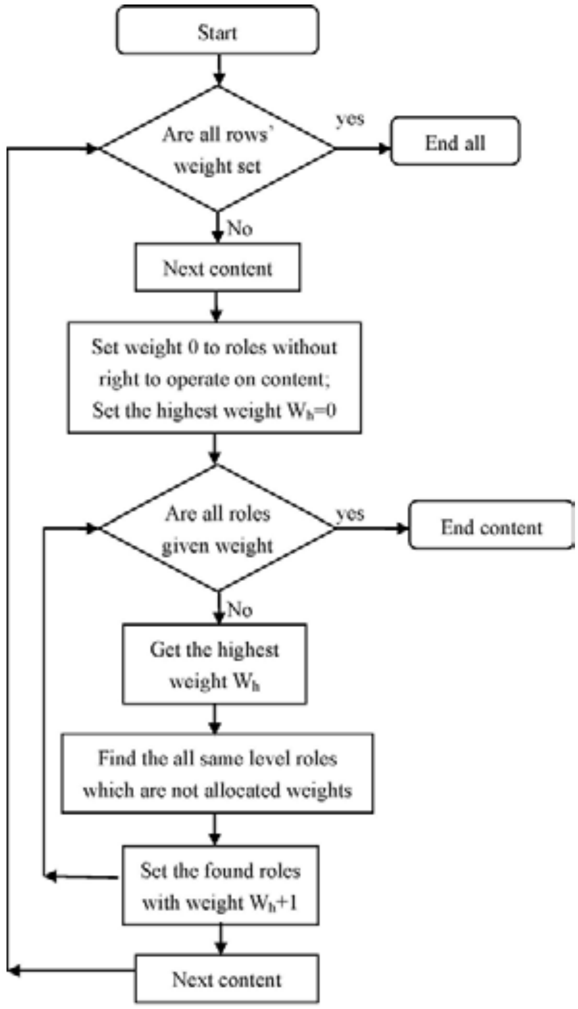


Figure 7.9: Weight arithmetic

Besides the weight of roles, there is another factor that is considered in the generation of rules, which is the number of participants who are filling same role and who hold the same opinion (collaboration content).

In this research, there are two rules used for conflict negotiation: one is aimed at single role conflict negotiation, while the other one relates to conflict negotiation between multiple roles.

Rule 1: Conflict reasoning for a single role

If there is a single role involved in the collaboration, take the number of participants who set a target with a value of V_1 as N_1, while N_2 is the number of participants who set targets with a value of V_2, and so on.

Supposing:

Max (N_1, N_2, …, N_n) — The maximum of N_1, N_2, …, N_n

n = The number which is how many of (N_1, N_2, …, N_n) is equal to the Max (N_1, N_2, …, N_n)

Rule 1 can be described as follows:

If n does not equal 1

Participants reset value again

Else

If N_i equals Max (N_1, N_2, …, N_n)

Final value $V = V_i$ and the number of participants who hold V_i is $N_1+N_2+ \ldots + N_n$

Rule 1 can be used for conflict negotiation on editing pollution sources, visualizing pollution results and analyzing pollution results. In application, the roles with the same weight can be treated as being the same role.

Rule 2: Conflict reasoning for multiple roles

Conflict between same role participants has been solved by rule 1. However, suppose there are N_1 participants with R_1 (Role 1) who set the collaboration value at V_1, while N_2 participants with R_2 set the collaboration value at V_2 and so on. At the same time, suppose R_1 has the highest priority for setting the value, followed by R_2

and then R_3, R_4 … R_n. This kind of multiple role conflict can be represented as in the matrix below.

$$\begin{bmatrix} R_1 & N_1 & V_1 \\ R_2 & N_2 & V_2 \\ \vdots & \vdots & \vdots \\ R_n & N_n & V_n \end{bmatrix}$$

Rule 2 can be given as:

If N_1 equal to 0

{

If N_2 equal to 0

{

If N3 equal to 0

{

……..

(Recursion here)

}

}

Else

The final value is V_2

}

Else

The final value is V_1

Rule 2 is a predominated principle emphasizing the highest priority role, which can be applied to conflict negotiation on air pollution modelling and computation.

(2) Communication based negotiation

Rule based negotiation has its disadvantages. For example, it is impossible to cover all of the conditions using these rules, especially when a lot of factors are considered. In addition, rule based negotiation is complex not only to understand, but also to implement.

Communication based negotiation, which has been widely used, is a more open and efficient way of resolving conflict. The forms of communication here include online text, audio, video and whiteboards. The communication is content independent, which means that collaboration can be used to solve any kind of conflict. Because it is content independent, once text, video, audio or whiteboard is established, they can be used to solve any form of collaborative conflict changing the implementation code.

7.5 Implementation

7.5.1 Geo-collaboration module distribution and implementation

The geo-collaboration modules for air pollution simulation are distributed into all of the nodes, including the system management node, participants' nodes and model computation nodes.

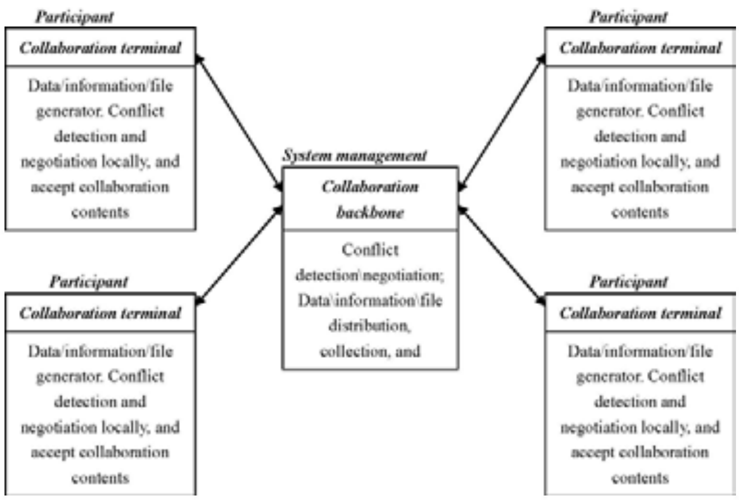


Figure 7.10: Geo-collaboration modules and distribution

Geo-collaboration is the backbone of the whole system, as is the system management node. The discussion in the sections above shows that conflict detection and negotiation can take place either centrally or dispersedly. If conflict detection and negotiation takes place on the system management node, the geo-collaboration type is central. However, if conflict and negotiation happen on participants' nodes, the geo-collaboration is distributed. Figure 7.10 integrates both kinds of geo-collaboration. When central conflict detection and negotiation is selected, all of the participants' nodes send information to the central node. The conflict detection and negotiation is then carried out on the system management node. Following this, the accepted content is sent back to the participants. If the distributed type of conflict detection and negotiation is chosen, the system management node will only relay data/information and files between participants. The distributed participants carry out two tasks: one is conflict detection and negotiation locally; and the other is generating information, data and files for collaboration. Figure 7.11 shows the interface for a collaboration module on the system management node. The interface is in fact the action monitor, while the collaboration process automatically takes place behind this.

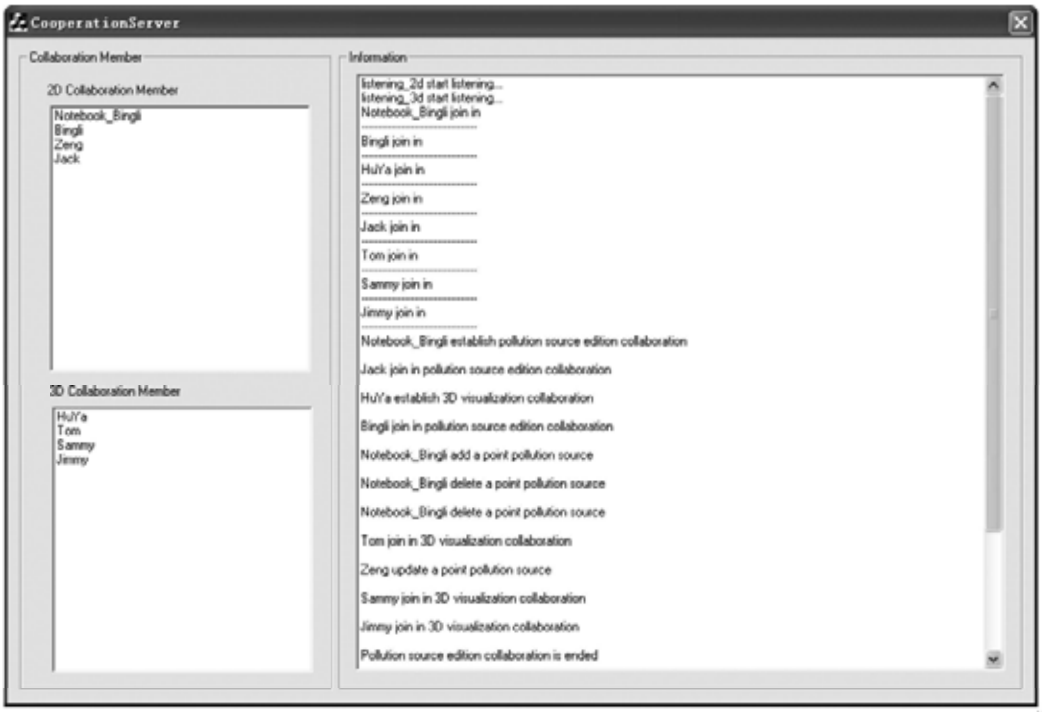


Figure 7.11: Interface of geo-collaboration on the system management node

7.5.2 Typical function implementation of geo-collaboration

(1) Pollution source collaborative compiling

Figure 7.12 shows collaborative compiling of pollution sources. Two participants from geographically distributed locations join to form a collaboration group. The upper participant deals with the pollution source within the top region, which has green boundary lines. The lower participant can edit the pollution source inside the polyline colored brown. Both editions can be viewed and detected by either of the two participants synchronously or asynchronously. The edition here includes a new pollution source addition, source deleting and source updating.

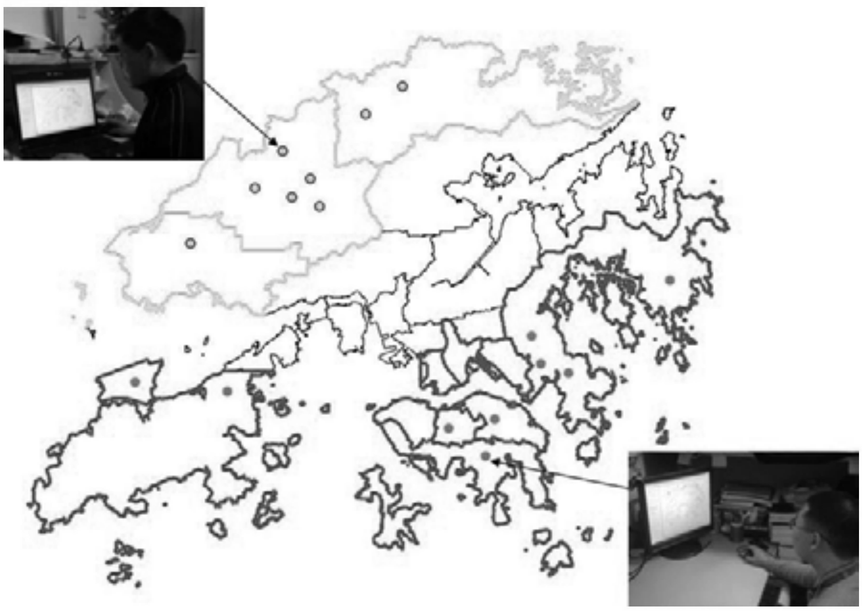


Figure 7.12: Pollution source collaboration compiling

Collaboration compiling of pollution sources can be carried out not only within same platform (2D or 3D), but also on different platforms at the same time (2D and 3D). Figure 7.13 shows collaboration between two kinds of environment (2D and 3D). In Figure 7.13, when a participant adds a new point source in the 2D environment, the source will automatically be added in the 3D environment.

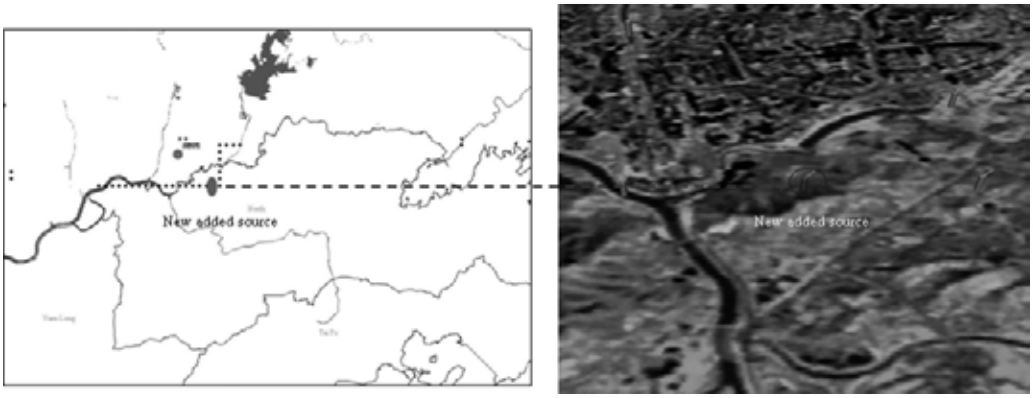


Figure 7.13: 2D and 3D environments consistency for adding a pollution source

(2) Collaborative setting on a Gaussian plume model

Two types of collaborations are implemented on the Gaussian plume model setting. The first is a parallel setting from multiple participants, as shown in Figure 7.14. In this type, the parameters are grouped. Each group is assigned to different participants according to their roles. In Figure 7.14, three participants (Xubingli, HuYa and ZengLiping) form a collaborative group to set the Gaussian model from geographically distributed locations. When the collaboration started, each participant registered his favorite parameters. The registration is based on a rule of first registering, first controlling, which means that if two participants register for the same parameter (such as wind speed), the one who registers earlier will be given the right to control the parameter. Therefore, the registration rule can be viewed as a methodology for avoiding conflict for this type of collaboration. During parameter setting, each participant can input his controlling values, while at the same time other participants' settings are also presented through online data. After all of the participants finish their settings, the parameters can then be input into the Gaussian plume model, which is run on a local computer or on parallel computers.

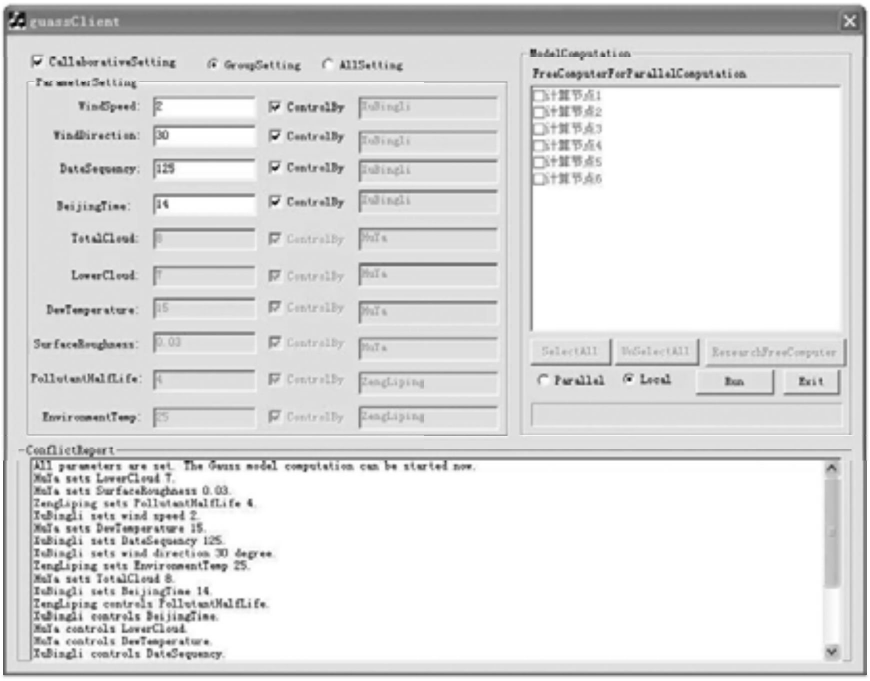


Figure 7.14: Gaussian plume model parameters are grouped and parallel-set (The parameters are grouped and assigned to different participants who carry out parallel initiation)

The second type of Gaussian plume model setting is monopolization, which means that each participant will monopolize all of the model parameters, as shown in Figure 7.15, and does their own setting. During this process, each participant cannot affect the others' values. After all participants finish their settings, the values will be submitted to a central conflict detection module that is installed in the management node. This module will detect any conflicts and report them back to all participants. In Figure 7.15, three conflicts are detected, which are wind speed, wind direction and environment temperature. After negotiation through online chat (text, video and audio) the final values are set.

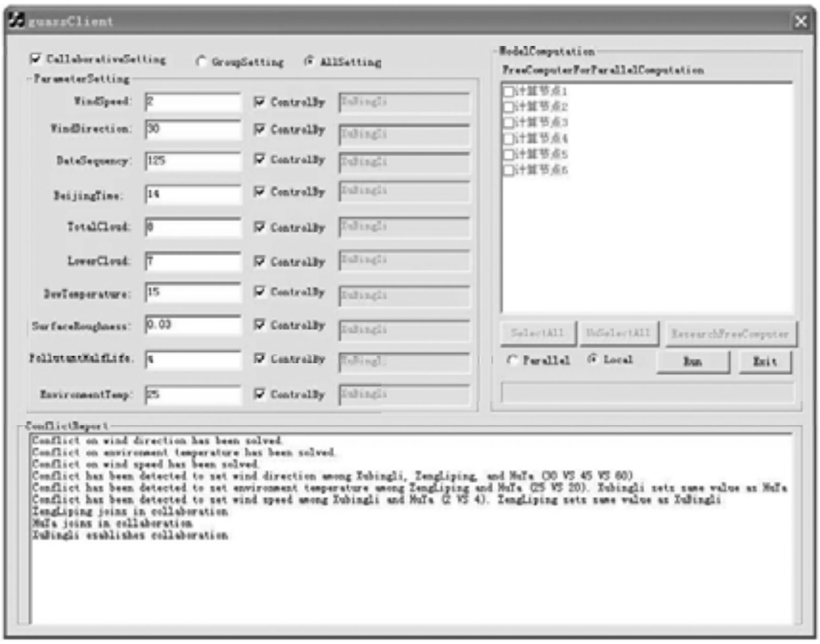


Figure 7.15: Gaussian plume model parameters under a monopolized setting (The parameters are grouped and assigned to different participants who carry out parallel initiation)

(3) Visualization collaboration

Once a participant establishes a visualization collaboration, whether in a 2D or 3D environment, he has the right to control the viewpoint (virtual camera), and can be viewed as the guide. In this situation, the others can follow the guide to see what he is doing. The operations of visualization collaboration are listed in Table 7.3.

Table 7.3: Operations of visualization collaboration

Collaboration operation	2D environment action	3D environment action
Move viewpoint down to terrain surface in vertical height	Map zoom in	Reduce virtual camera height
Move viewpoint far to terrain surface in vertical height	Map zoom out	Increase virtual camera height
Move viewpoint in pan	Map pan	Roam at a fixed height

Besides viewpoint collaboration, visualization content collaboration is implemented by realizing pollution layer distribution, which is shown in Figure 7.16.

In the figure, a controlling participant with a 2D environment set the pollution layer at a certain point in time. The operation is presented in the three scenes above including both 2D and 3D environments.

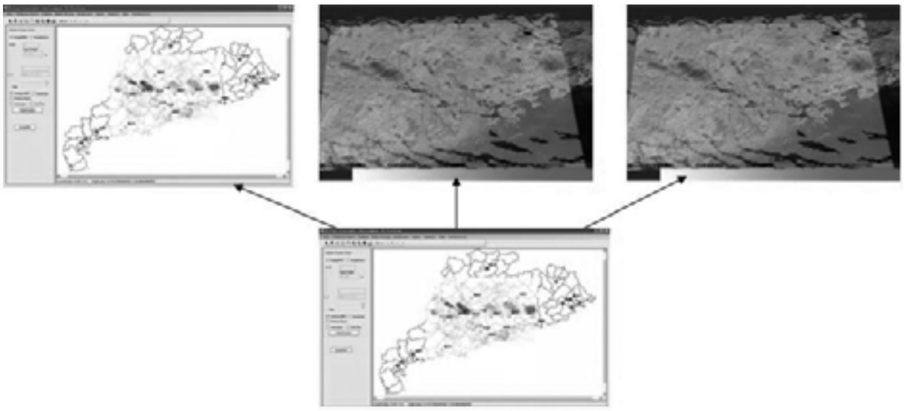


Figure 7.16: 2D environment control visualization collaboration

(4) Analysis collaboration

A typical analysis collaboration is shown in Figure 7.17. In this figure, there are four participants (A, B, C, D) who have joined to form a collaboration group. Each participant does his own analysis, including vertical pollution distribution on a point, and undertakes pollution analysis. The final analysis is the combination of the analysis collected from all participants.

