Final Report

Modeling, Monitoring and Fault Diagnosis of Spacecraft Air Contaminants

NASA Grant No. NAGW-4585

W. Fred Ramirez, Mikhail Skliar, Anand Narayan
Department of Chemical Engineering
University of Colorado
Boulder, Colorado 80309-0424

George W. Morgenthaler, Gerald J. Smith
Department of Aerospace Engineering
University of Colorado
Boulder, Colorado 80309-0429

Contents

Chapter 1

Introduction

1.1 Outline

Control of air contaminants is a crucial factor in the safety considerations of crewed space flight. Indoor air quality needs to be closely monitored during long range missions such as a Mars mission, and also on large complex space structures such as the International Space Station. This work mainly pertains to the detection and simulation of air contaminants in the space station, though much of the work is easily extended to buildings, and issues of ventilation systems.

Here we propose a method with which to track the presence of contaminants using an accurate physical model, and also develop a robust procedure that would raise alarms when certain tolerance levels are exceeded. A part of this research concerns the modeling of air flow inside a spacecraft, and the consequent dispersal pattern of contaminants. Our objective is to also monitor the contaminants on-line, so we develop a state estimation procedure that makes use of the measurements from a sensor system and determines an optimal estimate of the contamination in the system as a function of time and space. The real-time optimal estimates in turn are used to detect faults in the system and also offer diagnoses as to their sources.

This work is concerned with the monitoring of air contaminants aboard future generation spacecraft and seeks to satisfy NASA's requirements as outlined in their Strategic Plan document (Technology Development Requirements, 1996).

Most of this work was undertaken to satisfy the requirements of NASA's Advanced Environmental Monitoring and Control Program, with a view to developing an intelligent monitoring system for Space Station missions. The current NASA Strategic Plan has as one of its stated goals "to explore, use, and enable the development of space for human enterprise". The goal is to be accomplished in three time periods

- 1996-2002: Establish permanent presence in low Earth orbit by constructing and using the ISS,
- 2003-2009: Operate the ISS cost effectively, with a subgoal to "achieve advanced life support systems to close spacecraft air/water loops," and,
- 2010-2020 and beyond: Conduct international human missions to planetary bodies in our solar system.

Even though this work is targeted at future generation spacecraft and space stations, many of the specifications used in this work pertain to the International Space Station, scheduled for launch in late 1998. The International Space Station is a collaborative effort with participation by the governments of the United States, Canada, Europe, Japan, and Russia. The configuration will include a Hab and a Lab Element, two Nodes, and two International modules (the European Space Agency Attached Pressurized Module and the Japanese Experimental Module). Other relevant missions where this work applies include the manned missions to Mars, the Mars Short Visit, the Mars Human -Tended Outpost, and the Mars Permanent Outpost, where astronauts are expected to spend up to 600 days, and where the luxury of returning to earth for a cleanup in the case of a contaminant leak does not exist. Contaminants that are to be monitored include carbon dioxide, carbon monoxide, and volatile organics. According to NASA, primary chemicals of concern are nitrogen tetroxide, monomethyl hydrazine, ammonia, and Halon 1301. Studies aboard the Mir Station(Cole et al., 1996) have shown that about 45 compounds (32 target compounds and 13 non-target compounds) were consistently detected in air samples that were

collected during the missions, though none of the compounds were present at toxic levels. Still, contaminant monitoring remains an important focus area in ensuring the safety of humans in space.

Major sources of contaminants in the space station include off-gassing of cabin materials and hardware, use of utility chemicals, and metabolic waste products of crew members. Minor sources of contaminants include electrical equipment, microbial metabolism, leakage during experiments using chemicals, leakage from environmental or flight control systems, volatile food components, and reaction products from the environmental control and life support systems. Table 1.1 lists some of the substances being monitored and their sources aboard the spacecraft.

The National Research Council (National Research Council, 1984), in its various studies, has prescribed (National Academy of Sciences, 1981) spacecraft maximum allowable concentrations (SMACs), which are not to be exceeded under any circumstances (National Research Council, 1996). These concentrations are based on studies that link contaminant concentrations to the impairment of normal human activities.

The fault detection and diagnosis system is a synthesis of different mathematical procedures, which are functionally independent, but which come together to provide an effective structure that serves the purpose of monitoring the presence of airborne contaminants.

1.2 Space Station Environmental Control and Life Support System

In this section, the general layout of the International Space Station, and some of the components of the Air revitalization system are described.

Future missions, especially the long range missions, will increasingly have to be materially closed systems, since the cost of carrying spare oxygen would be prohibitive. Future missions will also involve growing food, and the complexity of the revitalization system will increase to account for many more possible contaminants.

Table 1.1: Some commonly monitored substances and their sources aboard spacecraft (National Research Council, 1992)

Monitored substance	Spacecraft source
Oxygen	Required component in cabin air
Carbon dioxide	Required component in cabin air
Carbon monoxide	Product of incomplete combustion
Nitrogen	Inert component in air
Halon	Diffusion from the Shuttle to the Station
1-Butanol	Off-gassing from flight hardware and from human metabolism
Diacetone alcohol	Off-gassing from paint that is not totally cured and from hardware off-gassing
Ethanol	Cleaner/disinfectant use
Ethyl benzene	Off-gassing from nonmetallic materials
Ethylene Glycol	antifreeze and off-gassing
Glutaraldehyde	Cell-Tissue fixatives
Trichloroethylene	Off-gassing
Xylene	Off-gassing

In addition, NASA's Advanced technology requirements outlines certain features that must be included in future life support systems. Some of those features that this work attempts to demonstrate include:

- Ambient air in the cabin must be monitored at selected locations, every 15 seconds, for the species O_2, CO_2, and CO.
- Toxicity of air must be reported in terms of specific major and trace species concentrations and their rates of rise.
- A computer model shall be available to predict the behavior and the contamination removal capabilities for contaminants that could suddenly be released into the atmosphere. The model must be experimentally verified and be capable of spatial resolution to the module level and temporal resolution to 0.5 hour.
- major air components shall be monitored on a near-continuous basis in the habitat atmosphere

The Environmental Control and Life Support system is divided into seven major subsystems, the temperature and humidity control (THC), atmosphere control and supply (ACS), atmosphere revitalization (AR), fire detection and suppression (FRS), water recovery and management (WRM), waste management (WM), and the Vacuum System (VS).