Figure 7.17: Analysis collaboration

(5) Online chat as a tool for negotiation

Chat style negotiation is a tool which is attached to the whole system. When collaborating, participants may launch both a collaboration system and a negotiation tool, as shown in Figure 7.18. During collaboration, if a conflict occurs, the negotiation tool can be used to discuss the conflict and means to solve it. On the negotiation tool, two videos are shown. One is the local video, and the other is of the remote participants who make a different setting as local participants. The negotiation tool supports not only video, but also audio and text communications.

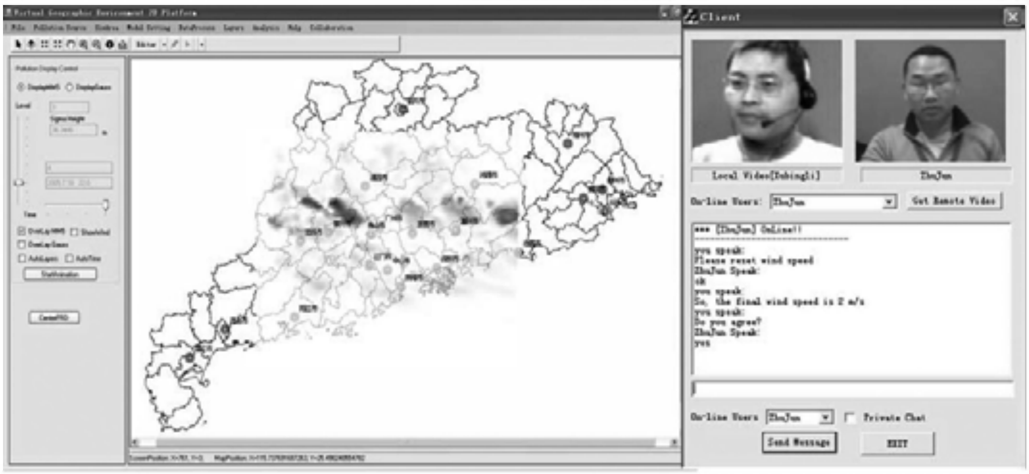


Figure 7.18: Chat style negotiation for participants

On the management node, the chat negotiation tool is installed to monitor participants, their state and operations.

Figure 7.19: Central monitor for chat negotiation

CHAPTER 8: PROTOTYPE BASED AIR POLLUTION SIMULATION AND DISCUSSION

Two objectives are focused on in this chapter. One is to respond to the questions about the status of air pollution simulation highlighted in the first chapter. The other is to test the feasibility and efficiency of the prototype system of CVGE that was established with the previous four chapters' technologies by simulating typical scenarios.

In responding to the questions highlighted in the first chapter, two scenarios will be simulated. One is the MM5 based air pollution episode simulation for the Pearl River Delta (PRD). In this episode, there are five steps taken for air pollution simulation: preparing air pollution sources, setting and computation MM5, visualization, analysis and model evaluation. Each of these is conducted step-by-step using the newly developed methodologies. At the same time, each step is accompanied by geo-collaboration. In this episode simulation, the geo-collaboration is a collaborative workflow, and there is no conflict. The second scenario is a simulation using the Gaussian plume model (GPM) of hypothetical accident (a ship wreckage) near Hong Kong Island. This simulation also has five steps. However, the geo-collaboration is different from the first simulation in that there are conflicts. Consequently, negotiation is applied to solve these conflicts.

The second objective, which is to test the feasibility and efficiency of the prototype system, is also met by in these two scenario simulations. The architecture, high performance computation based MM5 and GPM integration and computation, geo-visualization of air pollution modelling, analysis and geo-collaboration are all tested in these two simulations.

This chapter will first conduct an air pollution episode simulation for PRD based on MM5. Secondly, the Gaussian plume model based hypothetical scenario simulation will be presented. After these two simulations, the third task will be to discuss whether the efforts answer the questions posed in the first chapter. The merits and deficiencies of the CVGE prototype for air pollution simulation are also included in the third task.

8.1 MM5 Based Air Pollution Simulation for Pearl River Delta

8.1.1 Air pollution in PRD

The Pearl River Delta (PRD) occupies the low-lying areas alongside Southern China's Pearl River Estuary, where the Pearl River flows into the South China Sea. The PRD covers nine prefectures of Guangdong Province, namely Guangzhou, Shenzhen, Zhuhai, Dongguan, Zhongshan, Foshan, Huizhou, Jiangmen and Zhaoqing. The Greater Pearl River Delta (GPRD) covers the PRD and the Special Administration Regions (SAR) of Hong Kong and Macao. In order to avoid unnecessary complication of the narrative, the PRD and the GPRD are not distinguished. Thus, the PRD and the GPRD are analyzed as the same region.

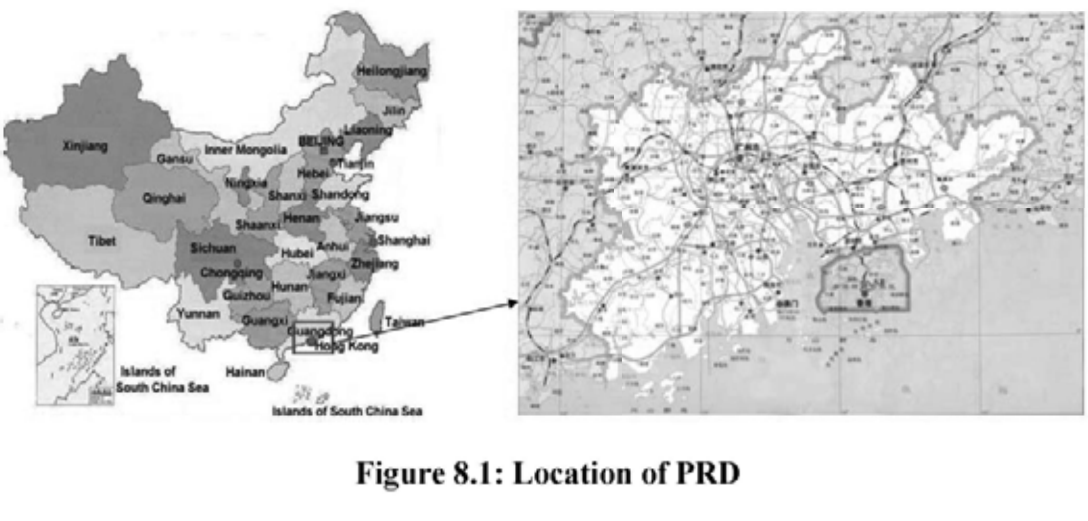


Figure 8.1: Location of PRD

The PRD is chosen to be the pioneer of reform when China's economic liberalization program began in the late 1970s. The delta area in Guangdong

Providence has become one of the leading economic regions and major manufacturing centers of the mainland. The region's Gross Domestic Product (GDP) grew from just over US$8 billion in 1980 to more than US$89 billion in 2000 and 1805.9 billion in 2005. During that period, the average actual rate of GDP growth in the Pearl River Delta Economic Zone was almost double of the total national GDP of the People's Republic of China (PRC) (www.stats.gov.cn/tjfx/fxbg/index.htm, access date 2009-4- 16).

In tandem with the fast economic growth rate, the region's population has rapidly expanded rapidly. In 2000, the PRD's floating population was approximately 21.61 million. According to the 2005 national census, it had a permanent population of 43 million and floating population of 24.8 million. Among these cities, Guangzhou and Shenzhen have almost reached their capacity.

The economic development and population expansion of the PRD economy bring negative impacts on this region's environment. More and more agricultural land is given over to the needs of urban expansion. Industrial emissions and effluent pollute soil, water and air. The problem of air pollution in the PRD is increasing, as evidenced by the following: The number of hazy days is increasing year by year (table1.1). On 27th August 2006, it was reported by the BBC, VOA and other media that "Hong Kong's polluted air is driving away foreign professionals and threatening international investment, according to a survey released by the American Chamber of Commerce in Hong Kong"(AmCham 2007). Time Magazine also reported the declining air quality in Hong Kong, which has become a threat to residents' health (Bryan 2006). In June 2008, Professor Nanshan Zhong, a fellow of the Chinese Academy of Engineering researching into public health, reported an increasing number of non-smokers in Guangzhou with pneumoconiosis (or "black lung") requiring medical and surgical intervention. He added the expiation that the air pollution was responsible. To summarize, air pollution has impacted negatively on the local economy and residents' health. The problem of air pollution must be resolved.

Table 8.1: Total hazy days in main cities of PRD from 2001 to 2005(Wen 2006)

City	2001	2002	2003	2004	2005	Mean
Guangzhou	56	65	87	142	132	96
Fushan	174	123	159	170	130	151
Dongguan	32	35	154	190	165	115
Zhaoqing	88	63	85	91	86	83
Zhongshan	26	37	26	17	14	24
Huizhou	3	1	6	5	19	7

How to control air pollution to improve this region's air quality has catch attentions of experts, government officials and publics. Currently, scientific researches and governmental managements on air pollution in PRD are both launched. The scientific researches reveal that sulphur dioxide (SO_2) is one of the main pollutants in this region (Wang et al. 2005b). Thus, the following simulation will focus on SO_2.

8.1.2 Episode

Tropical Cyclone (TC) is one of the typical climate types in PRD. TC and atmospheric aerosol events are sometimes related, because it could change its surrounding meteorological conditions, and thus influence local or regional air quality(Wu et al. 2005, Feng et al. 2007). Typhoon HaiTang, a damaging TC, can serve as a typical case. From 18th July 2005 to 20th July 2005, HaiTang passed Taiwan Strait and landed at Fuzhou of mainland of China, which also caused thousands of million economy damage on its path. Around HaiTang pathway, air quality was significantly influenced. During that period, PRD underwent bad air quality. For instance, in ShenZhen city, three continuous days were dominated by severe hazy from 19th to 21th in July 2005, which made visibility less than three kilometers.

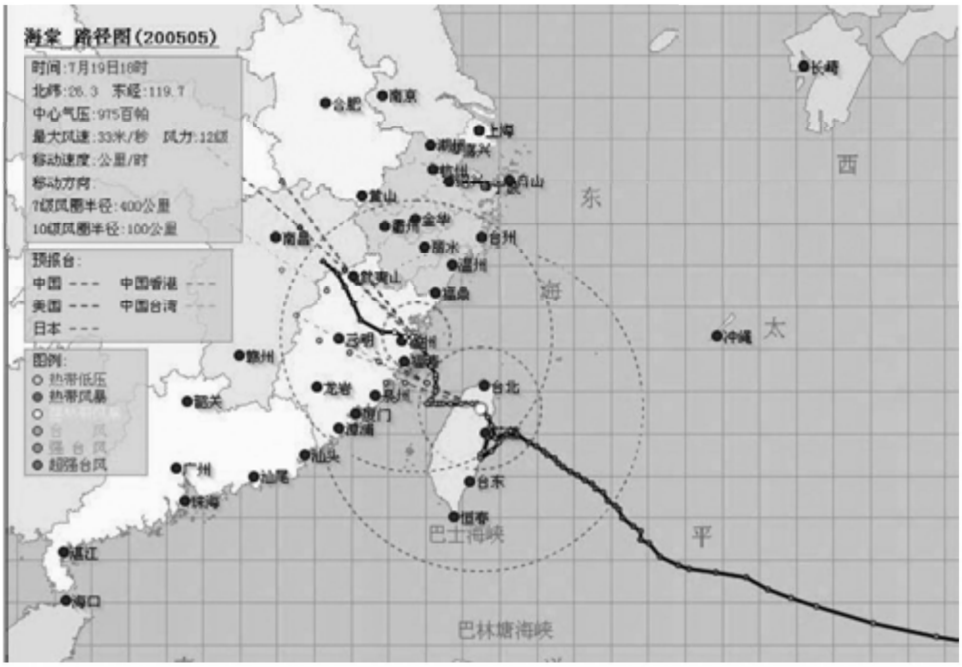


Figure 8.2: Pathway of typhoon HaiTang (http://slt.zj.gov.cn/typhoneweb/, access May 2009)

The period of the episode to be simulated is selected to covers from 2 o'clock on 19th July 2005 to 1 o'clock on 20th July 2005, which is part of the whole period of HaiTang lifetime. The atmosphere environment within the selected period is used as the environmental background.

8.1.3 Episode simulation based on prototype of CVGE

This simulation engages multiple participants who are a governmental official (GO), a geographer (G), an environmental scientist (ES), a computer technician (CT). The simulation based on CVGE prototype will be conducted step by step. The simulation process is based on the design of Figure 8.3. The model computation and geo-visualization use the achievements of chapter 6 and chapter 7.

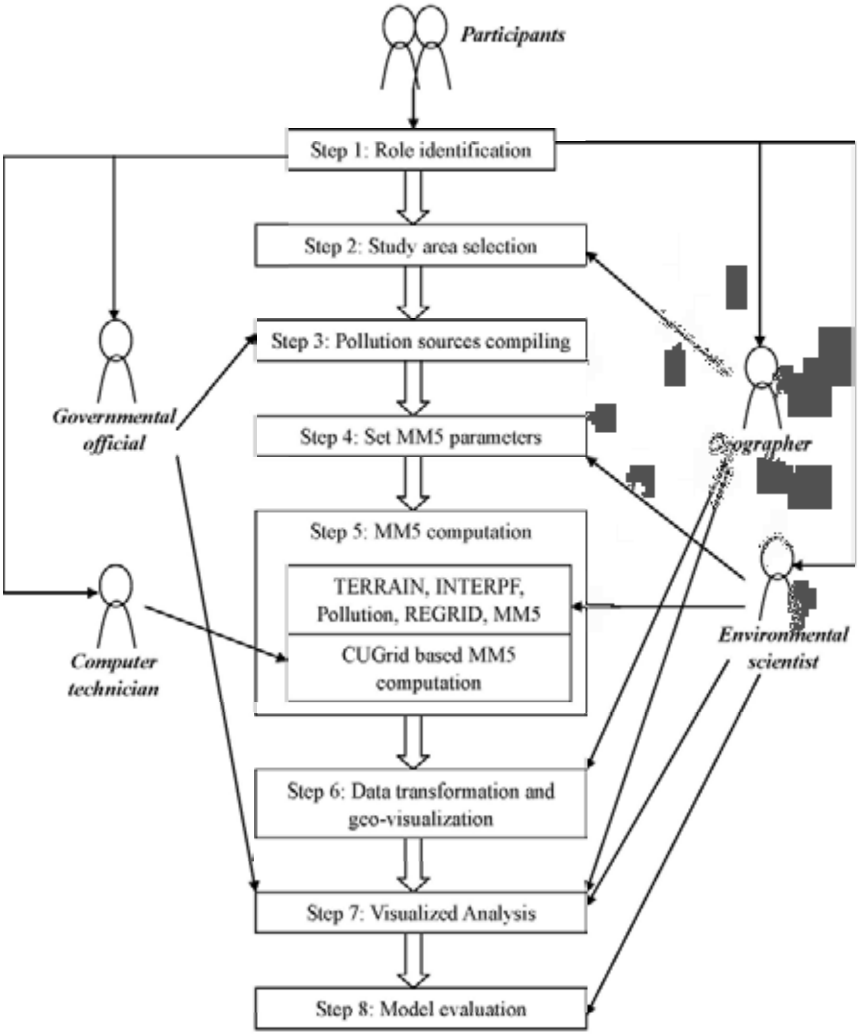


Figure 8.3: The collaborative air pollution simulation workflow based on MM5

Step 1: roles identification

Multiple participants from geographically dispersed places start the CVGE platform. The dialogue (Figure 8.4) of roles identification will first popped up. On this dialogue, participant must input his name and passwords, identify his role from selecting from drop-list, and determine his preferring interface. Once the role identified, a participant can only conduct authorized operations which are listed in table 7.1. At the same time, the online chat room (as Figure 7.8 left) is also launched when a participant enters into the system.

Figure 8.4: System login dialogue

Step 2: Select simulated region

Through online chat room, the governmental official asks the geographer to set the research area in a rectangle coving the whole PRD region. With the 2D environment, the geographer easily draws the required area as Figure 8.5 (left). After drawing, the detail geographic locations of low bottom corner and top right corner are restricted with high precision as (111.4266, 21.3241) and (115.5791, 24.3480) respectively. The geographer continues to extract geo-information to support air pollution modeling (as Figure 8.5 right).

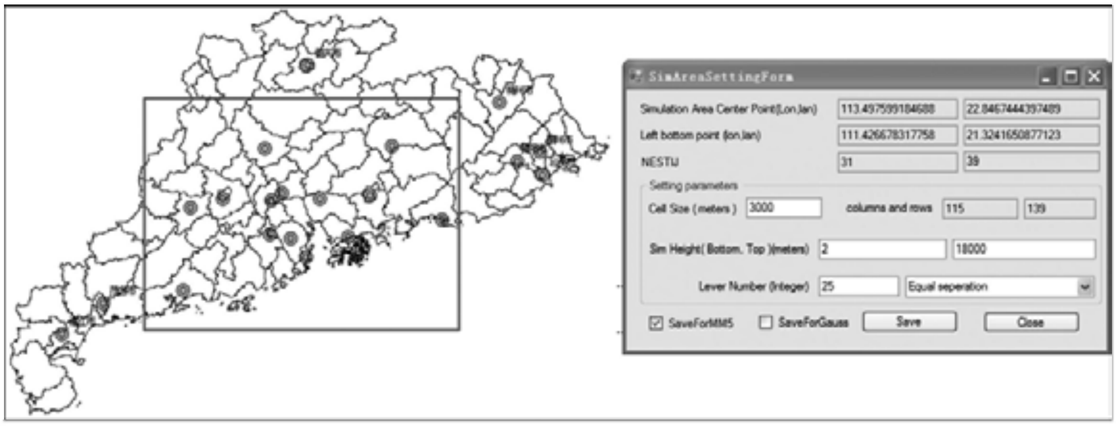


Figure 8.5: Simulation area selection and geo-information extraction conducted by geographer

The operation of drawing rectangle can be viewed synchronously in other participants' interface. After all the operations, the geographer sends out message to tell other participants that he has finished his work.

Step 3: compiling air pollution sources

When getting the message from geographer that he has finished the selection of study area, the governmental official start compiling air pollution sources. With the visualized interface of Figure 6.10, governmental official open all the pollution sources in PRD and save the setting for next step operation. After compiling, the governmental official tells other participants by text message or audio/video that he has finished his work.

Step 4: Setting MM5 parameters

Receiving the message of finishing pollution sources compiling, the environmental scientist starts his work to set the parameters of MM5.

The study area is automatically set according to step 2. The resolution of MM5 computation and the grid numbers in column and row are also automatically set according to step 2, whose values are 3000 meters, 139, and 115 respectively.

Table 8.2: Simulated layers' sigma heights

NO.	Sigma Height (m)
1	4.03631
2	12.1041
3	36.3446
4	76.8735
5	113.487
6	146.143
7	244.74
8	432.217
9	687.682
10	1016.34

The environmental scientist sets the simulation time period as 24 hours from 2 o'clock on 19th July 2005 to 1 o'clock on 20th July 2005. He also sets the computation time step, the output data saving time step, and the output time step with 60 minutes. 60 seconds, and 60 minutes respectively. In vertical direction, only the layers under 1 kilometer height are calculated because these layers are closely related with human activity. Thus, the environmental scientist set sigma heights as those listed in table 8.2.

Step 5: MM5 computation

Following step 4, this step will firstly preprocess modules of MM5 including TERRAIN, REGRID, INTERPF, Pollution, and MM5. The processing on these modules is conducted by the environmental scientist. The operation interfaces are as Figures 5.21, 5.22, 5.23, 5.24, and 5.25. Among these modules, the Pollution module is used to transform the compiled air pollution sources conducted by step 3 into a Fortran program file. Finishing running these modules, the environmental scientist sent out notification message.