Their functions include (Reuter, 1998) :

- Atmosphere Revitalization:
  - Control and disposal of carbon dioxide
  - Control of airborne trace contaminants
  - Oxygen (O_2) supply via generation

  * Atmospheric monitoring of primary constituents, including O_2, CO_2, nitrogen (N_2), hydrogen (H_2), methane (CH_4), and water vapor
  * Airborne particulate and microbial control

- Temperature and Humidity Control
  * Cabin air temperature and humidity control
  * Equipment air cooling
  * Inter- and intra-module ventilation for crew comfort and station level control of O_2, CO_2, and trace contaminants

- Atmosphere Control and Supply
  * Total pressure and O_2 partial pressure control
  * Total pressure monitoring and monitoring of loss of pressure
  * Stored gaseous N_2 and O_2 supply and replenishment

- Fire Detection and Suppression
  * Smoke detection
  * Fire extinguishment

- Waste Recovery and Management
  * Potable and hygienic water supply
  * Waste water and urine water collection, recovery, and disposal

- Waste Management:
  * Urine/fecal collection and recovery

- Vacuum System
  * Vacuum venting and maintenance for payload support

Some of the air and water quality requirements which are to be maintained in the ISS are shown in Table 1.2.

A test of the Life Support System was carried out in the Johnson Space Center's Integrated Life Support System Test facility, in which a four-member crew spent thirty continuous days in a closed chamber. All test objectives were accomplished (Lunar-Mars, 1997), and no SMACs were violated.

1.3 Previous work

Space environment monitoring has been in place for as long as there have been crewed missions, with the levels of sophistication changing with the complexity and duration of the missions, and along with the developments in computational and sensor technologies. Two computer models that represent the present generation of Space Environment monitoring models are the Trace Contaminant Control Simulation (TCCS) (Perry, 1993) and the Computer Aided System Engineering and Analysis (CASE/A) (CASE/A, 1990) modeling package. Both the packages model the space station modules as well-stirred tanks, where each module is represented by its average concentration. The CASE/A package is more flexible and user-friendly, and provides a means for simulation of a number of interconnected well-stirred tanks. Both the packages suffer from the limitation that stagnation points within the cabin, and non-uniform forced convection patterns cannot be represented within the model. A study of the inhalation risks aboard spacecraft (Todd et al., 1994), where the Space Station was modeled as a series of well stirred chemical reactors, improved on the lumped analysis by providing more information about spatial variations of contaminants aboard spacecraft. A comprehensive study of the sensor system was undertaken (Smith, 1996), which used a lumped system of modeling the transport in order to optimize the sensor selection process.

The first significant shift away from lumped analysis in contaminant dispersal modeling was in the work of Skliar and Ramirez (Skliar and Ramirez, 1997a),

Table 1.2: Air quality requirements for the ISS(Bacskay and DaLee, 1991)

Parameter	Units	Operational	Degraded	Emergency
CO_2 Partial Pressure	N/m^2	400 MAX	1013 MAX	1600 MAX
O_2 Partial Pressure	N/m^2 x 10^3	19.5-23.1	16.5 - 23.7	15.8 - 23.7
Total Pressure	N/m^2 x 10^3	99.9 - 102.7	99.9 - 102.7	99.9 - 102.7
Temperature	^oK	291.5 - 299.8	291.5 - 299.8	288.8 - 302.6
Dew point	^oK	277.6 - 288.8	274.8 - 294.3	274.8 - 294.3
Ventilation	m/sec	0.076 - 0.203	0.051 - 0.508	0.051 - 1.016
Airborne Trace Contaminants	mg/m^3	SMAC	SMAC	SMAC
Airborne Micro-organisms	CFU/m^3	1000 (max)	1000 (max)	1000 (max)
Airborne particulates	$Particles/m^3$	3,530,000 $\geq$ 0.5 micrometers		

where each cabin module was represented using a two dimensional mesh and a solution to the reaction-diffusion equation was obtained using a finite difference scheme. This approach has the advantage of providing information in two spatial directions.

Blackwell (Blackwell, 1998) has proposed a a fault detection and location procedure based on physical laws and modeling, analysis, and simulation. His work is based on optimal control theory, and uses analytical redundancy for the detection part, but the work does not specifically address the physics of the problem, and merely lays out the structure for the procedure.

1.4 Salient characteristics of this research

This research seeks to build on the two-dimensional model developed by Skliar (1996), and extend that work to three dimensions. It also seeks to use a more rigorous computational fluid dynamics solution to the flow equations that what has been used hitherto for this purpose, one that vastly reduces the modeling error. A fault detection algorithm is implemented with the ability to distinguish between process and sensor faults. The fault detection algorithm uses the concept of analytical redundancy to detect the faults.

The final part of the research focuses on diagnosing the fault in the system, primarily that of an unknown contaminant source in the cabin. This is an inverse problem that is ill-posed and has no unique solution, so the attempt here will be to obtain a satisfactory solution to a required degree of accuracy. To this end, an Extended Implicit Kalman Filter is developed, which is an extension of the Implicit Kalman Filter suitable for joint state and parameter estimation. The filter essentially maintains the same structure as the original filter, and many of the algorithms remain the same as before. The filter requires an initial guess for the location and capacity of the unknown source, for which purpose we use pre-calculated sensitivity matrices that contain information about the local variation in concentration for perturbations throughout the cabin. The combination of these two techniques makes

for an accurate and efficient algorithm for obtaining a solution to the inverse problem.

The proposed procedure for air contaminant monitoring is shown in Fig. 1.1. The same procedure can be applied for a variety of substances,, and some algorithms can even be used to monitor air pressure and some other parameters.

1.5 Outline of Report

This report is organized as follows. Chapter 2 discusses the flow modeling work that we performed in order to simulate the air flow aboard the International Space Station, and Chapter 3 describes the mathematical modeling of the contaminant dispersal process along with its numerical solution, and sample profiles that were obtained. The State estimation procedure is discussed in Chapter 4, and the use of the State Estimation Procedure for fault detection is described in Chapter 5. Chapter 6 contains the discussion of the fault diagnosis algorithms that are proposed, and finally Chapter 7 presents the conclusions and the significant contributions of this research.

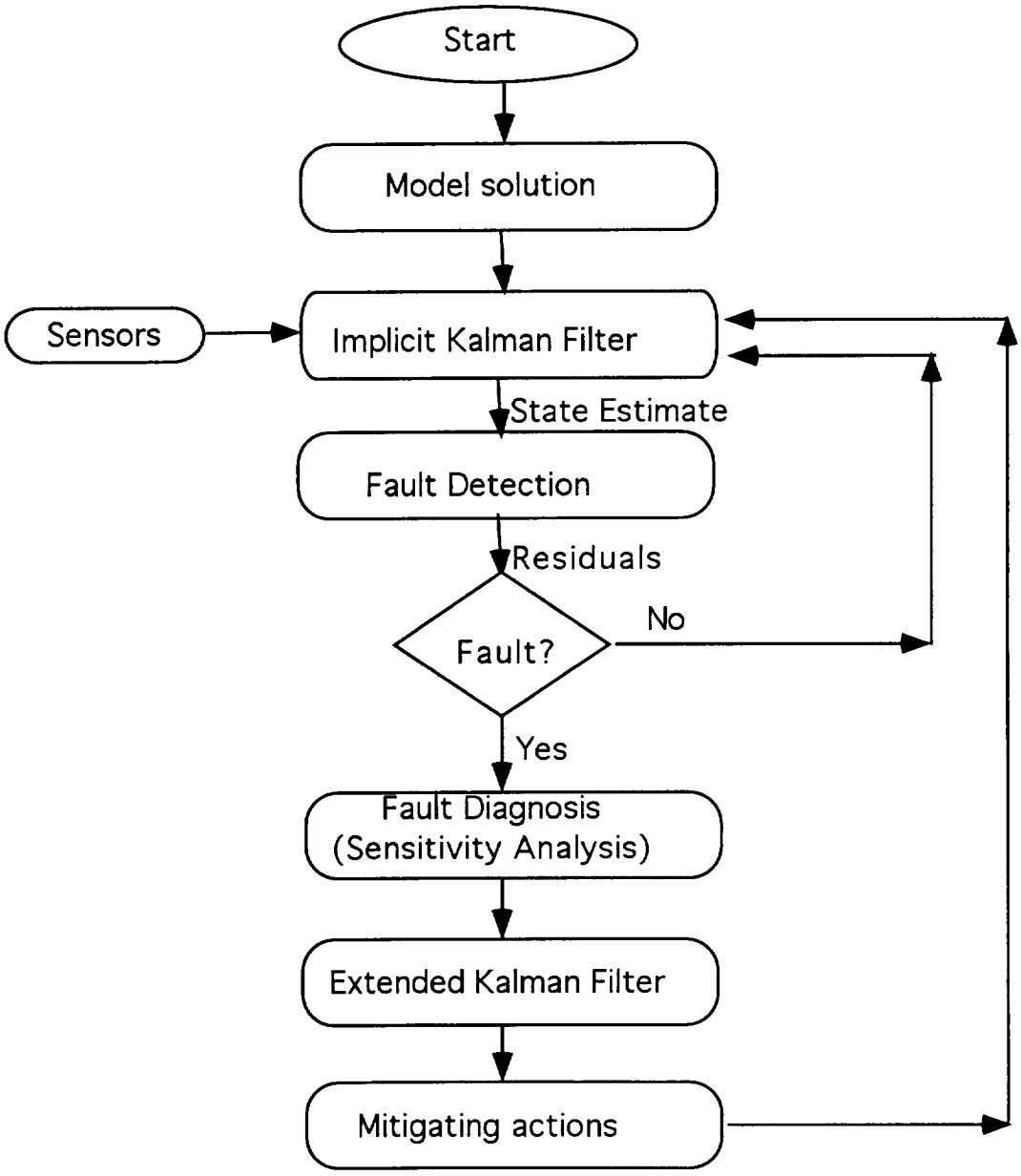

Figure 1.1: Air contaminant monitoring system

Chapter 2

Indoor air flow

2.1 Literature review

Air flows inside enclosures have been a subject of active interest within the building systems community. Computational fluid dynamics (CFD) has been used for predicting room air movement since the seventies. There even exists a public domain software (Kurabuchi et al., 1990; Said et al., 1995) called EXACT3, which is a three dimensional finite difference computer program for simulating buoyant turbulent airflow within buildings. In recent years, much effort has been made to enhance CFD as a reliable tool for the evaluation of air flows. CFD has been used in studying clean room air-flows (Kuehn, 1988; Yamamoto et al., 1988) because of the need to keep the clean room free of particles and particulates, and air flow becomes a critical parameter in such cases. Space application in the context of contaminant dispersion is very similar in scope to clean room applications, though there is little work reported in the literature that pertains to space applications. Recently (Tam, 1998), an interesting study evaluated the application of CFD in the software design of environmental control and life support systems, and found that atmospheric conditions within the Space Station could be adequately modeled using the Resource Utilization Planning and System Modeling (RUPSM) scheme.