When notified that the environmental scientist has finished running previous models, the computer technician will conduct running MM5 on CUGrid. The interface to support this operation is Figure 5.26.

Step 6: Data transformation and geo-visualization for results calculated by MM5

This step is conducted by the geographer. He will firstly transform the simulated result of air pollution into geographically referenced data which can be read by CVGE. The operation interface for data transformation is Figure 6.6. According to MM5 setting, the output data have ten layers and 24 time sequences. Thus, after transformation, there are 240 layer and time files for air pollution representation. Each layer will has 115*139 grids with the resolution of 3000 meters. After the data transformation, the geographer visualizes air pollution dispersion in 2D and 3D environments.

Before doing the displaying, pollution color will be designed to match the custom of the governmental official and the environmental scientist.

(1) Color-Bar design

The government official supplies the geographer a table (as table 8.3) of Air Pollution Index (API) defined by Environment Protection Department (EPD) of Hong Kong government. According to table 8.3, EPD sets five levels of API, which are Low, Medium, High, Very High, Severe, with the API regions of 0-25, 26-50, 51-100, 101-200, and 201-500 respectively.

Table 8.3: API Subindex Levels and Their Corresponding Concentrations (http://www.epd-asg.gov.hk/english/backgd/table122.php, access May 2009)

API Subindex Level	Corresponding Concentrations (μg/m^3)							
	RSP 24-hr	SO_2 24-hr	SO_2 1-hr	NO_2 24-hr	NO_2 1-hr	CO 8-hr	CO 1-hr	O_3 1-hr
0	0	0	0	0	0	0	0	0
25	28	40	200	40	75	2500	7500	60
50	55	80	400	80	150	5000	15000	120
100	180	350	800	150	300	10000	30000	240
200	350	800	1600	280	1130	17000	60000	400
300	420	1600	2400	565	2260	34000	90000	800
400	500	2100	3200	750	3000	46000	120000	1000
500	600	2620	4000	940	3750	57000	150000	1200

According API and five levels of air quality, color bar of sulphur dioxide concentration is designed by geographer as Figure 8.6. Pollution color is separated into five segments. Each segment is stretched into color region. The first segment presents Level Ⅰ with sulphur dioxide concentrations below 0.2mg/m^3. Following three pollution color segments are Level Ⅱ, Level Ⅲ, and Level Ⅳ with threshold pollution value of 0.4mg/m^3, 0.8mg/m^3, and 1.6mg/m^3 respectively. Left segments is Level Ⅴ to present sulphur dioxide concentration higher than 1.6mg/m^3.

Figure 8.6: Color bar design for sulphur dioxide concentration distribution (The unit is mg/ m3)

(2) Geographic visualization of sulphur dioxide dispersion in 2D environment

Stationary or animated geo-visualizations are both conducted. When the geographer visualizes concentration of SO_2, the visualized graphics can also be viewed by other participants.

Following series pictures (Figure 8.8) show SO_2 dispersion on layer 2 with sigma height 12 meters above terrain. These pictures share the same color bar which is shown in Figure 8.7.

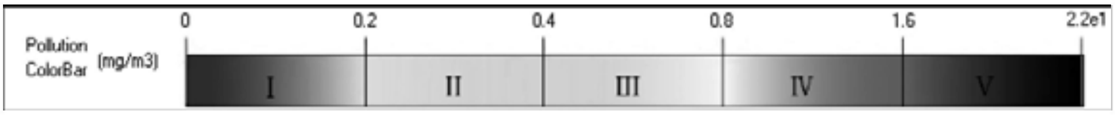
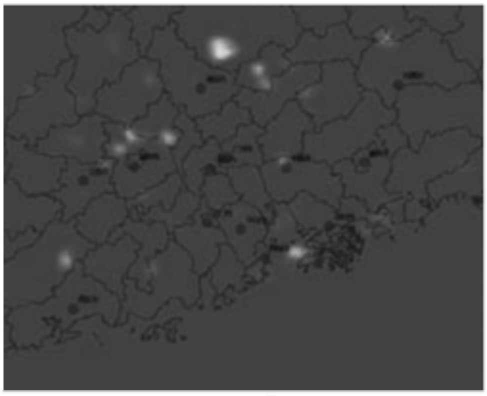
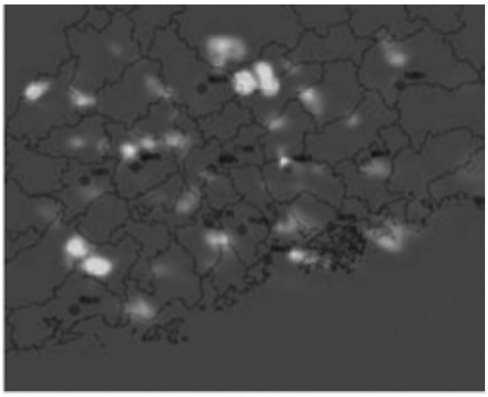


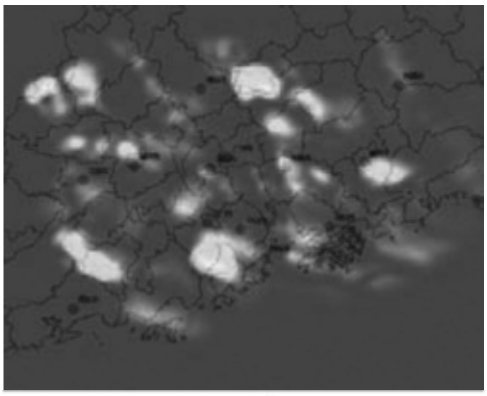
Figure 8.7: Pollution ColorBar for air pollution visualization with all air pollution sources opened



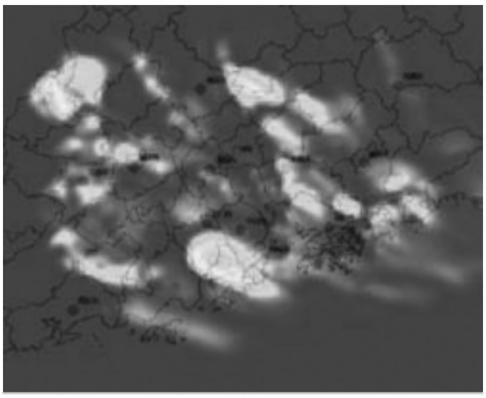
2:00 19th Jul 2005



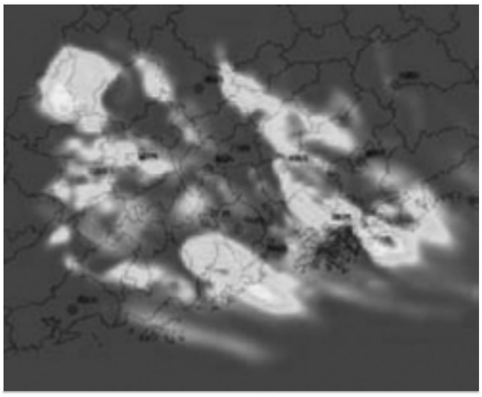
3:00 19th Jul



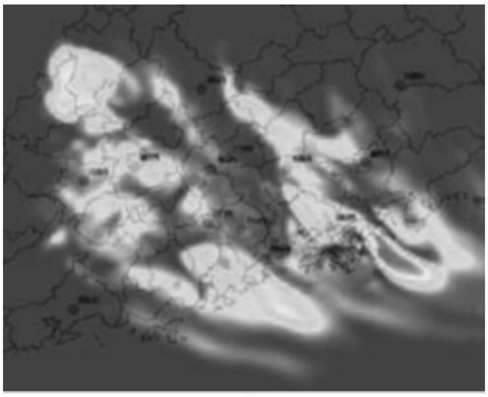
4:00 19th Jul



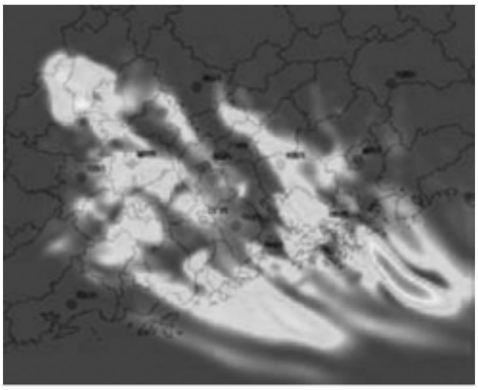
5:00 19th Jul



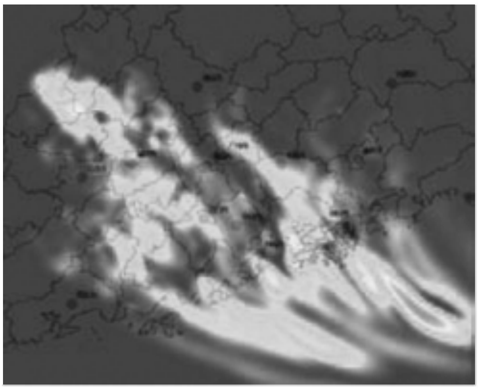
6:00 19th Jul



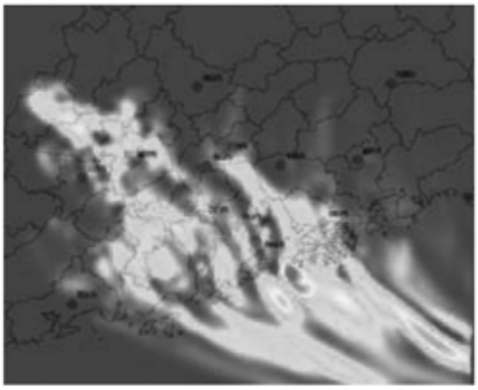
7:00 19th Jul



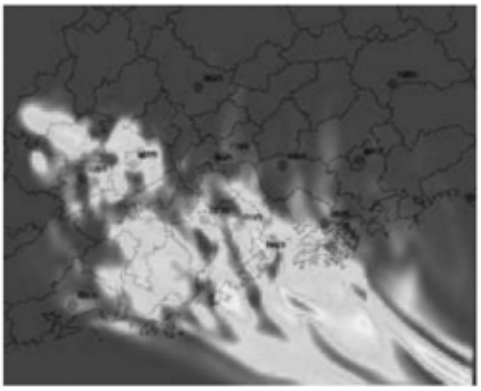
8:00 19th Jul



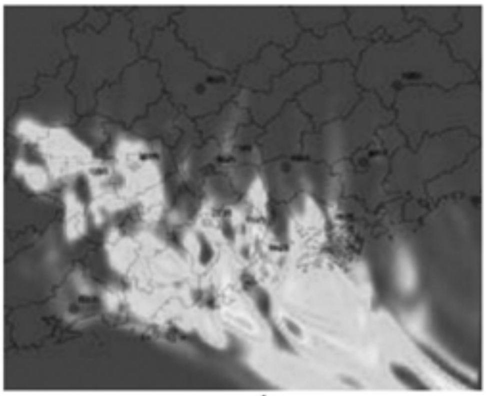
9:00 19th Jul



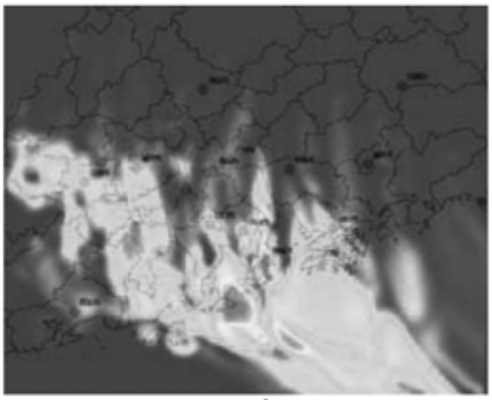
10:00 19th Jul



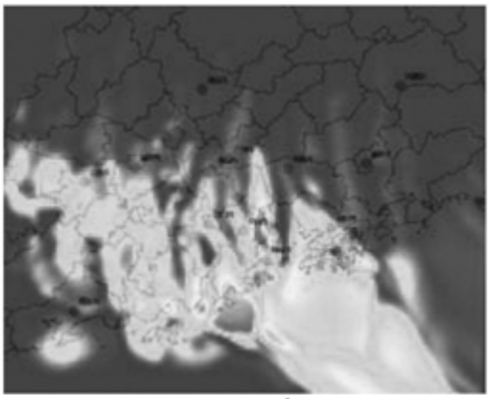
11:00 19th Jul

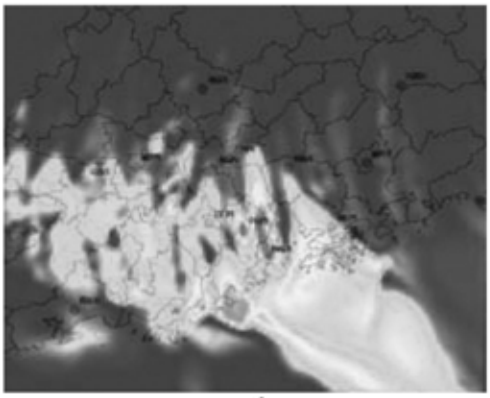


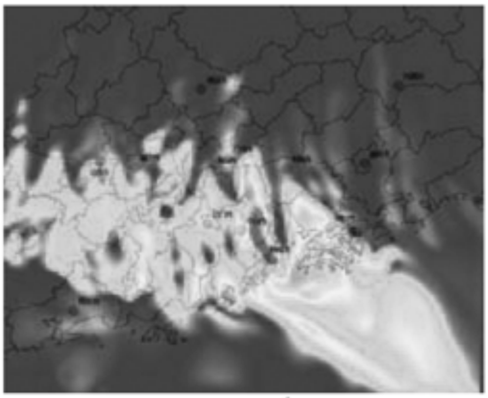
12:00 19th Jul

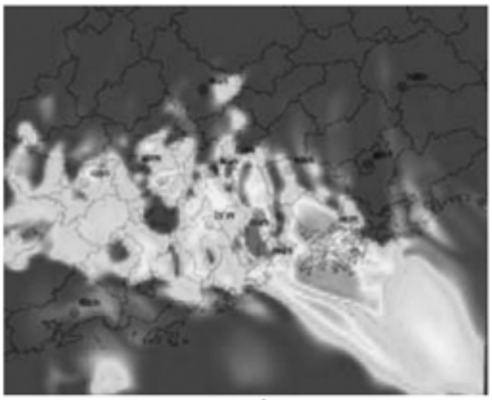


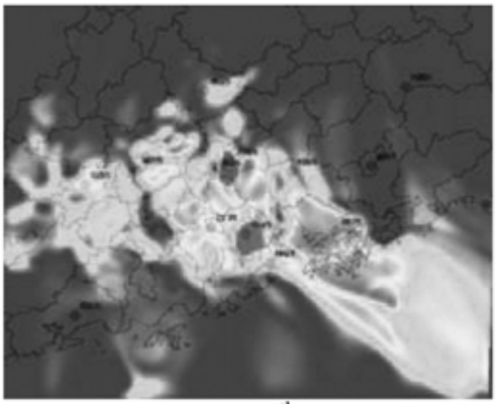
13:00 19th Jul

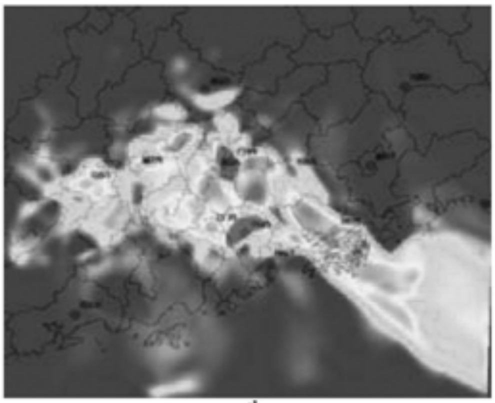

14:00 19th Jul

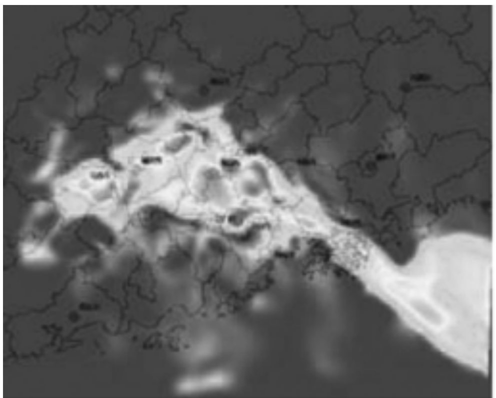

15:00 19th Jul

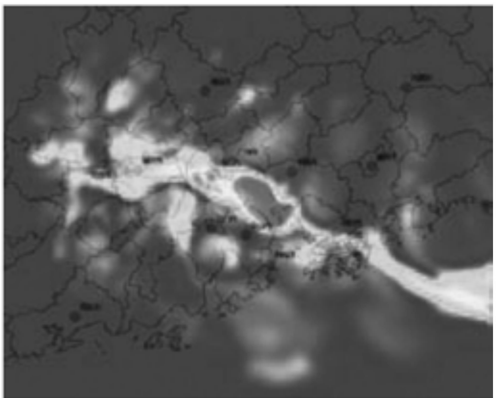

16:00 19th Jul

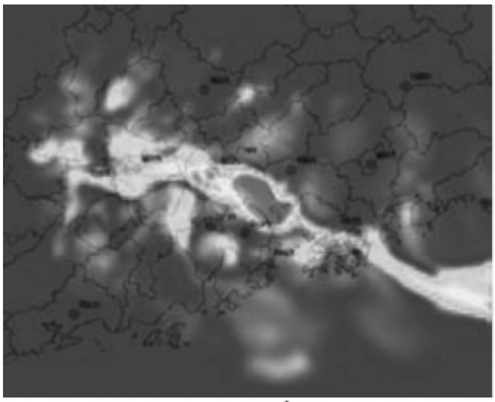

17:00 19th Jul


18:00 19th Jul

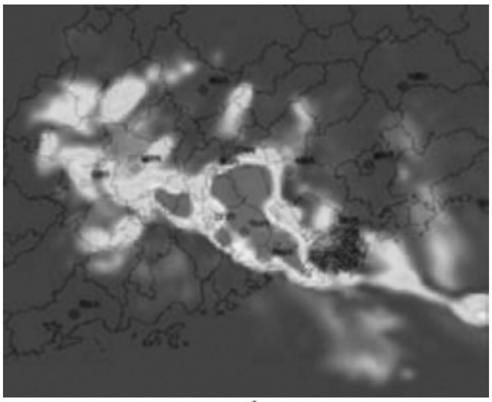

19:00 19th Jul


20:00 19th Jul

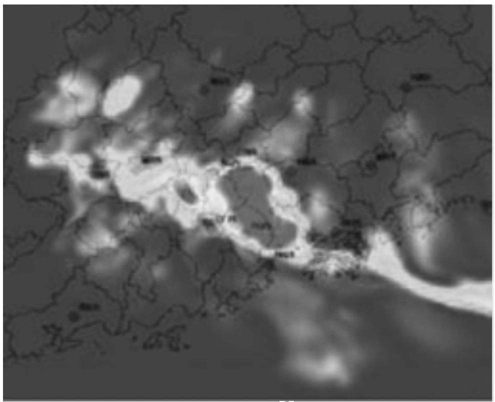

21:00 19th Jul



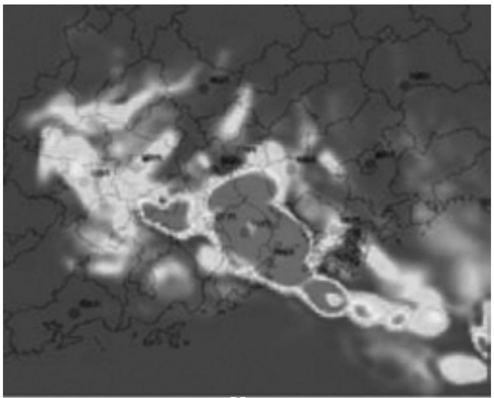
22:00 19th Jul



23:00 19th Jul



24:00 19th Jul



1:00 20th Jul

Figure 8.8: Simulated sulfur dioxide concentration dispersion from 2:00 19th July 2005 to 1:00 20th July 2005 on layer with sigma height of 12 meters above terrain.