Most of the work in the area has remained computational, though a few validating experimental results also exist in the literature. The International Energy

Agency, though a subsidiary research group, Annex 20, measured velocities, temperatures, and turbulence velocity scales in full scale rooms (Whittle and Clancy, 1991). Canadian researchers (Barber et al., 1982) did studies on the correlation between the velocity of the inlet jet and the floor velocity, and proposed a jet momentum number that would measure the energy contained in the diffuser jet relative to the room air volume. In this work, we use CFD to provide us with the information regarding the flow which is then used as an input to the mathematical model for the diffusion and for the procedure that estimates the current concentration of contaminants in the cabin. This development marks an important step in our estimation procedure since the accuracy of the procedure is largely dependent on the accuracy of the flow model since most of the transport is occurring through convective diffusion.

A detailed knowledge of the flow field is required in order to ensure that the ventilation system is performing adequately, and to provide information about local variations in the concentration profile of the contaminants. Another advantage in using CFD is that it enables the calculation of quantities like turbulence intensity which have direct effects on the comfort level of people inside the cabin. CFD is an inexpensive tool for such studies and has been used to study the effect of different ventilation techniques (Gan, 1995) on thermal comfort in buildings. Research has shown that lower turbulence intensities contribute to higher comfort levels (Zhang et al., 1992).

2.2 Air Flow Modeling

We assume that the air flow is steady and incompressible, and solve the three-dimensional Navier-Stokes equations for the appropriate boundary and initial conditions. Air under atmospheric pressure and for the low velocity flows that are characteristic of room air flows is expected to be incompressible, and we invoke the steady state assumption because solving for transient cabin flows is computationally too intensive to be used in a real-time operation. We tried simulating the air flow

for both 2-D and 3-D cases. Earlier work used a two-dimensional flow field (Skliar and Ramirez, 1997a) as an input to their state estimation procedure. In a significant development in this work, we consider modeling the three-dimensional geometry of the space station module. The 2-D case has the advantage of requiring far less computational time, whereas it suffers from a lack of information about the third dimension.

Air flows inside enclosures are usually turbulent, random, and highly recirculating (Zhang et al., 1992). In this work, we solve the equations for both laminar and turbulent cases. The geometry chosen here follows experimental test used in previous studies (Son and Barker, 1997) and accurately represents the US Space Station Lab module. We used this geometry for our simulations in order to have an experimental set of results to validate our simulations. The simulated cabin (Fig. 2.1) is 6 m long, 2 m wide, and 2 m high (approximately. 20' X 7' X 7'). There are two inlets and two outlets for the air. The Temperature and Humidity Control (THC) is the primary air supply which supplies regulated air into the cabin and is one of the primary subsystems in the Environmental Control and Life Support Systems for the Space Station (see Son and Barker, 1993). The Intermodule ventilation (IMV) airflow assemblies are used to interchange airflow between modules. One would expect that the THC air is relatively contaminant free since it is filtered, whereas the IMV could have trace contaminants generated in other modules, both routine contaminants and those released due to accidents.

The air-flow model is based on the continuity equation, the Navier-Stokes equation, thermal energy equation, and the concentration equation together with the $k-\epsilon$ turbulence model equations, for the case of turbulent flow. The $k-\epsilon$ model (Anderson et al., 1984; Whitaker, 1968) uses the kinetic energy of turbulence k, and its dissipation rate ϵ as the two scales. This introduces two additional transport equations. In the $k-\epsilon$ model, the turbulent viscosity μ_t also known as the eddy

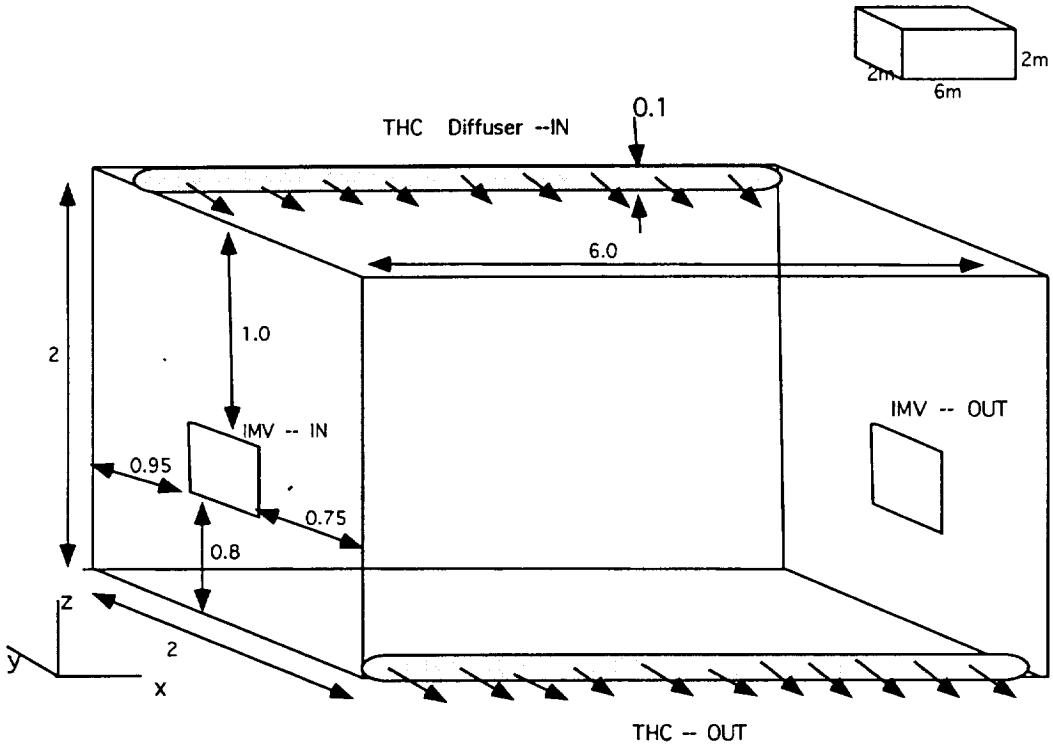

Figure 2.1: Sketch of the model cabin.

viscosity is defined by the relation:

$$\mu_t = c_\mu \rho k^2 / \epsilon \tag{2.1}$$

where c_μ is an empirical constant, and ρ is the density of the fluid.