From above simulation, the conclusions can be made for layer at sigma height of 12 meters above terrain as following:

(1) The air pollution cased by industries emission in PRD dispersed from north west to south east driven by wind field;

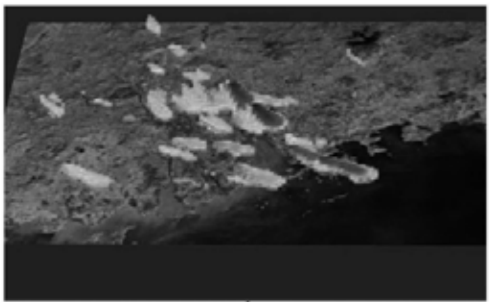
(2) In the simulated condition, four cities had severe air pollution level for several hours during simulation time, which included ZhuHai, JiangMen, ZhongShan, and FoShan. Parts of GuangZhou, DongGuan, ShenZhen were also polluted seriously;

(3) During the simulation time, Hong Kong was covered by very high and even severe pollution levels for about four hours from 15:00 o'clock to 19:00 o'clock on 19th July 2005.

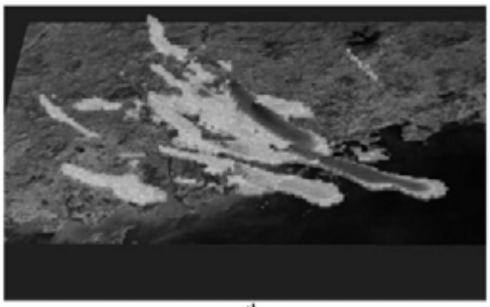
(3) Geographic visualization of sulphur dioxide dispersion in 3D environment

In chapter 6, participles system and pollution boxes are designed to present air pollutant which is regarded as fuzzy boundary volume object. Because of consuming so much computer resource, following visualization only applies pollution boxes to geographically visualize dispersion of sulphur dioxide. With the time going on, the dispersion of sulphur dioxide can be visualized as Figure 8.10. Following visualization shares color-bar of Figure 8.7.

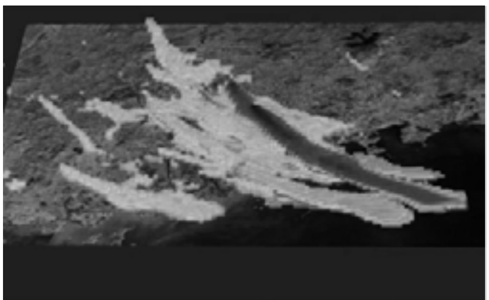
Just like geo-visualization in 2D environment, the 3D geo-visualization conducted by geographer can also be viewed by other participants.



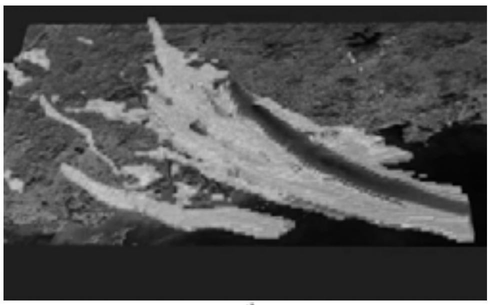
2:00 19^{th} Jul 2005



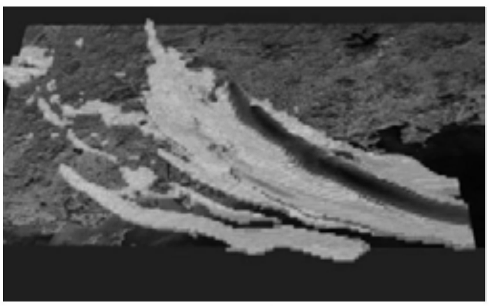
3:00 19^{th} Jul 2005



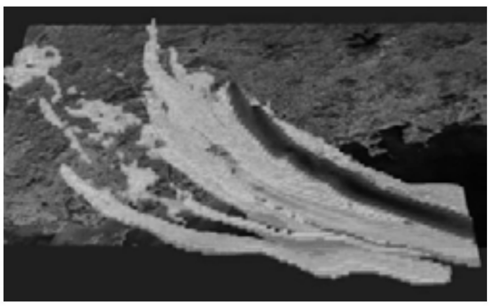
4:00 19^{th} Jul 2005



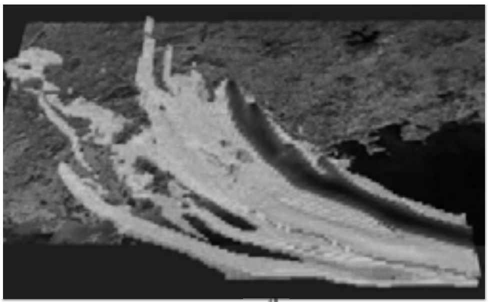
5:00 19^{th} Jul 2005

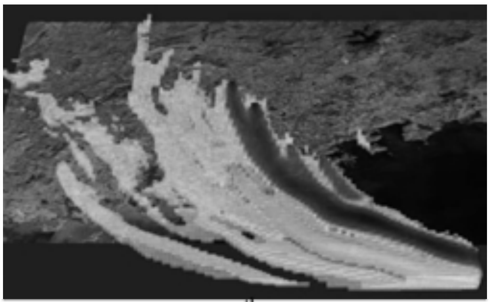


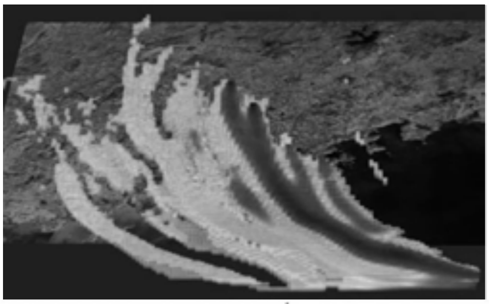
6:00 19^{th} Jul 2005

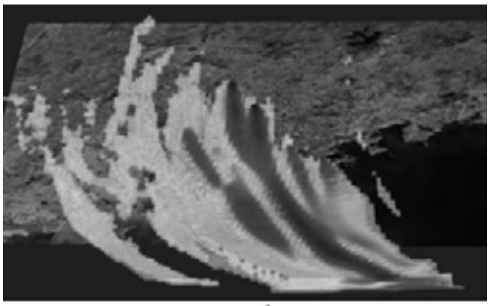


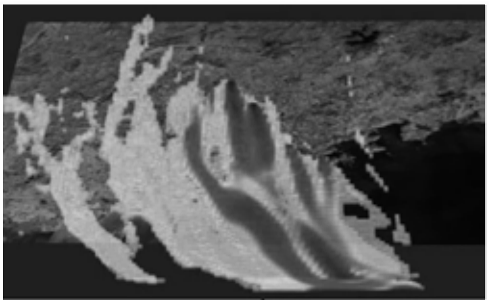
7:00 19^{th} Jul 2005

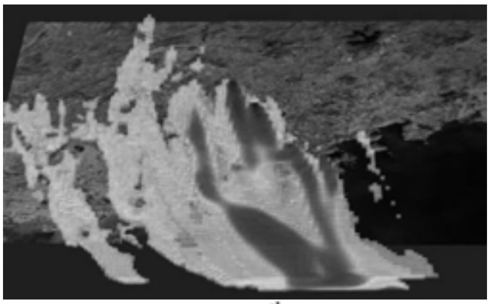

8:00 19th Jul 2005

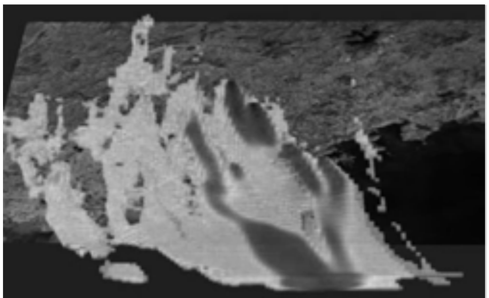

9:00 19th Jul 2005

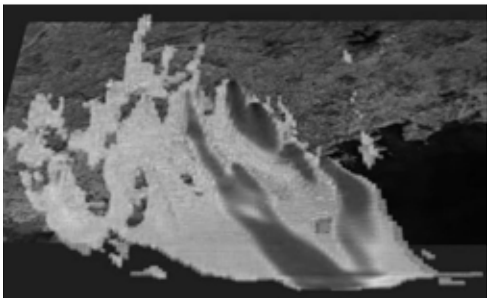

10:00 19th Jul

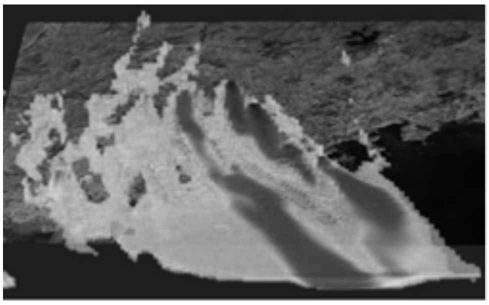

11:00 19th Jul

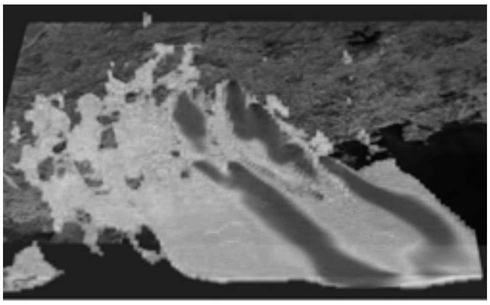

12:00 19th Jul

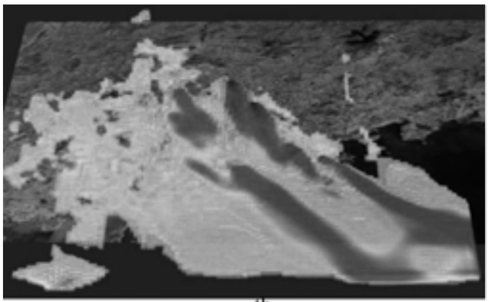

13:00 19th Jul


14:00 19th Jul

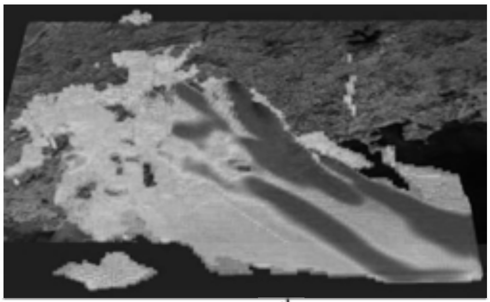

15:00 19th Jul


16:00 19th Jul

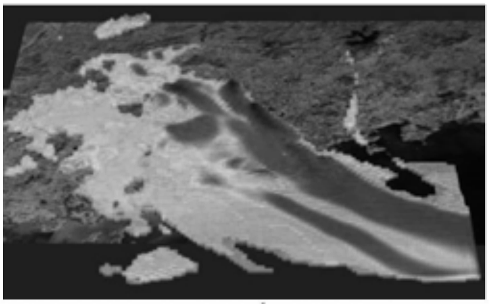

17:00 19th Jul



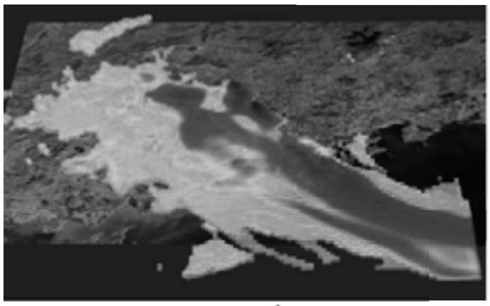
18:00 19th Jul



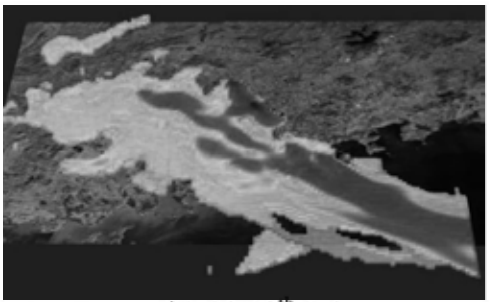
19:00 19th Jul



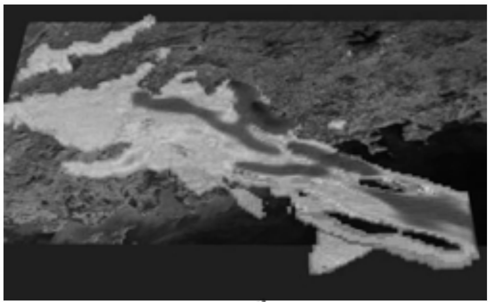
20:00 19th Jul



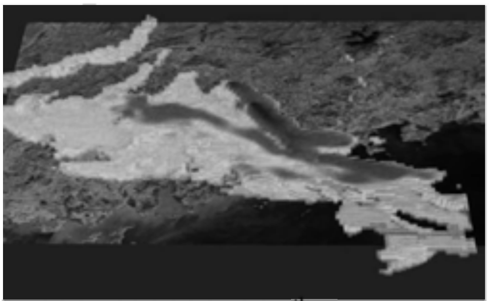
21:00 19th Jul



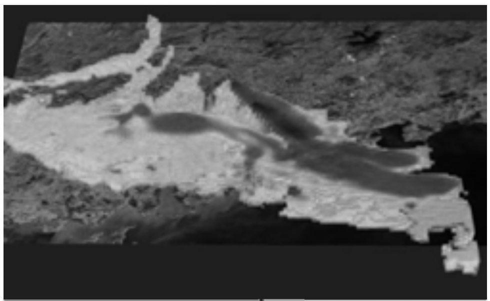
22:00 19th Jul



23:00 19th Jul



24:00 19th Jul



1:00 20th Jul 2005

Figure 8.9: Volume visualization of sulphur dioxide dispersion in 3D environment

With 2D and 3D environments, the geographer can zoom in/out, pan the map, , change view point, roam in 3D environment, and do many other operations. All of the operation effect can also be aware by other participants, whose environments will be automatically updated to keep synchronization with the geographer's.

Step 7: Collaborative analysis

Three roles collaboratively conduct this step. The analysis functions are these implemented in chapter 6, which are point profile and transection. The visualization effects are as Figure 7.17.

Step 8: Air pollution dispersion model evaluation

This step is controlled by the environmental scientist.

The simulated results are compared with monitoring data to evaluate the precision of air pollution dispersion model (SYSUM). The simulated result is gotten with the function of point profile analysis in 2D and 3D virtual environments. The monitoring data are restrained in Hong Kong, and downloaded from website of Environment Protection Department of Hong Kong government. The data are collected at eleven general stations and three roadside stations (Figure 8.10) every one hour. From 2:00 19th July 2005 to 2:00 20th July 2005, there were nine stations collected intact data, but the other five stations lost some records. So, the nine stations with intact data are selected to evaluate simulation, which are general stations located at ShaTin, Central/Western, Eastern, KwaiChung, KwunTong, ShamShuipo, TsuenWan, TapMun and TungChung. The comparison diagrams for these nine monitor station are listed as Figure 8.11.

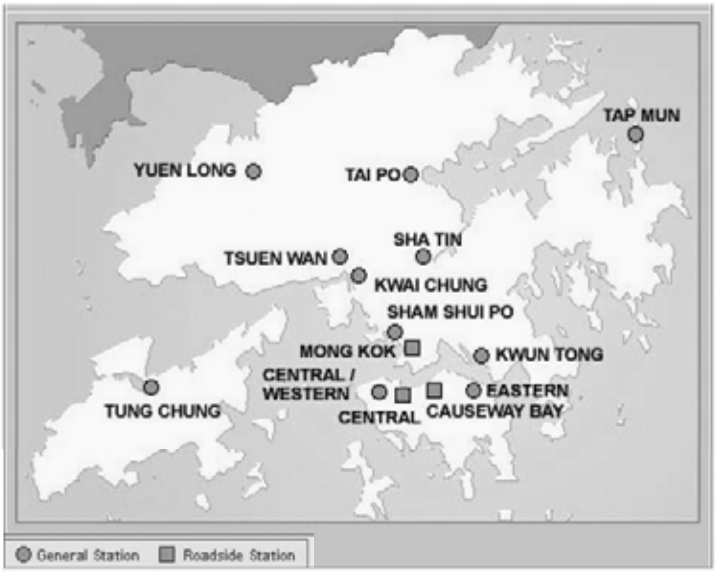
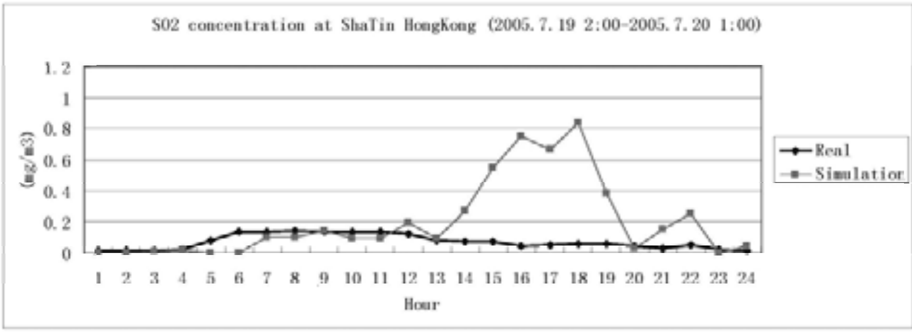
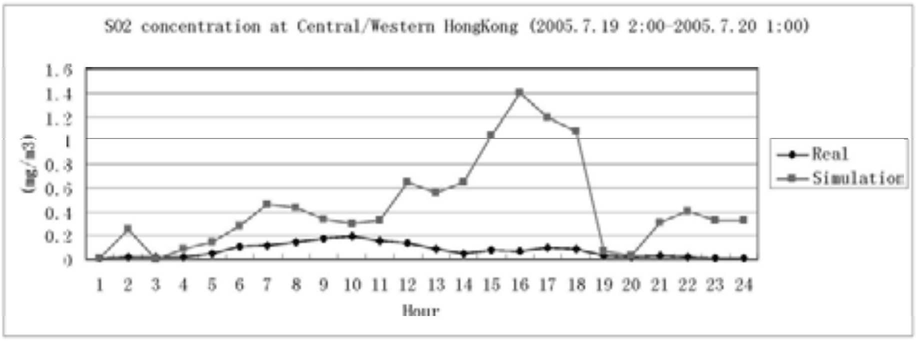
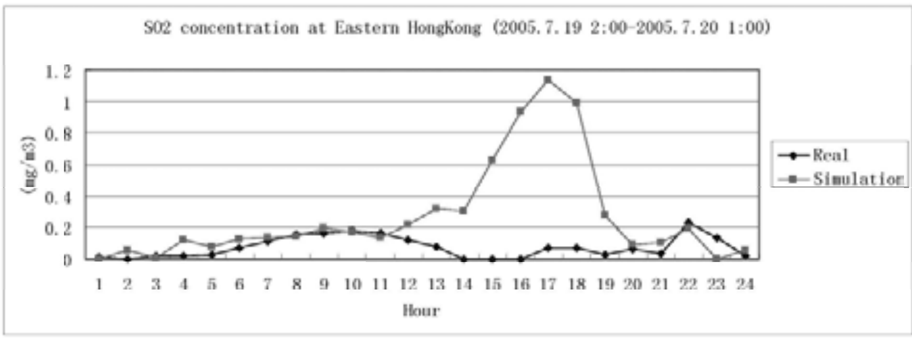
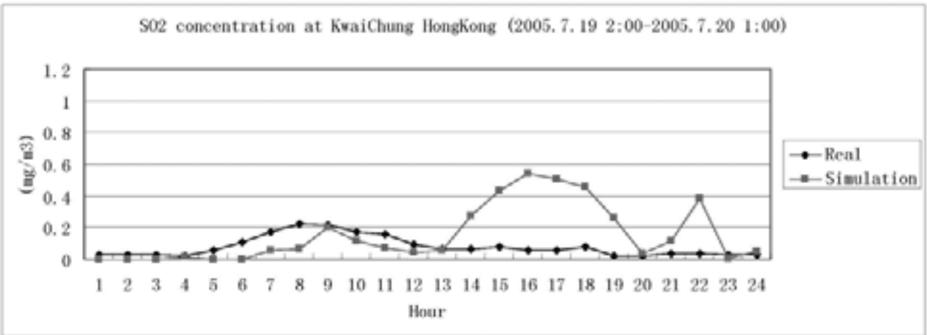


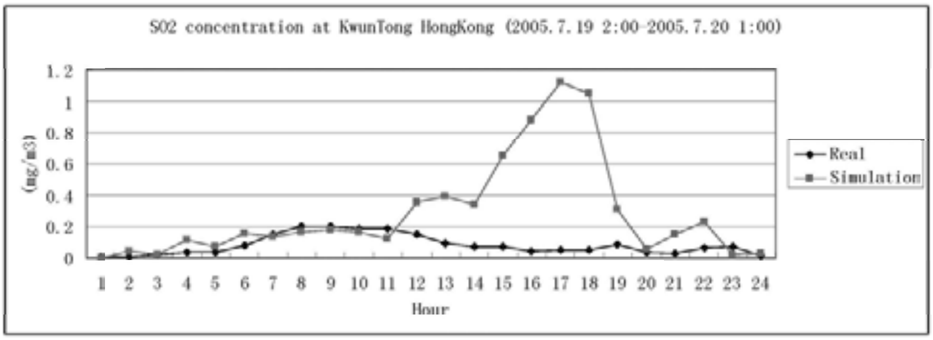
Figure 8.10: Air quality monitoring station in Hong Kong (www.epd-asg.gov.hk/english/backgd/quality.php, access May 2009)

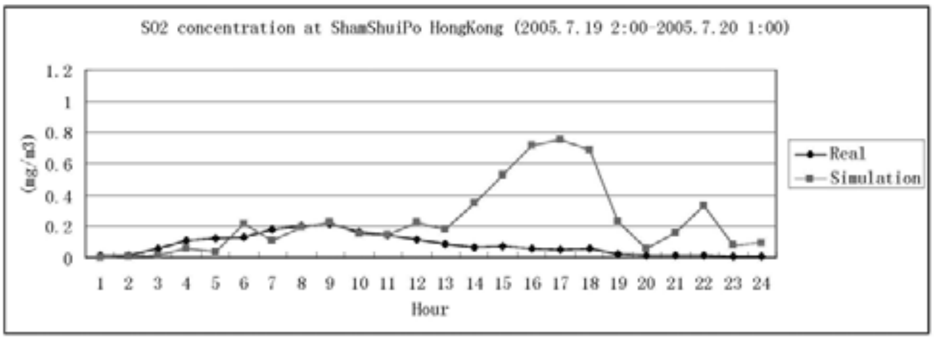


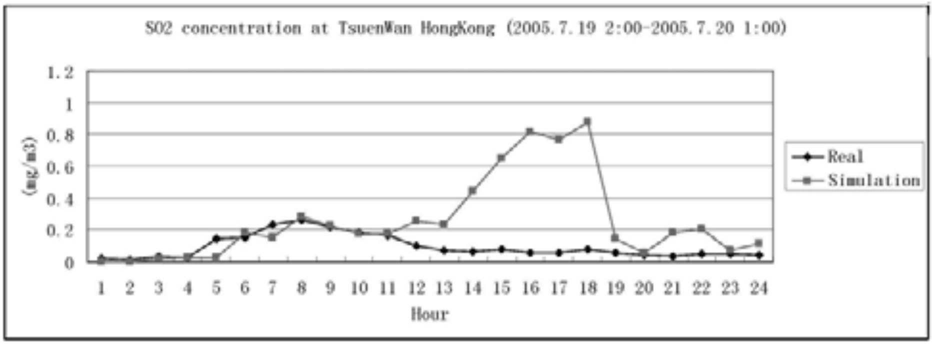


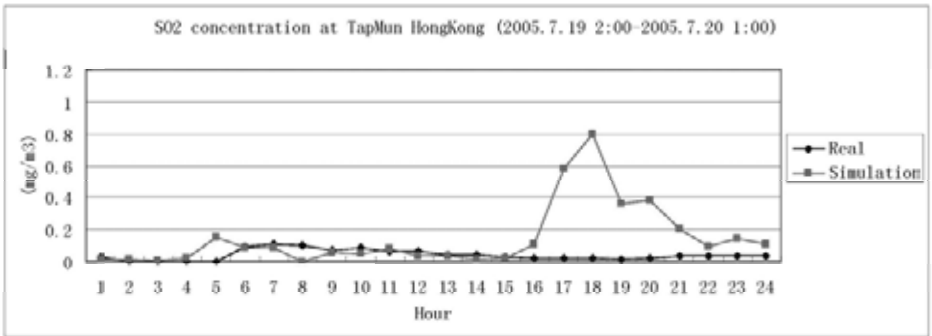



SO2 concentration at KwaiChung HongKong (2005.7.19 2:00-2005.7.20 1:00)
(mg/m3)
Hour
Real
Simulation


SO2 concentration at KwunTong HongKong (2005.7.19 2:00-2005.7.20 1:00)
(mg/m3)
Hour
Real
Simulation


SO2 concentration at ShamShuiPo HongKong (2005.7.19 2:00-2005.7.20 1:00)
(mg/m3)
Hour
Real
Simulation


SO2 concentration at TsuenWan HongKong (2005.7.19 2:00-2005.7.20 1:00)
(mg/m3)
Hour
Real
Simulation


SO2 concentration at TapMun HongKong (2005.7.19 2:00-2005.7.20 1:00)
(mg/m3)
Hour
Real
Simulation

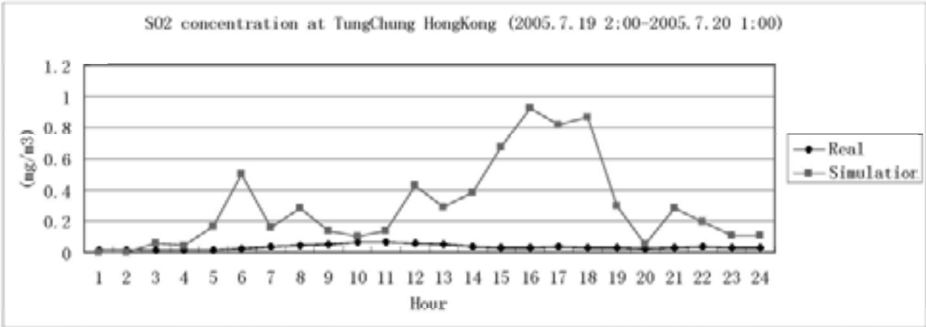


Figure 8.11: Evaluation of air pollution model by comparing simulated results with real monitoring data

From the comparison, it can be clearly found that the simulated results well matched with real monitoring data before 12:00. But after noon, the simulated results jumped to big values. When time came to 20:00 o'clock, the simulated value would come down again to meet monitoring value. So, the conclusion can be made as followings. In this simulation condition, the MM5 based air pollution dispersion model (SYSUM) can be used to simulate air pollution in PRD from morning to noon. But after noon, the simulated results are significantly higher than real conditions. But after 20:00, the simulated results would approach real values again.

8.2 Air Pollution Simulation Based on Gaussian Plume Model

8.2.1 Scenario definition

This scenario examines the hypothetic situation that a large ship wrecks near Hong Kong Island. The ship emits Particulate Matter with a diameter less than 10μm (PM_{10}) which is harmful to human health. Hong Kong government starts emergency response mechanism and organizes three organizations (Hong Kong Observatory—HKO, Environmental Protection Department—EPD, and the Institute of Space and Earth Information Science—ISEIS from the Chinese University of Hong Kong) to conduct a collaborative simulation. The simulation tries to find out the concentration distribution of PM_{10} in 3D space. The simulated results will be adopted by government to support decision of resident evacuation.

The participants are a geographer from ISEIS, an environmental scientist from EPD, and an atmospheric scientist from HKO. Everyone stays at their offices and connected together by the prototype CVGE. The locations of ship wreckage and the participants' offices are labeled in Figure 8.12.

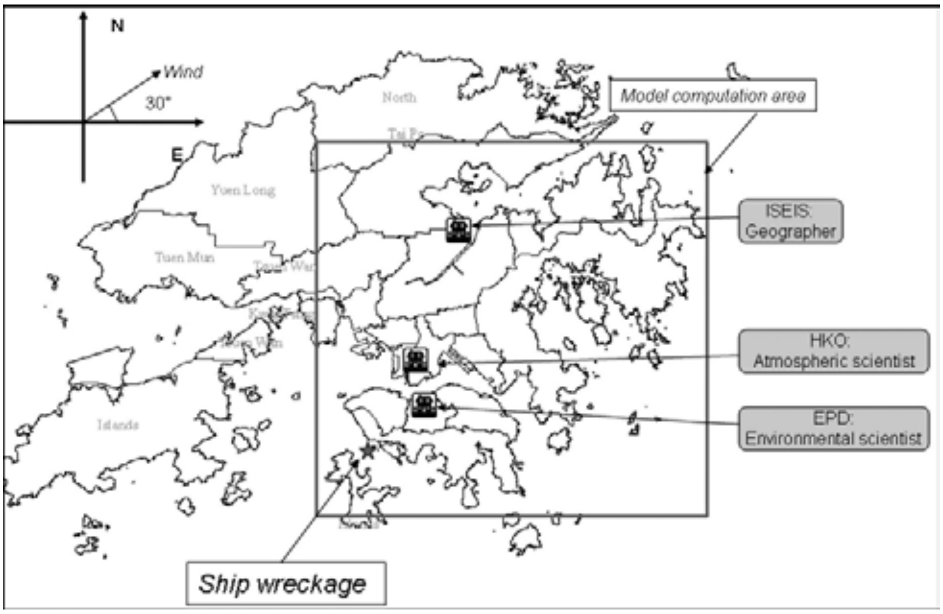


Figure 8.12: Locations of ship wreckage, participants' office, wind direction, and study area.

8.2.2 Scenario simulation based on Gaussian plume model

Step 1: System login and role identification

Participants from geographically dispersed locations login shared prototype CVGE by the interface as Figure 8.5. Participants from ISEIS, HKO and EPD are identified as geographer, atmospheric scientist, and environmental scientist respectively.

Step 2: Locating ship wreckage (pollution source compiling)

An on-the-spot staff surveys the location of ship wreckage, measures the emission intensity and emission height, and reports the three values to the office of EPD by telephone. With the help of pollution compiling function, the environmental scientist adds a new point source into the system. The point source has at least three attributes of emission location, emission height, and emission intensity. And the three attributes are assigned values according to on-the-spot report. The location is (114.1283°E, 22.25°N). The emission intensity of PM_{10} is 53kg/min. The emission height is 2 meters above sea surface. The pollution source is saved in a file which will be read by program of Gaussian plum model.

The added point source can be synchronously viewed by other two participants.

Step 3: Gaussian plume model collaborative setting and computation

The environmental scientist (with user name as "EPD") and atmospheric scientist (with user name as "HKO") synchronously set ten input parameters for Gaussian plume model. Among this parameters, wind speed and environment temperature are different, which mean the conflicts exist. "HKO" set wind speed and environment temperature as 2m/s and 25 degree Celsius respectively. However, "EPD" set wind speed and environment temperature as 4m/s and 20 degree Celsius respectively. Two conflicts are automatically detected by the program as Figure 8.13 (below part).

The mechanism of conflict detection discussed in chapter 7.4.1 is applied here to identify the differences among parameters. The online communication based chat, which is discussed in chapter 7.4.2 (2), is used as a tool to solve the conflicts. The parallel computation of Gaussian plume model, which is studied in chapter 5.3, is adopted to minimize computation time. After negotiation, the final input parameters are listed in table 8.4.

Table 8.4: Input parameters

Parameters	Value
WindSpeed	2m/s
WindDirection	30°
DateSequency	125
BeijngTime	14:00
TotalCloud	8
LowCloud	7
DewTemperature	15
SurfaceRoughness	0.03
PollutantHalfLifeTime	4Hour
EnvironmentTemperature	25°

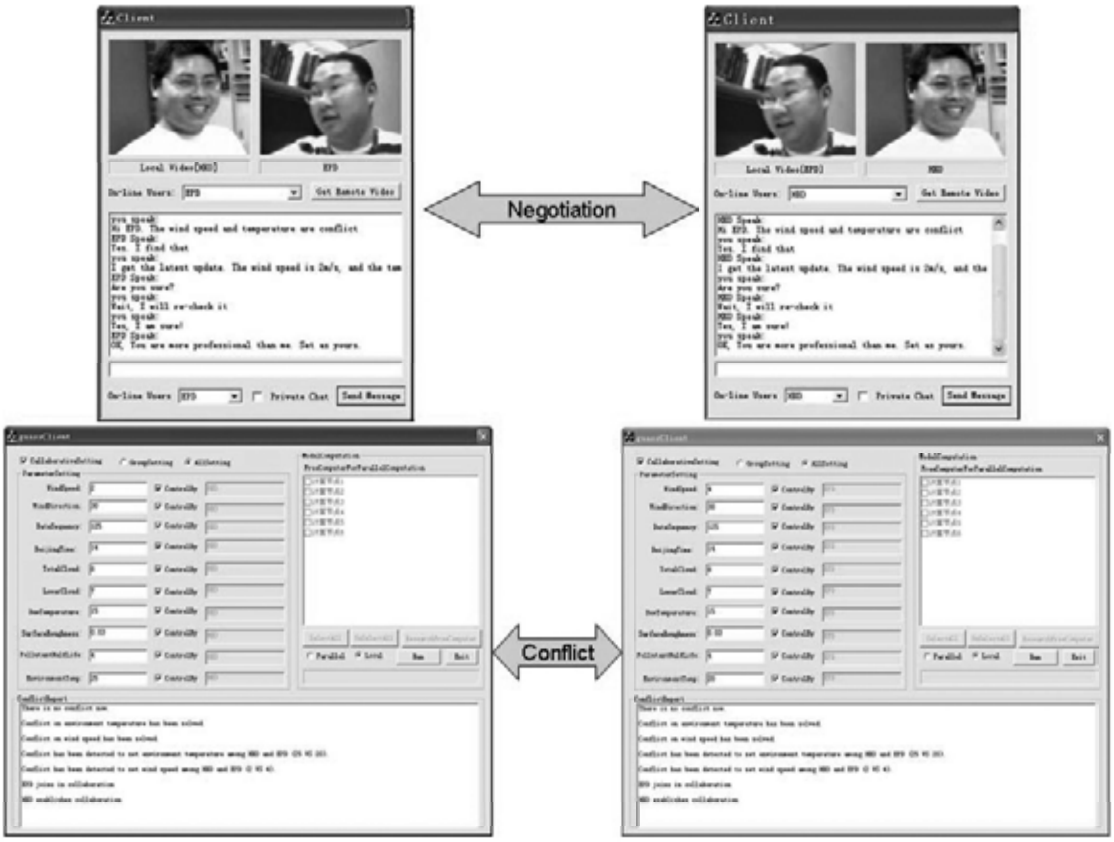


Figure 8.13: Conflict detection and resolution of collaborative setting on Gaussian plume model

Step 4: Geo-visualization of result from Gaussian plume model computation

Being notified that the computation of Gaussian plume model has been finished, the geographer takes over the work of geo-visualization of the computation result. He conducts the data transformation designed as Figure 6.1, creates layers and volume of pollution mass (PM_{10}), and geographically visualize the layers and volume in 2D and 3D environments respectively. The effects of geo-visualization are shown as Figure 8.14 and 8.15.

The geo-visualization conducted by the geographer can also be viewed by the other two participants.

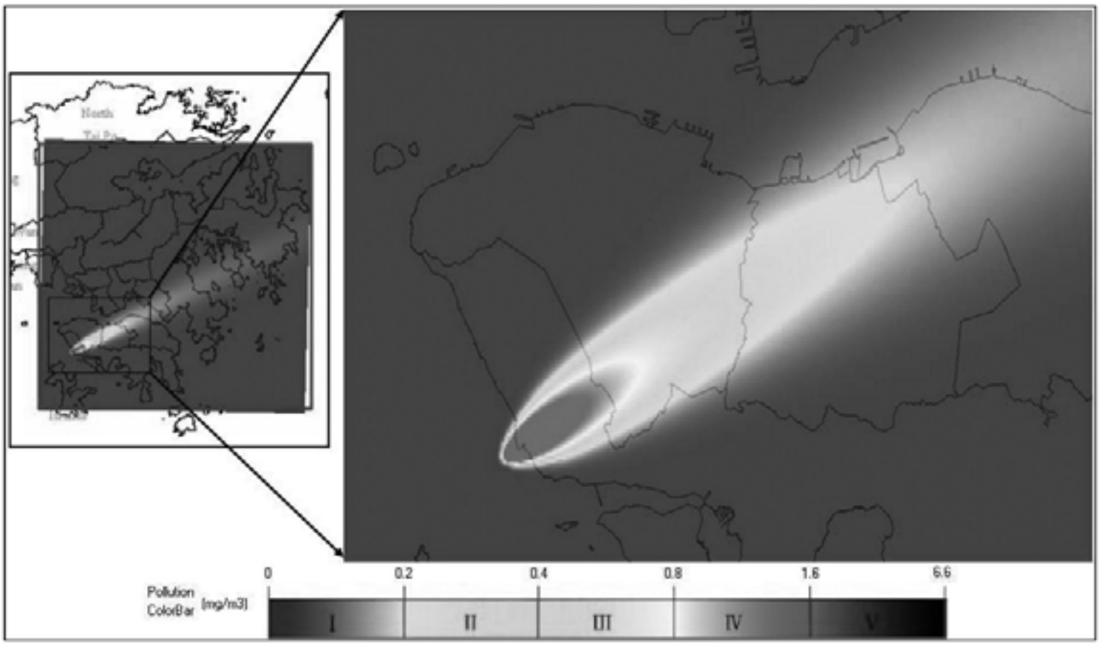


Figure 8.14: Geo-visualization of the result of Gaussian plume model computation in 2D environment

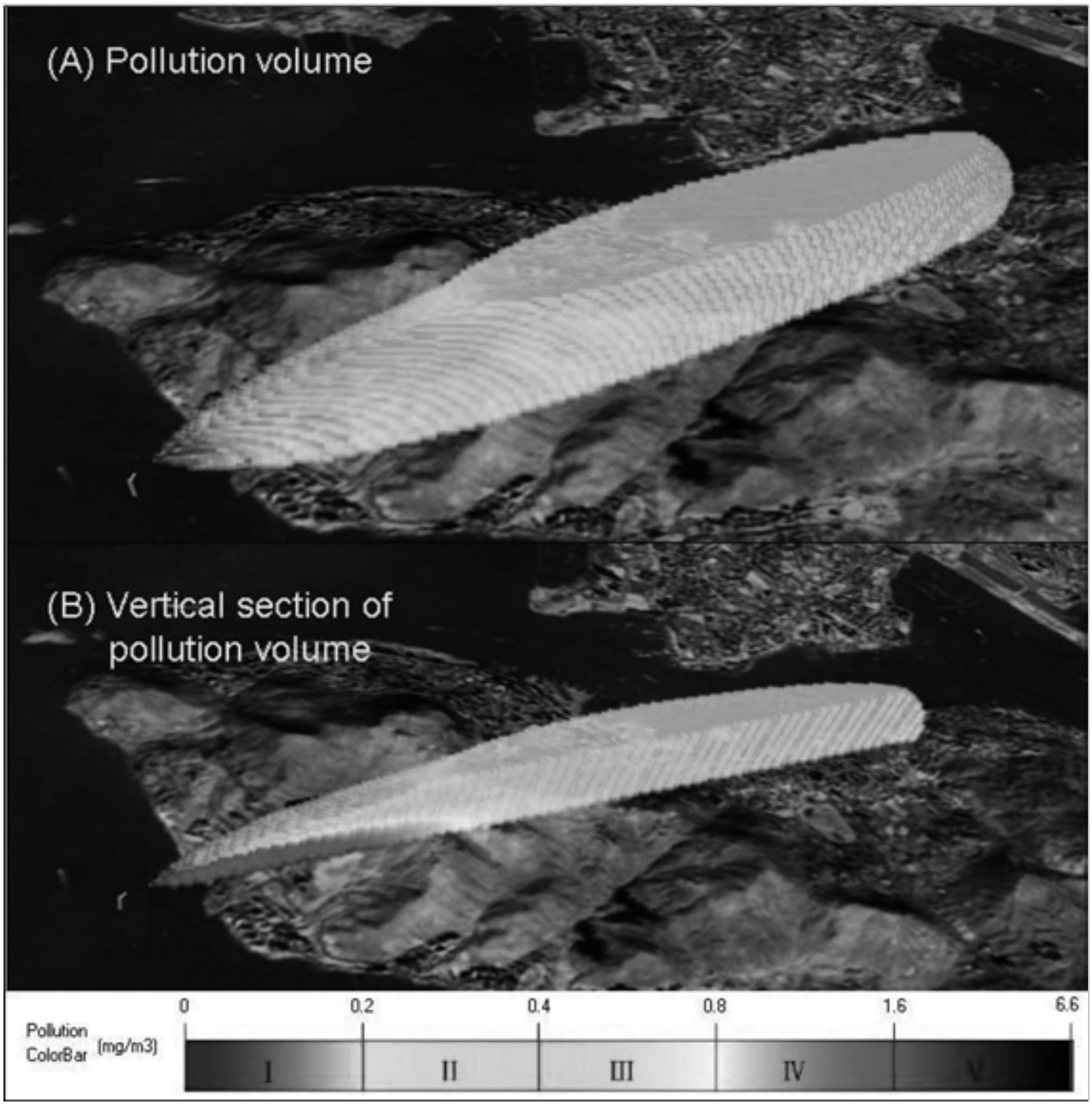


Figure 8.15: Geo-visualization of the result calculated by Gaussian plume model in 3D environment

Step 5: Spatial analysis to support decision making

The geographer continues his work into this step. Spatial overlay analysis is applied here to indentify the buildings severely polluted by PM_{10}. In Figure 8.16, the layer of Hong Kong district boundary, the layer of Hong Kong buildings, and the layer of distribution of PM_{10} are all overlaid together. Figure 8.16 (A) is the outline, while (B) is the enlarged part. In Figure 8.16 (B), the highlighted footprints of buildings indicate the space where the residents should be immediately evacuated.