For a rectangular cabin geometry, the equations of continuity, momentum, energy and mass for an incompressible flow are as follows. The symbols in the equations are defined in the nomenclature section at the end of this thesis.

Continuity:

$$\nabla \cdot \mathbf{u} = 0 \tag{2.2}$$

Momentum:

The momentum equation for the three co-ordinate directions is given below.

x-direction

$$\frac{\partial(\rho u)}{\partial t} + \frac{\partial(\rho u^2)}{\partial x} + \frac{\partial(\rho uv)}{\partial y} + \frac{\partial(\rho uw)}{\partial z} = \tag{2.3}$$
$$-\frac{\partial p}{\partial x} + \frac{\partial}{\partial x}(\lambda \nabla \cdot \mathbf{u} + 2\mu \frac{\partial u}{\partial x}) + \frac{\partial}{\partial y}[\mu(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y})] + \frac{\partial}{\partial z}[\mu(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x})] + \rho f_x$$

y-direction

$$\frac{\partial(\rho v)}{\partial t} + \frac{\partial(\rho uv)}{\partial x} + \frac{\partial(\rho v^2)}{\partial y} + \frac{\partial(\rho vw)}{\partial z} = \tag{2.4}$$
$$-\frac{\partial p}{\partial y} + \frac{\partial}{\partial x}[\mu(\frac{\partial v}{\partial z} + \frac{\partial u}{\partial y})] + \frac{\partial}{\partial y}(\lambda \nabla \cdot \mathbf{u} + 2\mu \frac{\partial v}{\partial y}) + \frac{\partial}{\partial z}[\mu(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z})] + \rho f_y$$

z-direction

$$\frac{\partial(\rho w)}{\partial t} + \frac{\partial(\rho uw)}{\partial x} + \frac{\partial(\rho vw)}{\partial y} + \frac{\partial(\rho w^2)}{\partial z} = \tag{2.5}$$
$$-\frac{\partial p}{\partial z} + \frac{\partial}{\partial x}[\mu(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x})] + \frac{\partial}{\partial y}[\mu(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z})] + \frac{\partial}{\partial z}(\lambda \nabla \cdot \mathbf{u} + 2\mu \frac{\partial w}{\partial z}) + \rho f_z$$

Energy:

The energy equation is given below for the sake of completeness. The Space Station environment is expected to be isothermal for the most part, and minor temperatures would not affect the flow field much since buoyancy is not an important factor under micro-gravity.

$$\rho\frac{D}{Dt}(e+\frac{\mathbf{u}^2}{2}) = \rho\dot{q} + \frac{\partial}{\partial x}(k\frac{\partial T}{\partial x}) + \frac{\partial}{\partial y}(k\frac{\partial T}{\partial y}) + \frac{\partial}{\partial z}(k\frac{\partial T}{\partial z}) \tag{2.6}$$
$$-p(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}) + \lambda(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z})^2 +$$
$$\mu[2(\frac{\partial u}{\partial x})^2 + 2(\frac{\partial v}{\partial y})^2 + 2(\frac{\partial w}{\partial z})^2 + (\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x})^2 + (\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x})^2 + (\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y})^2]$$

2.3 Solving the Navier-Stokes equations

The equations were solved in a coupled manner using the Fluid Dynamics Analysis Package (FIDAP Version 7.62)(FIDAP, 1993). A finite element mesh grid was developed for the two-dimensional and the three-dimensional problems with specified nodal boundary conditions. An eight-node brick was used as the basic finite element for the purpose of discretization. The velocity components were approximated using trilinear interpolation functions within the elements. The pressure was discretized in a piecewise continuous manner, with the pressure degree of freedom associated with the element centroid. A segregated solver was used to solve the resulting non-linear equations. The segregated solver decouples the equations for the purpose of solution, and sequentially solves them, using the results of one equation in the next, and so on. This increases the CPU time needed, but conserves memory, and has been found to be very useful for large mesh sizes. A variety of boundary conditions was tried, though for the sake of conciseness, only two cases will be discussed here. The Convergence criterion required the residuals of the equations (velocity components and the pressure) to be below 0.0001. Most of the computations were

performed on the SGI Power Challenge Array, in a parallel mode with either two or three processors, available to us through the National Center for Supercomputing Applications, Urbana-Champaign. The runs required varying times to converge, approximately in the range of 60 CPU hours for each.

2.4 Flow field results

Case 1: Laminar Flow- high flow rate

The geometry has already been described. For this case, we used an IMV with a duct velocity of 4.8 m/s and a THC with a velocity of 6.0 m/s. The cabin is assumed to be isothermal. The ducts leading to the outlet are modeled using free boundary conditions, i.e. the values of the velocities are allowed to float to satisfy the Navier-Stokes equations. The no slip boundary condition was invoked at all the walls. We assume that the flow is laminar, and that it is a steady flow field.

Figures 2.2 - 2.4 illustrate the flow profiles that we obtained. The contour of the magnitude of the velocities and the velocity vectors themselves are depicted. We show three horizontal slices of the box, one near the top and bottom, and one halfway up the cabin to illustrate the variations. The slice near the bottom is mainly dominated by the exit of the THC duct. Note that the flow leaves at an angle to the duct due to the blast of air that blows in the x-direction. The slice near the center clearly shows the profile near the inlet and outlet for the IMV ducts. The flow spreads out throughout the room. The laminar case shows no recirculation cells. The slice from the top of the cabin shows the flow entering the cabin, and the cells formed as the jet curves downward.