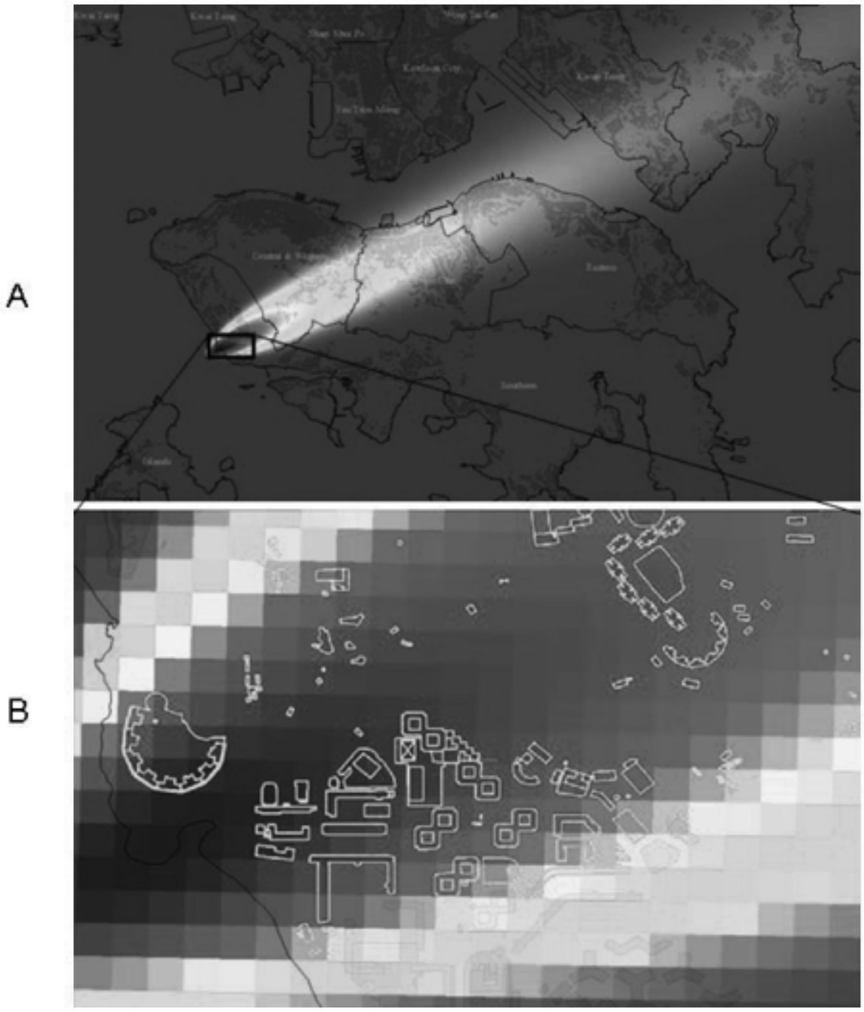


Figure 8.16: Spatial overlay analysis to identifying the buildings severely polluted to support decision of evacuating residents.

8.3 Discussion

8.3.1 Fulfilment of the Objectives

In the first chapter, the research questions regarding current air pollution simulation were highlighted. Chapter one also proposed a solution to these problems. In this section, the fulfilment of the prototype system of CVGE based on the proposed solution is discussed, along with a comparison table (Table 8.5).

Table 8.5: Objectives fulfillment—Comparison the implemented prototype system with research questions existing in air pollution simulation

Contents	Challenges in current air pollution simulation	Prototype system achieved	Comments
Data	● Depend on geospatial data process tools; ● Time consuming; ● No data management;	● Integrate geo-data, air pollution sources, the result of air pollution model computation; ● Information retrieved from integrated data can be directly input into atmospheric and environmental models; ● Data management and data process are both integrated;	The prototype system couples geo-data with geo-models.
Model	● So many models have been developed. The efficiency of models' computation is always the problem with computer technology moving on. ● Interface of MM5 is unfriendly, and it is also hardly to be integrated to support applications study;	● Adopt CUGrid to minimize computation time of MM5; ● Design the parallel arithmetic to decrease computation time of Gaussian Plume Model ● The operation interface of MM5 is friendly window style. It is portable and easy be integrated by applications. The interface also supports professional operation.	The geo-location stamped at results of air pollution modeling will open opportunities of the cross study between air pollution with other disciplines.

Visualization	● Lack of geo-information visualization; ● Being lowly interactive (do not support zoom in/out, query, roam, view point changing, and so on.); ● Low quality of visualization;	● Integrated geo-visualization of geo-data, geo-information, and air pollution distribution and dispersion; ● Multi-dimensional geo-visualization with diagram, 2D environment, 3D environment, and animation; ● Geo-visualization is high interactive; ● Particles system and pollution boxes to present air pollutant in natural way;	To present air pollution dispersion much clearer. Free visualization of air pollutant, which lays foundations for further research.
Spatial and temporal analysis	● Almost do not support spatial and temporal analysis;	● Support layer overlay; ● Support profile analysis at a click location; ● Support transection analysis; ● Support animation	Improve air pollution analysis.
Collaboration	● Few researches are carried out on collaboration of air pollution simulation ● Almost all of existing tools/software to facilitate air pollution simulation are/is designed for use by individual.	● Pollution sources collaboration; ● Collaborative setting models (MM5 and Gaussian plume model); ● Collaborative geo-visualization; ● Collaborative analysis	This effort will benefit both scientific research on CVGE and daily air quality management.

Although the prototype system is immature, the merits are obvious. The following section will discuss these merits, while at the same time highlighting the deficiencies for further research and development.

8.3.2 Merits

(1) The integrated platform of air pollution modeling and computation, visualization and analysis

The prototype system has integrated five steps (data preprocess, air pollution modeling and computation, visualization, model evaluation, and analysis) of air pollution simulation. It avoids the copy and paste effort for data exchange among the tools to support one or several steps.

(2) Geographic information coupled with air pollution modelling

Air pollution is a geographic process. Currently, modelling air pollution is mainly carried out by experts who have majored in atmosphere science, and they almost ignore geographic information. When performing spatial and temporal analysis, some GIS tools are used. But unfortunately, the incompatibility of the data formats between the model for the output of air pollution and GIS creates an obstacle to performing any analysis. This prototype has made some contribution to this problem. In this system, the output from the air pollution model computation is transformed into geographic referenced data, which can be seamlessly integrated with geographic information. This paves the road for spatial and temporal analysis of air pollution. At the same time, some settings from the air pollution model computation are carried out on geographic information supported by digital maps, which makes model operation easy and intuitive.

(3) High performance computation integration into CVGE to support air pollution modelling and computation

The integration and management of geo-models are remarkable characteristics for CVGE. This prototype system attempts the most difficult approach using the complex and computation intensive model (MM5). Tests show that CVGE integrated with the CUGrid can significantly improve complex model computation. This function contributes not only to further research and development on CVGE, but also to MM5 computation for professional study. The parallel arithmetic for the Gaussian plume model has similar merits to CUGrid based MM5 computation.

(4) Supporting Geo-collaboration

Geo-collaboration, which is supported by this prototype system, has been partly implemented for group work on air pollution issues. Collaborative compiling of pollution sources, collaborative setting for Gaussian plume model computation, collaborative visualization and collaborative analysis are also addressed in this prototype system. Collaboration on compiling pollution sources supplies an efficient way for distributed management on source data scattered over a large area. Collaborative settings for Gaussian plume model computation assures the most reasonable configuration of parameters. Geo-collaborations on visualization and analysis show a way for transforming geo-knowledge and finding the optimal solution.

(5) Multi-dimension geo-visualization

Air pollution is visualized in the geographically referenced environments. This prototype system not only implements the widely-used 2D geo-visualization, but also adds some functions for a 3D geographical environment. Both environments can be used to present pollution distribution and animated dispersion. However, the two kinds of visualization can also act collaboratively, which enhances the ability to explore potential rules.

(6) High interactive interface

The most widely-used tools for air pollution visualization today include GrADS for 2D visualization and Vis5D for high-dimension visualization. However, interactions with both tools are very limited. GrADS only supports a scale-fixed visualization display and does not accept any running interactions. Vis5D is also scale-fixed. It does support transection operation, but there is no free roam, no zoom in/out and no visualized querying.

Besides implementing the main functions of GrADS and Vis5D, this prototype system offers many new interactions for operating model computation, data processing, visualization and analysis. The system discards the tradition operations of MM5 computation using command lines, and achieves interactive windows. It also allows visualization scale, roaming, zoom in/out and visualized querying, among others.

Although the merits are extracted from typical applications, the methodologies behind these can be applied not only for further research and development on CVGE, but also for air pollution solving for PRD and for other regions.

8.3.3 Deficiencies

CVGE and air pollution simulation are both major topics. It is impossible to develop a system that satisfies all users. The functions of this prototype system are limited, but do cover almost every step of air pollution simulation. In the system test, some deficiencies are found.

(1) Rationality of the air pollution dispersion model

This research emphasizes the integration of an air pollution model with CVGE, but not the creation of air pollution dispersion models. However, the system can be used to ascertain the reasonability of such a model. System tests indicate that an improved MM5 is good for air pollution distribution computation before noon. After noon, the errors are significant between the model computation results and the real monitoring records. These errors have nothing to do with the methodology of system creation; but if the system is applied to find a scientific solution, the air pollution dispersion model must be improved to meet more rational ones.

(2) Low efficiency of some function of visualization

CVGE supports multi-dimensional visualizations. The high-dimension and high -definition visualizations are the most powerful way of visually exploring geo-information. However, they do also consume more computer resources. Some visualization functions require too much memory to be rendered. This problem occurred in this research. In a 3D environment, the particle system was applied to present the fuzzy boundary volume of an air pollutant. This method is sound for pollution mass rendering in a small region, such as the whole of Hong Kong. But when moving to a large area such as the whole PRD, because the number of particles is very high, the rendering frame rate is almost decreased to zero. Because of this problem, the final method for presenting air pollution mass for the whole PRD is based on a pollution box with no particles inside. This simplified rendering can present a volume feature, but not a fuzzy feature.

(3) Simple rules for negotiation on geo-collaboration conflicts

Rule based negotiation is very complex. This thesis only considers two factors, which are the roles and the number of the same roles with shared opinions. This deficiency is complemented by another method of negotiations, which is communication based negotiation. For further research, this deficiency should be resolved.

CHAPTER 9: CONCLUSIONS, CONTRIBUTIONS AND FUTURE RESEARCH

9.1 Conclusions

Air pollution is a global problem. Deep understanding, accurate prediction and efficient control of air pollution have become the focus of scientists, governmental officials and the public all over the world. Air pollution simulation is one major method that can achieve these goals. Only a computer supported type of simulation is examined in this research.

Air pollution simulation has several components: data preparation, modelling and computing atmospheric circulation and air pollution dispersion, visualizing the results of model computations, analysis and model evaluation. In order to support decision making, air pollution simulation should add another two components: goal identification and a strategy of emission control. Of these components, there has been research into modelling and computing atmospheric circulation and air pollution dispersion, but the other components are weak to a greater or lesser extent. To address these weaknesses, this thesis proposes a new solution which is to integrate data, model and analysis into a multi-dimension, virtual, geographically referenced environment. This environment can retrieve the input information to support air pollution modelling, visualize air pollution distribution/dispersion in a virtual geographic space and conduct spatial and temporal analysis. Most importantly, collaboration is added to solve the problem of multi-disciplinary knowledge requirements for conducting air pollution simulation. From the perspective of the geography community, the data, models, visualization, analysis and collaboration are viewed as geo-data, geo-models, geo-visualization, spatial and temporal analysis and geo-collaboration, respectively. A platform which effectively integrates all of these elements is conceptualized in this thesis as a collaborative virtual geographic environment (CVGE).

There are two focuses in this thesis. One is the development of a conceptual framework and prototype of CVGE from practice in air pollution simulation. The other is to apply this conceptual framework to facilitate air pollution simulation. The key works done by this research are as follows:

(1) Defining the concept of CVGE, developing the conceptual framework of CVGE and discussing the primary theories of CVGE. The thesis proposes a CVGE central double-core of geo-data and geo-model, and engages multiple participants from geographically dispersed locations to study a geo-spatial problem in a shared virtual environment. The research objectives for CVGE include the theories and methodologies required to construct CVGE. The theories of CVGE come from the pillar sciences, including geography, earth system science, social science, information science, computer science and applied fields.

(2) Designing the architecture of a CVGE prototype to facilitate air pollution simulation. The tier architecture, functional architecture and hardware architecture are all designed. The key issues that exist in these architectures include: ① complex air pollution modelling and computation based on high performance computation; ② geographically visualizing the professional results that are calculated by the air pollution dispersion models so that they can easily be understood; and ③ geo-collaboration on air pollution simulation in its lifecycle. These three key issues are highlighted and studied in detail in three separate chapters.

(3) Integration and computation of a complex atmospheric circulation model and an air pollution dispersion model based on high performance computation. Two scales of air pollution modelling are considered. One is on a regional scale, with the support models of MM5 (an atmospheric circulation model) and SYSUM (an air pollution dispersion model). MM5 is complex and computationally intensive. In order to minimize the computation time, the CUGrid is adopted. The interface for MM5 operation is improved from the original command lines to a windows style, which makes MM5 integration into CVGE easy. The other scale is the Gaussian plume model, which is on a local scale. When calculating the multi-layer and multi-grid for each layer impacted by the multiple pollution sources, the Gaussian plume model also faces a problem of time consumption. In order to decrease the computation time of the Gaussian plume model, a parallel arithmetic is designed and run based on the PC cluster.

(4) Geo-visualization of air pollution distribution and dispersion calculated by air pollution dispersion models. Geo-visualization attempts to combine tables, diagrams, 2D digital maps, 3D virtual environments and animation to present geo-spatial phenomena and processes in a geographically referenced virtual space. The air pollution distribution and dispersion data that is calculated by the air pollution dispersion models on different scales are not compatible with geo-data and geo-information. Data conversion and transformation are thus unavoidable. After transformation, the air pollution distribution and dispersion data is stamped with geo-spatial locations, which can be read and visualized in both 2D and 3D environments. Both 2D and 3D environments can geographically visualize not only geo-data and geo-information, but also air pollution distribution and dispersion. An air pollution mass is a fuzzy boundary volume object. An air pollution mass is also dynamic, because its position and shape change from time to time. This thesis uses a methodology of pollution boxes wrapped particles to present air pollution distribution and dispersion in a 3D environment. Several methods are also studied to improve the efficiency of rendering an air pollution mass.

(5) Geo-collaboration on air pollution simulation. The geo-collaboration on air pollution simulation occurs on many levels, including data collaboration, modelling and computation collaboration, geo-visualization collaboration and analysis collaboration. The mode of geo-collaboration for air pollution simulation is different place, same/different time collaboration. The mediator for this geo-collaboration is the distributed, multi-dimension virtual environment. The thesis designs the architecture for organizing multi-level geo-collaboration. In order to detect conflict among the many participants, a comparison mechanism is used. The conflict resolution is based on two methods of negotiation: one is rules based negotiation, and the other is online, chat based negotiation.

(6) CVGE prototype based air pollution simulation. Two objectives are focused on in this section: one is to respond to the questions regarding the status of air pollution simulation; the other is to test the feasibility and efficiency of the prototype system by simulating typical scenarios. Two scenarios are simulated. One is an MM5 based air pollution (SO_2) simulation for the Pearl River Delta (PRD). In this episode, there are five steps of air pollution simulation: preparing air pollution sources, setting and computation of MM5, visualization, analysis and model evaluation. Each of these is conducted step-by-step using the newly developed methodologies. At the same time, each step is accompanied by geo-collaboration. In this episode simulation, the geo-collaboration is a collaborative workflow, and there is no conflict. The second scenario is a Gaussian plume model based hypothetical accident (ship wreckage) simulation near Hong Kong Island. This simulation is intended to identify buildings severely polluted by PM_{10}, and also has five steps. However, the geo-collaboration is different from the first simulation in that conflicts exist. Consequently, negotiation is applied to solve these conflicts. The second main objective, which is to test the feasibility and efficiency of the prototype system, is also achieved by these two scenario simulations.

The following conclusions can be drawn from the simulations: ① the conceptual framework of CVGE is valuable for facilitating air pollution simulation by allowing geo-data integration, geo-model integration and computation, geo-visualization and geo-collaboration; ② from the perspective of the geography community, the prototype system designed for air pollution simulation proves that the conceptual framework and key issues of CVGE that are studied in this thesis are rational and efficient.

9.2 Contributions

On the one hand, the research is a major investigation of the typical application of CVGE (air pollution simulation); on the other hand, it also applies the thought behind CVGE to facilitate practice of air pollution simulation. The contribution can thus be concluded from the two aspects of CVGE and application.

From the perspective of CVGE, the thesis:

- Develops the conceptual framework of CVGE;
- Designs the architecture of a CVGE prototype to facilitate air pollution simulation;
- Distributes MM5 into multiple nodes, modifies the interface of MM5 so that it is more user-friendly, integrates MM5 into CVGE and conducts high performance computation for MM5 based on a computational grid;
- Designs a parallel arithmetic for the Gaussian plume model and conducts parallel computation on a PC cluster;
- Proposes the concept of a "fuzzy boundary volume object" and designs a solution of a particle system wrapped with pollution boxes to represent this;
- Highlights the level of geo-collaboration required for air pollution simulation.

In terms of air pollution simulation, the contributions of this thesis are:

- Integrates air pollution sources, geo-data, an atmospheric circulation model, an air pollution dispersion model, geo-visualization and analysis into a collaborative, virtual, geographic environment, which supplies a new research methodology and platform for air pollution simulation;

- Offers a new platform which is scalable and free from restrictions of operation on visualization, paving the way for further extension;
- Couples air pollution dispersion models with geo-information, which opens the opportunity for cross-study between air pollution and other research areas, such as the economy, public health and urban planning.

9.2 Future Research

Air pollution research and management are major topics, and this research cannot cover and solve them completely. There are thus many opportunities for research that remain for future study. The motivation for future research can also be drawn from two different aspects—CVGE and application.

From the perspective of CVGE, future research can be focused on:

- Systematic discussion of CVGE, which is at an initial stage; there is still a long way to go on the concept, primary theories and methodologies of CVGE;
- The efficient integration and management of heterogeneous geo-models into CVGE in standardization;
- The efficient rendering of a complex structured object in CVGE;
- Further work on the scope, profundity and usability of CVGE.

From the perspective of practice, future research can be focused on:

- Extending air pollution dispersion models;
- Improving the efficiency of air pollutant rendering with a pollution boxes wrapped particle system. In future research, parallel visualization is one option for improving visualization efficiency, and a Tiled Display Wall (TDW) could be used for a large-scale, high-resolution display.
- Conducting cross-research.

Appendices A: Publication, Academic Activities, And Awards During the Period of PhD Research

Selected Publications:

Hui Lin, Jun Zhu, Jianhua Gong, **Bingli Xu**, Hua Qi. A Grid-based Collaborative Virtual Geographic Environment for the Planning of Silt Dam Systems. International journal of geographic information science. (In press)

Bingli Xu, Hui Lin, Jun Zhu, Wenshi Lin, Sammy Tang, Ya Hu, Jianbin Wu, Liping Zeng. Construction of a Virtual Geographic Environment for Pearl River Delta Air Pollution Simulation. Journal of Geomatics and Information Science of Wunan University. 2009.6 (In Chinese)

Xu Bing-Li, Lin Hui,Gong Jian-Hua,Zang Yong-Qiang, Huang Ming-Xiang. Water Pollution Visualization Simulation on Beijing-Hangzhou Canal within Yangzhou City. Journal of System Simulation. 2009.6 (In Chinese)

Hui Lin, Jun Zhu, **Bingli Xu**, Wenshi Lin and Ya Hu, A Virtual Geographic Environment for a Simulation of Air Pollution Dispersion in the Pearl River Delta (PRD) Region.Lecture Notes in Geoinformation and Cartography, chapter1 in 3D Geo-Information Sciences 2008,pp:3-11.

Bingli Xu, Hui Lin, Jun Zhu, Sammy Tang, Wenshi Lin, and Jianbin Wu. Grid based model computation of virtual geographic environment: application in Pearl River Delta air pollution visualization. Proc. SPIE Vol. 7143, 714325 (Nov. 3, 2008)

Jun Zhu, Hui Lin, **Bingli Xu**, and Ya Hu.Real-time visualization of virtual geographic environment using the view-dependent simplification method. Proc. SPIE Vol. 7143, 71432F (Nov. 3, 2008)

LIN Hui, **XU Bingli**. Some Thoughts on Virtual Geographic Environments.

Geography and Geo-Information Science. Vol13, No.2,2007 (In Chinese)

Xu Bingli, Lin Hui, Gong Jianhua. Architecture of HLA Based Distributed Virtual Geographic Environment. Geo-spatialInformation Science (Quarterly). Volume9, Issue2, June 2006

Bingli Xu, Jianhua Gong,Hui Lin,Wenhang Li,Jianqin Zhang,Jun Zhu,Xian Wu.**Virtual geographic environment database design and collaboration.** Geoscience and Remote Sensing Symposium. 2005. IGARSS '05. Proceedings. 2005 IEEE International，Volume 2, 25-29 July 2005 Page(s):4 pp. Digital Object Identifier 10.1109/IGARSS.2005.1525240

Awards:

2008. Second best paper award. Conference of Urban GIS across Taiwan Strait. Tianjin, 2008.9

2007. Receipents of Yuen Yuen remote sensing technology scholarship. The Chinese university of Hong Kong and Yuen Yuen Institute.

2006. Receipents of Yuen Yuen remote sensing technology scholarship. The Chinese university of Hong Kong and Yuen Yuen Institute.

Academic Activities:

2006.12-2009.11. Project of “A Collaborative Virtual Geographic Environment System for Air Pollution Simulation for Clustered Cities in Pearl River Delta”. The National High-Tech Research and Development Program ("863"Program) of China. (2006AA12Z207). Works include proposal writing, research and development coordination, system design, annual report and defense, MM5 integration, parallel arithmetic design of Gaussian plume model, geo-visualization design and 2D environment implementation, geo-collaboration design and partly implementation.

2007-2009. Project of "Collaborative Virtual Geographic Environment Based Air Pollution in Pearl River Delta". General Research Fund—Competitive Earmarked Research Grant. 447807. Works include proposal writing, research and development coordination, system design, MM5 integration, parallel arithmetic design of Gaussian plume model, geo-visualization design and 2D environment implementation, geo-collaboration design and partly implementation.

May 2009. Oral presentation on "Design and Implement of a Distributed Virtual Geographic Environment for Pearl River Delta Air Pollution Simulation". The second international conference on earth observation for global change. ChengDu, China.

Mar. 2009. Attend. Conference on Spatial Integrated Humanities and Social Science (In Chinese). The Chinese University of Hong Kong, Hong Kong.

Sep.2008. Oral presentation on "Collaborative Virtual Geographic Environments Based Air Pollution Simulation for Pearl River Delta". Symposium on Theory and Technology of GIS (In Chinese). Nanzhou, China.

Sep. 2008. Oral presentation on "Collaborative Virtual Geographic Environments Based Air Pollution Simulation for Pearl River Delta". Conference of Urban GIS across Taiwan Strait(In Chinese). Tianjin, China.

May 2008. Attend. The Joint Conference of CyberCity and Virtual Geographic Environments(In Chinese). Wuhan, China.

Mar.2008. Oral presentation on "Grid based model computation of virtual geographic environment: application in Pearl River Delta air pollution visualization". GeoInfromatics 2008. Sun Yat-Sen University, Guangzhou, China.

Jan. 2008. Attend. Virtual Geographic Environments - An International Conference on Developments in Visualization and Virtual Environments in Geographic

Information Science. The Chinese University of Hong Kong, Hong Kong.

Dec.2007. Attend. Symposium on Theory and Technology of GIS. Sun Yat-Sen University, Guangzhou (In Chinese) , China.

Apr.2007. Oral presentation on "Water Pollution Visualization Simulation on Beijing-Hangzhou Canal within Yangzhou City". The 12th Inter-University seminar on Asian Mega-cities -- Problems and visions of Asian metropolitan areas in the 21st century. Seoul National university, HanYang University. Korea.

Dec.2006. Attend. The 6th International Workshop on Web and Wireless Geographical Information Systems. The Chinese University of Hong Kong, Hong Kong.

REFERENCES

Ai, Z. M., X. Chen, M. Rasmussen & R. Folberg (2005) Reconstruction and exploration of three-dimensional confocal microscopy data in an immersive virtual environment. *Computerized Medical Imaging and Graphics,* 29, 313-318.

Alberola, C., R. Cardenes, M. Martin, M. A. Martin, M. A. Rodriguez-Florido & J. Ruiz-Alzola. 2000. diSNei: A collaborative environment for medical images analysis and visualization. In *Medical Image Computing and Computer-Assisted Intervention - Miccai 2000,* 814-823.

AmCham, H. K. A. C. o. C. 2007. Air Pollution Affects Recruitment, Survey Shows.

Andrade, N., D. A. Palmer & L. A. Lenert (2006) A WiFi public address system for disaster management. *AMIA Annu Symp Proc*, 846.

Andrienko, G., N. Andrienko, P. Jankowski, D. Keim, M. J. Kraak, A. Maceachren & S. Wrobel (2007) Geovisual analytics for spatial decision support: Setting the research agenda. *International Journal of Geographical Information Science,* 21, 839-857.

Antonarakis, A. S., K. S. Richards & J. Brasington (2008) Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment,* 112, 2988-2998.

Arora, N. K., A. L. Dans & M. A. D. Lansang (2007) The JCE-INCLEN collaboration: knowledge sharing in action. *Journal of Clinical Epidemiology,* 60, 538-539.

Arystanbekova, N. K. (2004) Application of Gaussian plume models for air pollution simulation at instantaneous emissions. *Mathematics and Computers in Simulation,* 67, 451-458.

Austerlitz, N. & A. Sachs (2006) Community collaboration and communication in the design studio. *Open House International,* 31, 25-32.

Bal, M. & M. Hashemipour (2007) Applications of virtual reality in design and simulation of holonic manufacturing systems: A demonstration in die-casting industry. *Holonic and Multi-Agent Systems for Manufacturing, Proceedings,* 4659, 421-432.

Balram, S. & S. Dragicevic (2003) Integrating geographic information systems and agent-based modeling techniques for simulating social and ecological processes. *Professional Geographer,* 55, 301-302.

Barcelo, J. A. (2001) Virtual Reality and Scientific Visualization. Working with models and hypotheses. *International Journal of Modern Physics C,* 12, 569-580.

Battersby, S. J., D. J. Brown & H. M. Powell (2000) Virtual reality game simulations for the education of people with learning disabilities. *Simulation in Industry'2000*, 395-399.

Batty, M. 2005. Approaches to modeling in GIS: spatial representation and temporal dynamics. In *GIS, Spatial Analysis and Modeling,* eds. D. J. Maguire, M. Batty & M. F. Goodchild. California: ESRI Press.

Beavis, A., J. Ward, P. Bridge, R. Appleyard & R. Phillips (2006) An immersive virtual environment for training of radiotherapy students and developing clinical experience. *Medical Physics,* 33, 2164-2165.

Benenson, I. (2003) Integrating geographic information systems and agent-based modeling techniques for simulating social and ecological processes. *Environment and Planning B-Planning & Design,* 30, 158-160.

Bhowmick, T., A. L. Griffin, A. M. MacEachren, B. C. Kluhsman & E. J. Lengerich (2008) Informing geospatial toolset design: Understanding the process of cancer data exploration and analysis. *Health & Place,* 14, 576-607.

Bishop, I., C. Stock & K. Williams (2009) Using virtual environments and agent models in multi-criteria decision-making *LAND USE POLICY* 26, 87-94.

Boejen, A., A. Beavis, K. Nielsen, R. Phillips, T. S. Soerensen, J. Ward & C. Grau (2007) Training of radiation therapists using a 3D virtual environment. *Radiotherapy and Oncology,* 84, S275-S275.

Boukerche, A., H. Maamar & A. Hossain (2008) An efficient hybrid multicast transport protocol for collaborative virtual environment with networked haptic. *Multimedia Systems,* 13, 283-296.

Bouras, C., E. Giannaka, A. Panagopoulos & T. Tsiatsos (2006) A platform for virtual collaboration spaces and educational communities: the case of EVE. *Multimedia Systems,* 11, 290-303.

Bryan, W. (2006) Let There Be Light. *Time*.

Cai, G., L. Bolelli, A. M. MacEachren, R. Sharma, S. Fuhrmann & M. McNeese. 2004. GeoCollaborative Crisis Management (GCCM): Building better systems through advanced technology and deep understanding of technology-enabled group work In *5th Annual NSF Digital Government Conference*. Los Angeles.

Cai, G., A. M. MacEachren, I. Brewer, M. McNeese, R. Sharma & S. Fuhrmann. 2005. Map-mediated GeoCollaborative Crisis Management. eds. P. Kantor, G. Muresan, F. Roberts, D. D. Zeng, F. Y. Wang, H. Chen & R. C. Merkle, 429-435.

Canty, M. J. (2009) Boosting a fast neural network for supervised land cover classification. *Computers & Geosciences,* In Press, Corrected Proof.

Chai, S., Y. F. Zhang, Y. J. Teng, X. J. Jiao, C. B. Ren & S. F. Wang (2008) Virtual Assembly and Dynamic Simulation of Tractor Gearbox Based on Virtual Reality Technology. *Proceedings of First International Conference of Modelling and Simulation, Vol V*, 292-295.

Chen, C. H., J. C. Yang, S. Shen & M. C. Jeng (2007) A desktop virtual reality earth motion system in astronomy education. *Educational Technology & Society,* 10, 289-304.

Chen, Y. J., Y. M. Chen & H. C. Chu (2008) Enabling collaborative product design through distributed engineering knowledge management. *Computers in Industry,* 59, 395-409.

Cheng, F.-Y. & D. W. Byun (2008) Application of high resolution land use and land cover data for atmospheric modeling in the Houston-Galveston metropolitan area, Part I: Meteorological simulation results. *Atmospheric Environment,* 42, 7795-7811.

Cheng, F.-Y., S. Kim & D. W. Byun (2008) Application of high resolution land use and land cover data for atmospheric modeling in the Houston-Galveston Metropolitan area: Part II: Air quality simulation results. *Atmospheric Environment,* 42, 4853-4869.

Chertov, O., A. Komarov, A. Mikhailov, G. Andrienko, N. Andrienko & P. Gatalsky (2005) Geovisualization of forest simulation modelling results: A case study of carbon sequestration and biodiversity. *Computers and Electronics in Agriculture,* 49, 175-191.

Chou, B. & V. L. Handa (2006) Simulators and virtual reality in surgical education. *Obstetrics and Gynecology Clinics of North America,* 33, 283-+.

Coleman, D. & S. Levine. 2008. *Collaboration 2.0--Technology and best practices for successful collaboration in a web 2.0 world.* Happy About

Collett, R. S. & K. Oduyemi (1997) Air quality modelling: a technical review of mathematical approaches. *Meteorological Applications,* 4, 235-246.

Cristaldi, L., A. Ferrero, A. Monti, F. Ponci, W. McKay & R. Dougal (2005) A virtual environment for remote testing of complex systems. *Ieee Transactions on Instrumentation and Measurement,* 54, 123-133.

Daoud, R. M., M. A. El-Dakroury, H. H. Amer, H. M. Elsayed & M. El-Soudani (2007) WiFi architecture for traffic control using MIPv6. *2007 Mediterranean Conference on Control & Automation, Vols 1-4*, 617-621.

DeFanti, T. A., G. Dawe, D. J. Sandin, J. P. Schulze, P. Otto, J. Girado, F. Kuester, L. Smarr & R. Rao (2009) The StarCAVE, a third-generation CAVE and virtual reality OptIPortal. *Future Generation Computer Systems-the International Journal of Grid Computing Theory Methods and Applications,* 25, 169-178.

Dewiyanti, S., S. Brand-Gruwel & W. Jochems (2005) Applying reflection and moderation in an asynchronous computer-supported collaborative learning environment in campus-based higher education. *British Journal of Educational Technology,* 36**,** 673-676.

Dodge, M., M. McDerby & M. Turner. 2008. The power of geographic visualizations. In *Geographic Visualization:Concepts, Tools and Applications,* eds. M. Dodge, M. McDerby & M. Turner. John Wiley & Sons, Ltd.

Doty, B. 1995. *The Grid Analysis and Display System.*

Dudhia, J., D. Gill, K. Manning, W. Wang & C. Bruyere (2005) PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and Users' Guide (MM5 Modeling System Version 3). *http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html.*

Dykes, J., K. Moore & J. Wood (1999) Virtual environments for student fieldwork using networked components. *International Journal of Geographical Information Science,* 13**,** 397-416

Edmans, J. A., J. R. F. Gladman, S. Cobb, A. Sunderland, T. Pridmore, D. Hilton & M. F. Walker (2006) Validity of a virtual environment for stroke rehabilitation. *Stroke,* 37**,** 2770-2775.

Encyclopedia, W.-t. f. (2009) Particle system. *http://en.wikipedia.org/ wiki/Particle_system*

EPA, U. S. 2006. USER'S GUIDE FOR THE AMS/EPA REGULATORY MODEL - AERMOD. ed. O. o. A. Q. P. a. S. U.S. ENVIRONMENTAL PROTECTION AGENCY, Air Quality Assessment Division. North Carolina.

EPA, U. S. E. P. A. 2008. National Air Quality - Status and Trends through 2007.

Evennou, F. & F. Marx (2006) Advanced integration of WiFi and inertial navigation systems for indoor mobile positioning. *Eurasip Journal on Applied Signal Processing.*

Eynard, B., S. Lienard, S. Charles & A. Odinot (2005) Web-based collaborative engineering support system: Applications in mechanical design and structural analysis. *Concurrent Engineering-Research and Applications,* 13**,** 145-153.

Feng, Y., A. Wang, D. Wu & X. Xu (2007) The influence of tropical cyclone Melor on PM10 concentrations during an aerosol episode over the Pearl River Delta region of China: Numerical modeling versus observational analysis. *Atmospheric Environment,* 41**,** 4349-4365

Finlayson, D. P. & D. R. Montgomery. 2003. Modeling large-scale fluvial erosion in geographic information systems. 147-164. Elsevier Science Bv.

Fortner, B. (2006) Virtual environment to aid tunnel design and construction. *Civil Engineering,* 76**,** 32-33.

Fraser, M. (2006) The ethics of reality and virtual reality: Latour, facts and values. *History of the Human Sciences,* 19**,** 45-72.

Freund, E., J. Rossmann & M. Schluse (2000) Projective virtual reality in space applications: A telerobotic ground station for a space mission. *Sensor Fusion and Decentralized Control in Robotic Systems Iii,* 4196**,** 279-290.

Fukui, K., K. Honda & K. Okada (2004) Promotion of multiparty conference in collaborative virtual environments. *Ieice Transactions on Information and Systems,* E87D**,** 2540-2547.

GA, G., D. J & S. DR (1994) A discription of the fifth generation Penn State/NCAR mesoscale model (MM5). *NCAR technical note, NAVR/TN-398+STR, National Center for Atmopheric Research, Boulder***,** 117.

Gao, J., Y. Xia & X. You. 1999. *Virtual Reality application in terrain simulation (虚拟现实在地形环境仿真中的应用).* Beijing(北京): People's Liberation Army Press(解放军出版社).

Gong, J., Y. Li, D. Wang, M. Huang & W. Wang (2008) Characteristics and Applications of Knowledge Maps in Geographic Knowledge Visualization:A Case Study of Spatial Planning of Dam Systems in Watersheds (地理知识可视化中知识图特征与应用——以小流域淤地坝系规划为例). *Journal of Remote Sensing(遥感学报),* 12**,** 355-361.

Gong, J. & H. Lin. 2001. *Virtual Geographic Environments-A Geographic Perspective on Online Virtual Reality(虚拟地理环境—在线虚拟现实的地理学透视)*. Beijing(北京): Higher education press(高等教育出版社).

---. 2006. A Collaborative Virtual Geographic Environment: Design and Development. In *Collaborative Geographic Information Systems,* eds. S. Balram & S. Dragicevic, 186-206. IGI Global,Hershey.PA.

Goodchild, M. F. (2000) Communicating geographic information in a digital age. *Annals of the Association of American Geographers,* 90**,** 344-355.

---. 2005. GIS and modeling overview. In *GIS, Spatial Analysis, and Modeling,* eds. D. J. Maguire, M. Batty & M.F. Goodchild, 1-18. Redlands, CA: ESRI Press.

Grudin, J. (1994) COMPUTER-SUPPORTED COOPERATIVE WORK - HISTORY AND FOCUS. *Computer,* 27**,** 19-26.

Gupta, S. & G. Kaiser (2005) P2P video synchronization in a collaborative virtual environment. *Advances in Web-Based Learning - Icwl 2005,* 3583**,** 86-98.

Hansson, A. (2000) Space tourism (Virtual reality, computers). *Architectural Design***,** 26-29.

Hibbard, B., J. Kellum, B. Paul, D. Santek, A. Battaiola & S. G. Johnson. 1999. Vis5D Documentation.

Hudson-Smith, A., R. Milton, J. Dearden & M. Batty. 2007. Virtual Cities: Digital Mirrors into a Recursive World. In *Working paper series*. London: Centre for Advanced Spatial Analysis

Hutanu, A., G. Allen, S. D. Beck, P. Holub, H. Kaiser, A. Kulshrestha, M. Liska, J. MacLaren, L. Matyska, R. Paruchuri, S. Pkohaska, E. Seidel, B. Ullmer & S. Venkataraman (2006) Distributed and collaborative visualization of large data sets using high-speed networks. *Future Generation Computer Systems-the International Journal of Grid Computing Theory Methods and Applications,* 22, 1004-1010.

ISLAM, M. A. (1999) Application of a Gaussian Plume Model to Determine the Location of an Unknown Emission Source *Water, Air, and Soil Pollution,* 112.

Jiang, W. 2004. *Air Pollution Meteorology(空气污染气象学)* NanJing: Nanjing University Press(南京大学出版社).

Jin, Y., M. Huang, H. M. Lin & J. Guo (2005) Towards collaboration: The development of collaborative virtual reference service in China. *Journal of Academic Librarianship,* 31, 287-291.

Johnson, A., J. Leigh, P. Morin & P. V. Keken (2006a) GeoWall: Stereoscopic Visualization for Geoscience Research and Education[J]., 2006,26(6):p10-14. *IEEE Computer Graphics and Applications,* 26, 10-14.

Johnson, A., J. Leigh, P. Morin & P. Van Keken (2006b) GeoWall: Stereoscopic Visualization for Geoscience Research and Education. *IEEE Computer Graphics and Applications,* 26, 10-14.

Jones, M. T. 2007. Google Earth: A Commitment to Community. In *The fifth international symposium on digital earth*. California US.