Case 2: Turbulent Flow- high flow rate

Turbulence, in a sense, is still an unsolved problem. The presence of a length and a time scale much smaller than the physical problem presents a scenario where the exact solution to equations cannot be obtained. In addition, for room and cabin flows, it is difficult to predict the onset of turbulence. A statistical approach

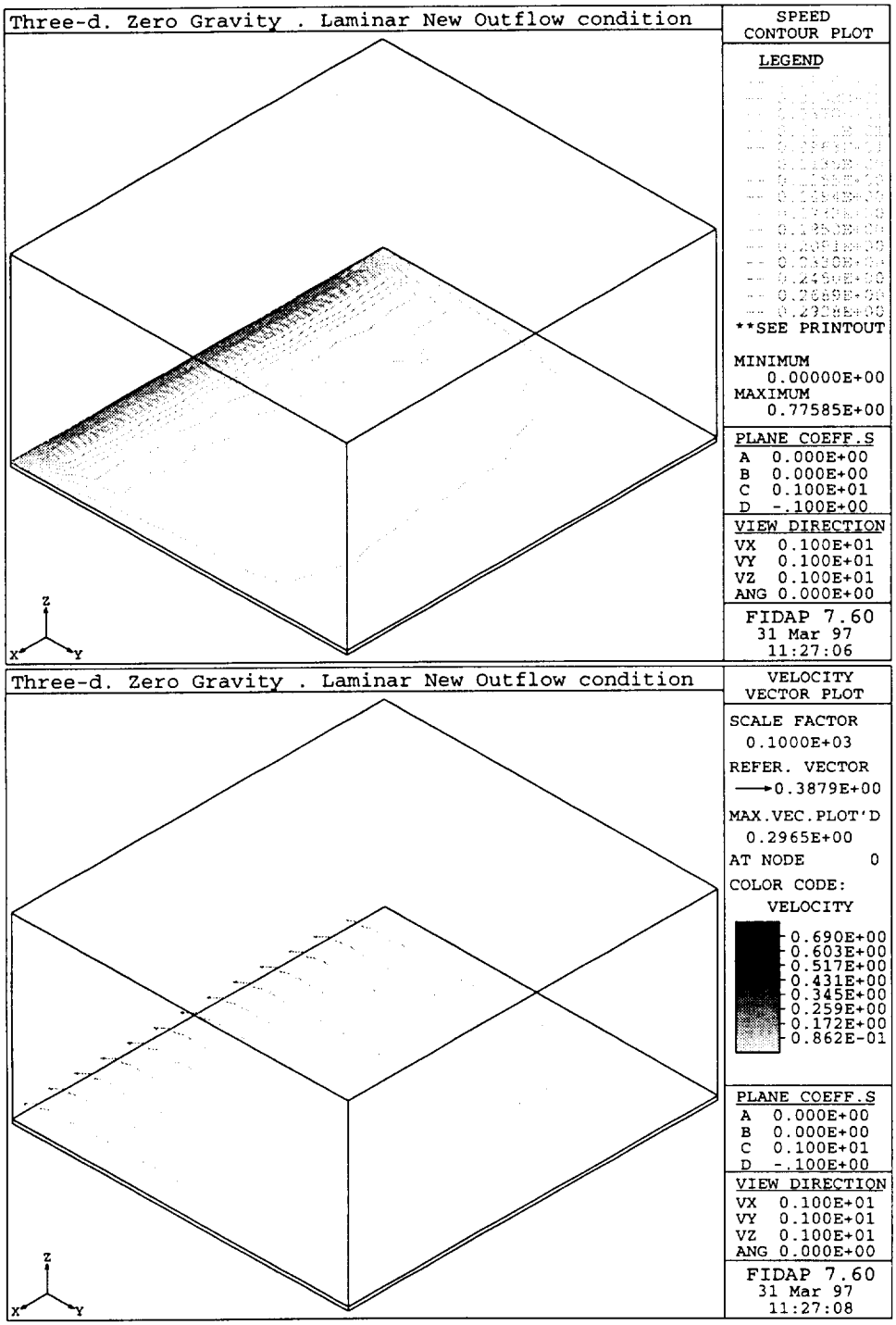

Figure 2.2: Speed Contour and Velocity Vector for the bottom slice under laminar flow conditions for a THC flow of 6.0 ms^{-1} and an IMV flow of 4.8ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