Kandrika, S. & P. S. Roy (2008) Land use land cover classification of Orissa using multi-temporal IRS-P6 awifs data: A decision tree approach. *International Journal of Applied Earth Observation and Geoinformation,* 10, 186-193.

Kao, K. J., C. E. Seeley, S. Yin, R. M. Kolonay, T. Rus & M. Paradis (2004) Business-to-business virtual collaboration of aircraft engine combustor design. *Journal of Computing and Information Science in Engineering,* 4, 365-371.

Kaufmann, H. & D. Schmalstieg (2006) Designing immersive virtual reality for geometry education. *Ieee Virtual Reality 2006, Proceedings*, 51-58.

Khan, M. N. 2003. Development and application of an adaptive grid air quality model. In *School of Civil and Environmental Engineering*. Georial Institute Technology.

Khoshafian, S. & M. Buckiewicz. 1995. *Introduction to groupware, workflow, and workgroup computing*. John Wiley and Sons.

Kim, S. J., F. Kuester & K. H. Kim (2005) A global timestamp-based approach to enhanced data consistency and fairness in collaborative virtual environments. *Multimedia Systems,* 10**,** 220-229.

Kotlarsky, J. & I. Oshri (2005) Social ties, knowledge sharing and successful collaboration in globally distributed system development projects. *European Journal of Information Systems,* 14**,** 37-48.

Kraak, M. J. (2003) Geovisualization illustrated. *Isprs Journal of Photogrammetry and Remote Sensing,* 57**,** 390-399.

--- (2006) Beyond geovisualization. *Ieee Computer Graphics and Applications,* 26**,** 6-9.

Lara, M. A. & S. Y. Nof (2003) Computer-supported conflict resolution for collaborative facility designers. *International Journal of Production Research,* 41**,** 207-233.

Larsson, P., D. Vastfjall & M. Kleiner (2001) The actor-observer effect in virtual reality presentations. *Cyberpsychology & Behavior,* 4**,** 239-246.

Lecuyer, A., F. Lotte, R. B. Reilly, R. Leeb, M. Hirose & M. Slater (2008) Brain-Computer Interfaces, Virtual Reality, and Videogames. *Computer,* 41**,** 66-+.

Li, D., Q. Zhu & X. Li (2000) CvberCity:Concepts, Technique Support and Typical Applications(数码城市:概念、技术支撑和典型应用). *Geomatics and Information Science of Wuhan University(武汉测绘科技大学学报),* 25.

Lim, C. H., Y. H. Wan, B. P. Ng & C. M. S. See (2007) A real-time indoor WiFi localization system utilizing smart antennas. *Ieee Transactions on Consumer Electronics,* 53**,** 618-622.

Limniou, M., D. Roberts & N. Papadopoulos (2008) Full immersive virtual environment CAVE (TM) in chemistry education. *Computers & Education,* 51**,** 584-593.

Lin, C.-C., M.-J. Chiu, C.-C. Hsiao, R.-G. Lee & Y.-S. Tsai (2006a) Wireless Health Care Service System for Elderly With Dementia. *IEEE Transactions onInformation Technology in Biomedicine,* 10**,** 696 - 704

Lin, H. & J. Gong (2002) On Virtual Geographic Environments(论虚拟地理环境). *Acta Geodaetica et Cartographica Sinica(测绘学报),* 31.

Lin, H., J. Gong & J. Shi (2003) From Maps to GIS and VGE--A Discussion on the Evolution of the Geographic Language(从地图到地理信息系统与虚拟地理环境--试论地理学语言的演变). *Geography and Geo-Information Science(地理与地理信息科学),* 19**,** 18-23.

Lin, H., F. Huang & G. Lv (2009) Development of Virtual Geographic Environments and the New Initiative in Experimental Geography (虚拟地理环境研究的兴起与实验地理学新方向). *ACTA Geographica Sinica (地理学报),* 64**,** 7-20.

Lin, H. & B. Xu (2007) Some Thoughts on Virtual Geographic Environments(关于虚拟地理环境研究的几点思考). *Geography and Geo-Information Science(地理与地理信息科学),* 23.

Lin, H. & Q. Zhu (2005a) The Linguistic Characteristics of Virtual Geographic Environments(虚拟地理环境的地理学语言特征). *Journal of Remote Sensing(遥感学报),* 9.

---. 2005b. *Virtual Geographic Environments.* Florida: CRC Press.

---. 2005c. Virtual Geographic Environments. In *Large-scale 3D Data Integration: Challenges and Opportunities,* eds. S. Zlatanova & D. Prosperi, 211-231. Florida: CRC Press.

Lin, Q., L. Zhang, I. Kusuma & N. Neo (2006b) Mobile agent based large-scale collaborative virtual environment system. *Euromedia '2006***,** 105-112.

Lococo, A. & D. C. Yen (1998) Groupware: computer supported collaboration. *Telematics and Informatics,* 15**,** 85-101.

Loo, G. S. L., D. H. S. Ly & B. C. P. Tang (2000) Mobile agent applications in a distributive interactive virtual environment: The happy paradise (HAPA). *11th International Workshop on Database and Expert Systems Application, Proceedings***,** 925-929.

Maceachen, A. M. & M. J. Kraak (2001) Research challenges in geovisualization. *Cartography and Geographic Information Science,* 28**,** 3-12.

MacEachren, A. M. (2000) Cartography and GIS: facilitating collaboration *Progress in Human Geography,* 24**,** 445-456.

--- (2001) Cartography and GIS: extending collaborative tools to support virtual teams. *Progress in Human Geography,* 25**,** 431-444.

MacEachren, A. M. & I. Brewer (2004) Developing a conceptual framework for visually-enabled geocollaboration. *International Journal of Geographical Information Science,* 18**,** 1-34.

MacEachren, A. M., I. Brewer, G. Cai & J. Chen. 2003. Visually-Enabled Geocollaboration to Support Data Exploration and Decision-Making In *International Cartographic Conference*, 394-401. Durban, South Africa

MacEachren, A. M., G. Cai, M. McNeese, R. Sharma & S. Fuhrmann. 2006. GeoCollaborative Crisis Management: Designing Technologies to Meet Real-World Needs In *7th Annual National Conference on Digital Government Research: Integrating Information Technology and Social Science Research for Effective Government*. San Diego.

MacEachren, A. M., M. Gahegan, W. Pike, I. Brewer, G. R. Cai, E. Lengerich & F. Hardisty (2004) Geovisualization for knowledge construction and decision support. *Ieee Computer Graphics and Applications,* 24, 13-17.

Maguire, D. J., M. Batty & M. F. Goodchild. 2005. *GIS, Spatial Analysis, and Modeling*. Redlands, CA: ESRI Press.

Manesh, H. F., M. Hashemipour & M. Bal (2008) Applications of virtual reality in computer integrated manufacturing systems. *27th Computers and Information in Engineering Conference, Vol 2, Pts a and B 2007*, 403-410.

Mendikoa, I., M. Sorli, J. I. Barbero, A. Carrillo & A. Gorostiza (2008) Collaborative product design and manufacturing with inventive approaches. *International Journal of Production Research,* 46, 2333-2344.

Mockus, J. (2006) Investigation of examples of E-education environment for scientific collaboration and distance graduate studies, Part 1. *Informatica,* 17, 259-278.

--- (2008) Investigation of examples of E-education environment for scientific collaboration and distance graduate studies. Part 2. *Informatica,* 19, 45-62.

Montello, D. R., J. S. Neil & B. B. Paul. 2001. Spatial Cognition. In *International Encyclopedia of the Social & Behavioral Sciences,* 14771-14775. Oxford: Pergamon.

Neo, H. K., Q. P. Lin & K. M. Liew (2005) A grid-based mobile agent collaborative virtual environment. *2005 International Conference on Cyberworlds, Proceedings*, 335-339.

Ocana, M., L. M. Bergasa, M. A. Sotelo & R. Flores (2005) Indoor robot navigation using a POMDP based on WiFi and ultrasound observations. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vols 1-4*, 503-508.

Odeyale, T. O. & V. Balogun (2006) Virtual reality as a tool for graphic presentation in architectural education. *eWork and eBusiness in Architecture, Engineering and Construction*, 597-603.

Ohno, N. & A. Kageyama (2007) Scientific visualization of geophysical simulation data by the CAVE VR system with volume rendering. *Physics of the Earth and Planetary Interiors,* 163, 305-311.

Okulicz, K. (2004) Virtual reality-based approach to manufacturing process planning. *International Journal of Production Research,* 42, 3493-3504.

Olson, J. S. & G. M. Olson. 1999. Computer supported cooperative work. In *Handbook of applied cognition,* eds. F. T. Durso, R. S. Nickerson & S. T. Dumais. John Wiley & Sons, Ltd.

Oppenraaij., R. H. F. v., A. H. J. Koning, B. A. Lisman, K. Boer, M. J. B. van den Hoff, P. J. van der Spek, E. A. P. Steegers & N. Exalto (2009) Vasculogenesis and Angiogenesis in the First Trimester Human Placenta: An Innovative 3D Study Using an Immersive Virtual Reality System. *Placenta,* 30, 220-222.

Ou, S. C., W. T. Sung & S. J. Hsiao (2002) Development of intelligent virtual reality web-based robotics for manufacturing applications. *Ieee Icit' 02: 2002 Ieee International Conference on Industrial Technology, Vols I and Ii, Proceedings*, 348-353.

Pappas, M., V. Karabatsou, D. Mavrikios & G. Chryssolouris (2006) Development of a web-based collaboration platform for manufacturing product and process design evaluation using virtual reality techniques. *International Journal of Computer Integrated Manufacturing,* 19, 805-814.

Pendergast, M. & S. Hayne (1999) Groupware and social networks: will life ever be the same again? *Information and Software Technology,* 41, 311-318.

Peterka, T., D. J. Sandin, J. H. Ge, J. Girado, R. Kooima, J. Leigh, A. Johnson, M. Thiebaux & T. A. DeFanti (2006) Personal Varrier: Autostereoscopic virtual reality display for distributed scientific visualization. *Future Generation Computer Systems-the International Journal of Grid Computing Theory Methods and Applications,* 22, 976-983.

Phillips, B. C. (2007) An education-service collaboration to address a perceived graduate RN readiness gap. *Nursing Outlook,* 55, 112-113.

Record, P. (2005) Teaching the art of fault diagnosis in electronics by a virtual learning environment. *Ieee Transactions on Education,* 48, 375-381.

Reif, R. & D. Walch. 2008. Augmented & Virtual Reality applications in the field of logistics. 987-994. Springer.

Rhee, S. M., J. Park & M. H. Kim. 2006. Multi-user live video in a shared virtual world for enhanced group collaboration. In *Technologies for E-Learning and Digital Entertainment, Proceedings*, 1131-1140. Berlin: Springer-Verlag Berlin.

Romain, C., B. Christian & R. Jerome (2006) Real time P2P network simulation for very large virtual environment. *DS-RT 2006: Tenth IEEE International Symposium on Distributed Simulation and Real-Time Applications, Proceedings*, 35-42.

Ruthenbeck, G. S., H. Owen & K. J. Reynolds (2008) A virtual reality throat examination simulation. *Stud Health Technol Inform,* 132, 433-5.

Sampaio, A. Z., P. G. Henriques & P. S. Ferreira (2006) Virtual reality models used in civil engineering education. *Proceedings of the IASTED International Conference on Internet and Multimedia Systems and Applications*, 119-124.

Sandoval-Ruiz, C. A., J. L. Zumaquero-Rios & O. R. Rojas-Soto (2008) Predicting geographic and ecological distributions of triatomine species in the southern Mexican state of Puebla using ecological niche modeling. *Journal of Medical Entomology,* 45, 540-546.

Santoro, F. M., M. R. S. Borges & E. A. Rezende (2006) Collaboration and knowledge sharing in network organizations. *Expert Systems with Applications,* 31, 715-727.

Schafer, W. A., C. H. Ganoe & J. M. Carroll (2007) Supporting Community Emergency Management Planning through a Geocollaboration Software Architecture. *Computer Supported Cooperative Work: The Journal of Collaborative Computing,* 16, 501-537.

Schepers, J., A. de Jong, M. Wetzels & K. de Ruyter (2008) Psychological safety and social support in groupware adoption: A multi-level assessment in education. *Computers & Education,* 51, 757-775.

SEDRIS. 2008. SEDRIS Technologies-The Source for ENvironmental Data Representation & Interchange. http://www.sedris.org/.

Shin, Y. (2005) Conflict Resolution in Virtual Teams. *Organizational Dynamics,* 34, 331-345.

Shu, Y., G. Zhu & C. Chen (2004) Generation of Virtual Forest(虚拟森林场景的构建). *Geomatics and Information Science of Wuhan University(武汉大学学报(信息科学版)),* 29.

Sidlar, C. L. & C. Rinner (2008) Utility assessment of a map-based online geo-collaboration tool. *Journal of Environmental Management,* In Press, Corrected Proof.

Sousa Santos, B., P. Dias, A. Pimentel, J. W. Baggerman, C. Ferreira, S. Silva & J. Madeira (2009) Head-mounted display versus desktop for 3D navigation in virtual reality: a user study. *Multimedia Tools and Applications,* 41, 161-181.

Sportisse, B. (2007) A review of current issues in air pollution modeling and simulation. *Computational Geosciences,* 11, 159-181.

Strouse, J. (2004) Meeting of minds - A collaboration between friends gives way to an expansive Florida apartment (Interior architecture by Brian O'Keefe, interior design by Marjorie Shushan for Carol Soffer's high-rise in Aventura). *Architectural Digest,* 61, 244-+.

Sun, J. (1999) An Exploration of Virtual Recreation Environment on Resources and Environmental Sciences(资源环境科学虚拟创新环境的探讨). *Resources science(资源科学),* 21.

Sun, Q. Z. (2008) Application of virtual reality technology in football education field. *Proceedings of First Joint International Pre-Olympic Conference of Sports Science and Sports Engineering, Vol Iii*, 288-293.

Taesombut, N., X. Wu, A. A. Chien, A. Nayak, B. Smith, D. Kilb, T. Im, D. Samilo, G. Kent & J. Orcutt (2006) Collaborative data visualization for Earth Sciences with the OptIPuter. *Future Generation Computer Systems-the International Journal of Grid Computing Theory Methods and Applications,* 22**,** 955-963.

Tandler, P. (2004) The BEACH application model and software framework for synchronous collaboration in ubiquitous computing environments. *Journal of Systems and Software,* 69**,** 267-296.

Tann, W. & H. J. Shaw (2007) The collaboration modelling framework for ship structural design. *Ocean Engineering,* 34**,** 917-929.

ter Beek, M. H., S. Gnesi, D. Latella, M. Massink, M. Sebastianis & G. Trentanni (2009) Assisting the design of a groupware system. *Journal of Logic and Algebraic Programming,* 78**,** 191-232.

Thompson, J. A., S. E. Carozza & L. Zhu (2008) Geographic risk modeling of childhood cancer relative to county-level crops, hazardous air pollutants and population density characteristics in Texas. *Environmental Health,* 7**,** 14.

Thrush, E. A. & M. Bodary (2000) Virtual reality, combat, and communication. *Journal of Business and Technical Communication,* 14**,** 315-327.

Topaloglu, U. & C. Bayrak (2008) Secure mobile agent execution in virtual environment. *Autonomous Agents and Multi-Agent Systems,* 16**,** 1-12.

Urquhart, C. J., A. M. Cox & S. Spink (2007) Collaboration on procurement of e-content between the National Health Service and higher education in the UK. *Interlending & Document Supply,* 35**,** 164-170.

van Dam, A., D. H. Laidlaw & R. M. Simpson (2002) Experiments in immersive virtual reality for scientific visualization. *Computers & Graphics-Uk,* 26**,** 535-555.

Varela, J. M. & C. G. Soares (2007) A virtual environment for decision support in ship damage control. *Ieee Computer Graphics and Applications,* 27**,** 58-69.

Wang, B. (2007) The participial virtual reality game and the development of Chinese family sports. *Proceedings of UK-China Sports Engineering Workshop***,** 137-143.

Wang, C. H. & S. Y. Chou (2008) Entities' representation modes and their communication effects in collaborative design for SMEs. *International Journal of Advanced Manufacturing Technology,* 37**,** 455-470.

Wang, W., L. Yu, W. Pei & F. Yang (2004) Review of Emission Dispersion Models Based on Gaussian Line-Source Model(基于高斯线源模式的主要尾气扩散模型综述). *Traffic Environment Protection(交通环保)*, 25.

Wang, X., Y. Liu & J. Zhang (2005a) Geo-Spatial Cognition:An Overview(地理空间认知综述). *Geography And Geo-Information Science(地理与地理信息科学)*, 21, 1-10.

Wang, X. M., G. Carmichael, D. L. Chen, Y. H. Tang & T. J. Wang (2005b) Impacts of different emission sources on air quality during March 2001 in the Pearl River Delta (PRD) region. *Atmospheric Environment*, 39, 5227-5241.

Wegman, E. J. (2000) Affordable environments for 3D collaborative data visualization. *Computing in Science & Engineering*, 2, 68-+.

Wen, Z. (2006) The suspended particle in the air became the city environmental damage in the Pearl River Delta(灰霾:珠三角城市公害). *Environment(环境)*, 3, 30-33.

Wenshi, L., D. Dongsheng, Z. Liping, F. Qi, W. Xuemei & W. Di. 2009. Numerical Simulation of Sea-land Breeze and Plume Dispersion in the Complex Coastal Terrain of South China. Guangzhou, China.

Wu, D., X. Tie, C. Li, Z. Ying, A. K.-H. Lau, J. Huang, X. Deng & X. Bi (2005) An extremely low visibility event over the Guangzhou Region: a case study. *Atmospheric Environment*, 39, 6568-6577.

Wu, Z., G. Liu, G. Xu, G. Zhang & G. Zeng (2004) Appl ication of Gaussian Model in Analysis of Sulfur Dioxide Distribution Regularity in Urban District of Zhuzhou(高斯模式在株洲市城区 SO2 分布规律研究中的应用). *HuNan Nonferrous Metals(湖南有色金属)*, 20.

Xavier, S. C. C., V. C. Vaz, P. S. D'Andrea, L. Herrera, L. Emperaire, J. R. Alves, O. Fernandes, L. F. Ferreira & A. M. Jansen (2007) Mapping of the distribution of Trypanosoma cruzi infection among small wild mammals in a conservation unit and its surroundings (Northeast-Brazil). *Parasitology International*, 56, 119-128.

Xu, B., H. Lin & J. Gong (2006) Architecture of HLA Based Distributed Virtual Geographic Environment. *Geo-spatial Information Science*, 9.

Zampardi, V., P. Polidori, F. Venuti, S. Bavetta, C. Carollo, R. Di Stefano, A. Provenzani, M. G. Sidoti & H. J. Johnson (2007) Collaboration of clinical pharmacists in nursing education. *Pharmacy World & Science*, 29, 459-460.

Zeng, L., W. Lin, S. Situ & X. Wang. 2009. Numerical simulation of NOx diffusion in the Pearl River Delta(珠江三角洲地区 NOX 扩散的数值模拟). In *The First National Atmospheric Boundary Layer Physics and Chemistry Academic Seminar --LAPC 2008 Annual meeting(第一届全国大气边界层物理和大气化学学术研讨会暨 LAPC2008 年年会会议)*. XiaMen.

Zha, X. F. & H. Du (2006) Knowledge-intensive collaborative design modeling and support Part II: System implementation and application. *Computers in Industry,* 57**,** 56-71.

Zhang, J. 2006. Research on Monitor and Simulation of Schistosomiasis around Poyang Lake Based on Remote Sensing and Agent(基于遥感和智能体的鄱阳湖区血吸虫病监测和模拟). Beijing(北京): Institute of Remote Sensing Applications(中国科学院遥感应用研究所).

Zhou, B., J. Blum & A. Eskandarian (2005) Virtual reality visualization of microscopic traffic simulations. *Human Performance; Simulation and Visualization***,** 159-166.

Zhu, J., J. H. Gong, W. G. Liu, T. Song & J. Q. Zhang (2007a) A collaborative virtual geographic environment based on P2P and Grid technologies. *Information Sciences,* 177**,** 4621-4633.

--- (2007b) A collaborative virtual geographic environment based on P2P and Grid technologies. *Information Sciences,* 177**,** 4621-4633.

---END---