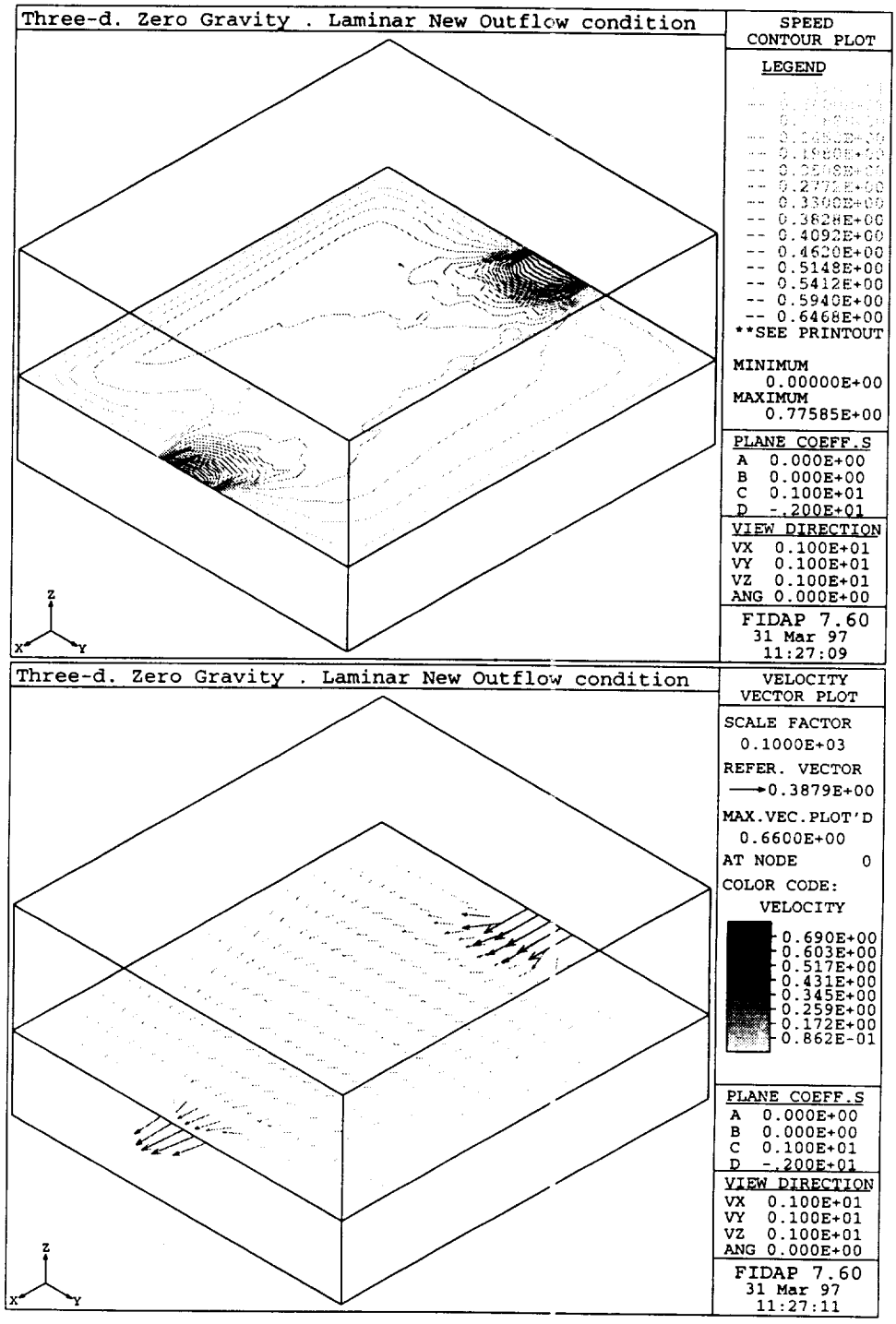

Figure 2.3: Speed Contour and Velocity Vector for the middle slice under laminar flow conditions for a THC flow of 6.0 ms^{-1} and an IMV flow of 4.8ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

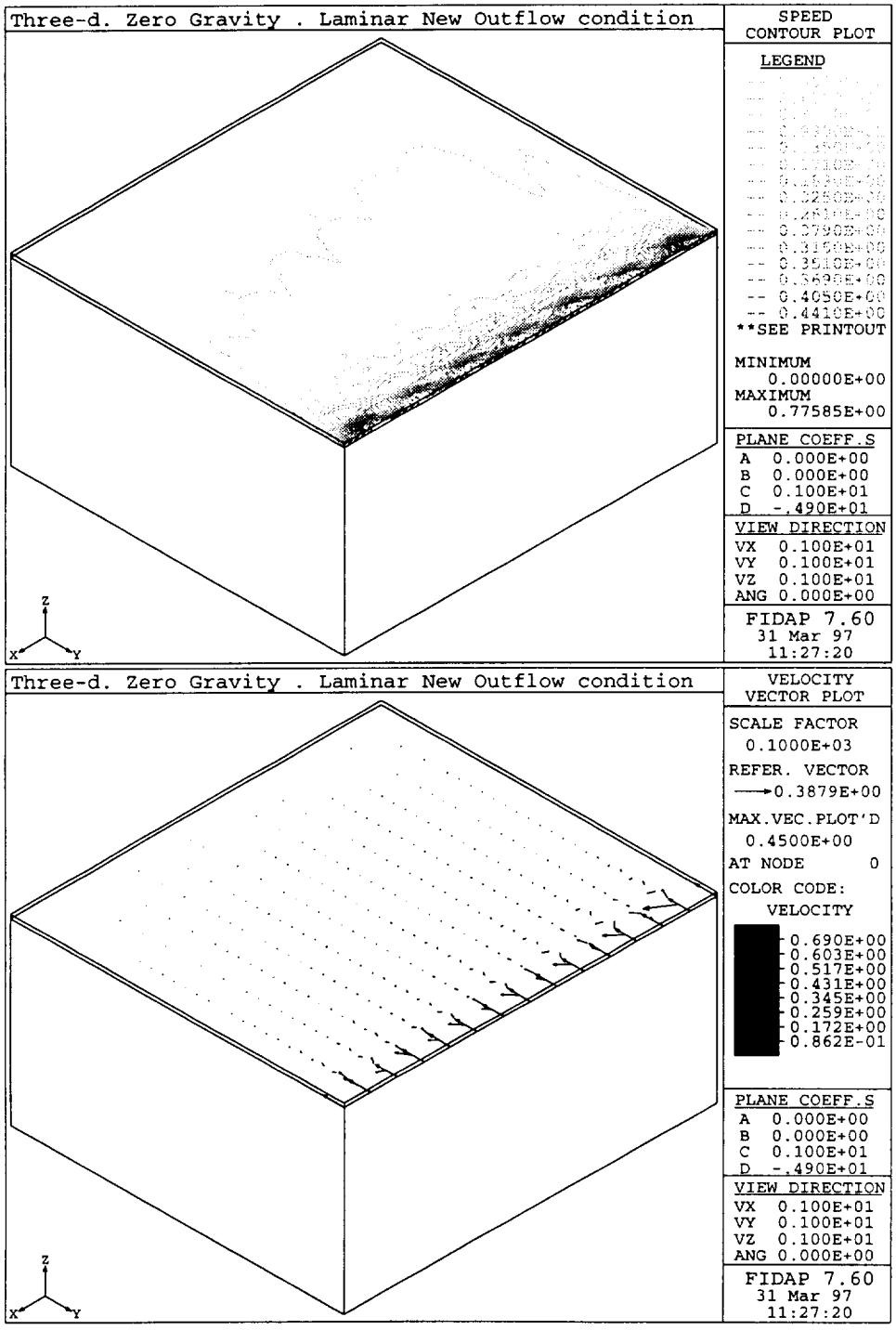

Figure 2.4: Speed Contour and Velocity Vector for the top slice under laminar flow conditions for a THC flow of 6.0 ms^{-1} and an IMV flow of 4.8ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

is usually used, and the equations are averaged over a time scale that is long compared to that of turbulent motion. The resulting averaged equations then describe the distribution of the mean velocity, pressure, temperature and the other quantities of interest. Detailed derivations of the equations can be found in any advanced book on fluid motion (Anderson et al., 1984). We use the two-equation model, briefly touched upon earlier. For isothermal flow with no mass transfer, the recommended set of model parameters were used for these empirical constants defined in the nomenclature section.

$$c_\mu = 0.09, \sigma_k = 1.00, \sigma_\epsilon = 1.30, c_1 = 1.44, c_2 = 1.92 \tag{2.7}$$

Figures (2.5-2.7) are graphical representations of the turbulent flow simulations.

Case 3: Laminar Flow -Low flow rate

For the same geometry as before, the flow rates were decreased, and the steady state flow profiles were recalculated for an IMV flow rate of 0.15 m/s and a THC flow rate of $0.3m/s$. All other conditions were maintained at previous levels. The profiles are shown in Fig. 2.8-2.10.

Case 4: Turbulent Flow -Low flow rate

The low velocity flow field calculation was repeated for turbulent flow. Figures 2.11-2.13 show the contours for the profiles that were obtained.

No major differences were noticed in the flow profiles obtained for high and low velocity duct flows. The patterns of flow essentially remained the same. A more thorough study of inlet velocities and their effect on cabin air flows is in order but beyond the scope of this work. The turbulent flow profiles closely resemble those obtained previously in experiments (Son and Barker, 1997) in the Space Station simulator. The differences can be attributed to the minor differences in the geometry in the region of the hatches connecting the modules, and the roundedness of their hatches. The existence of recirculation cells is the significant difference between the

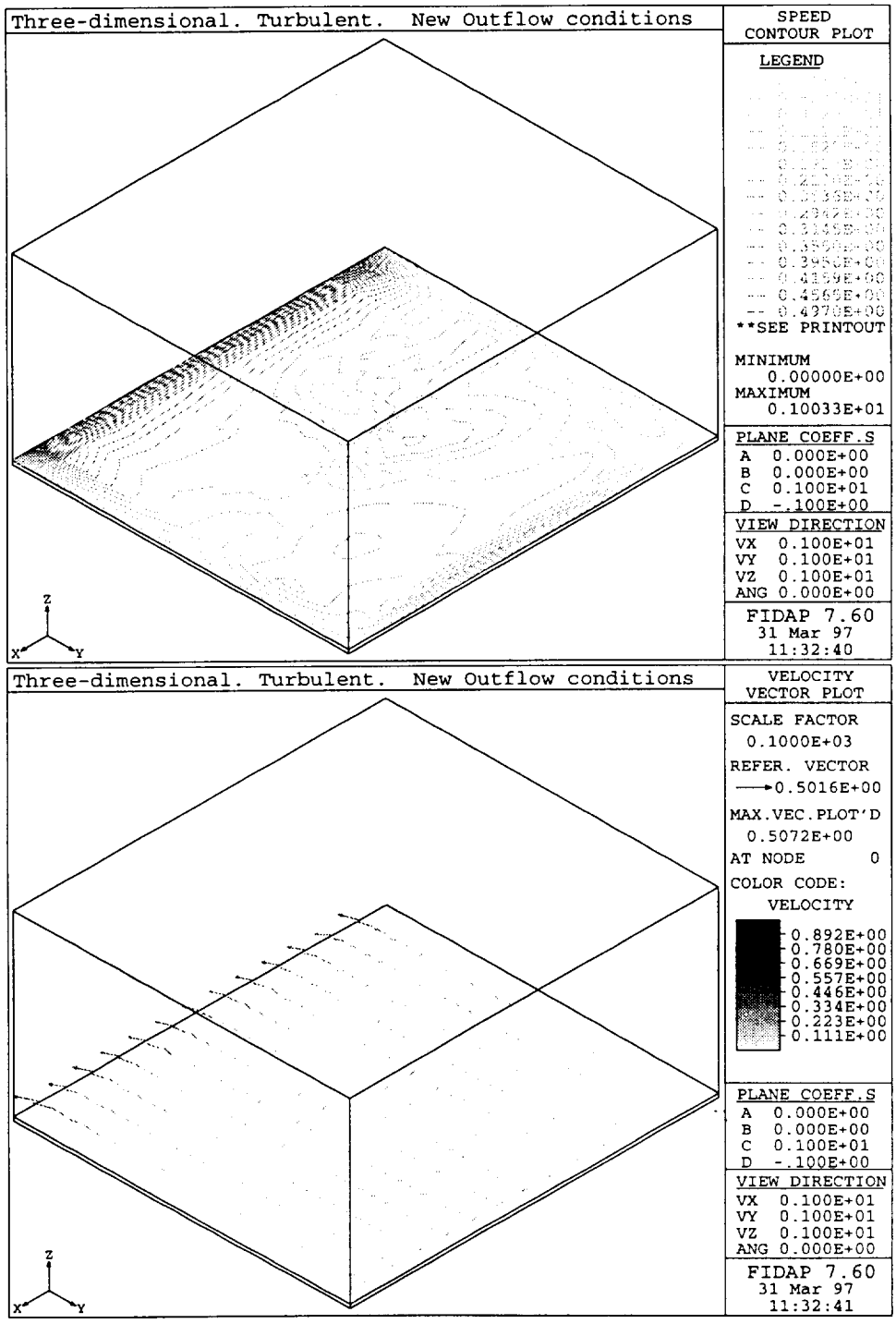

Figure 2.5: Speed Contour and Velocity Vector for the bottom slice under turbulent flow conditions for a THC flow of 6.0 ms^{-1} and an IMV flow of 4.8ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

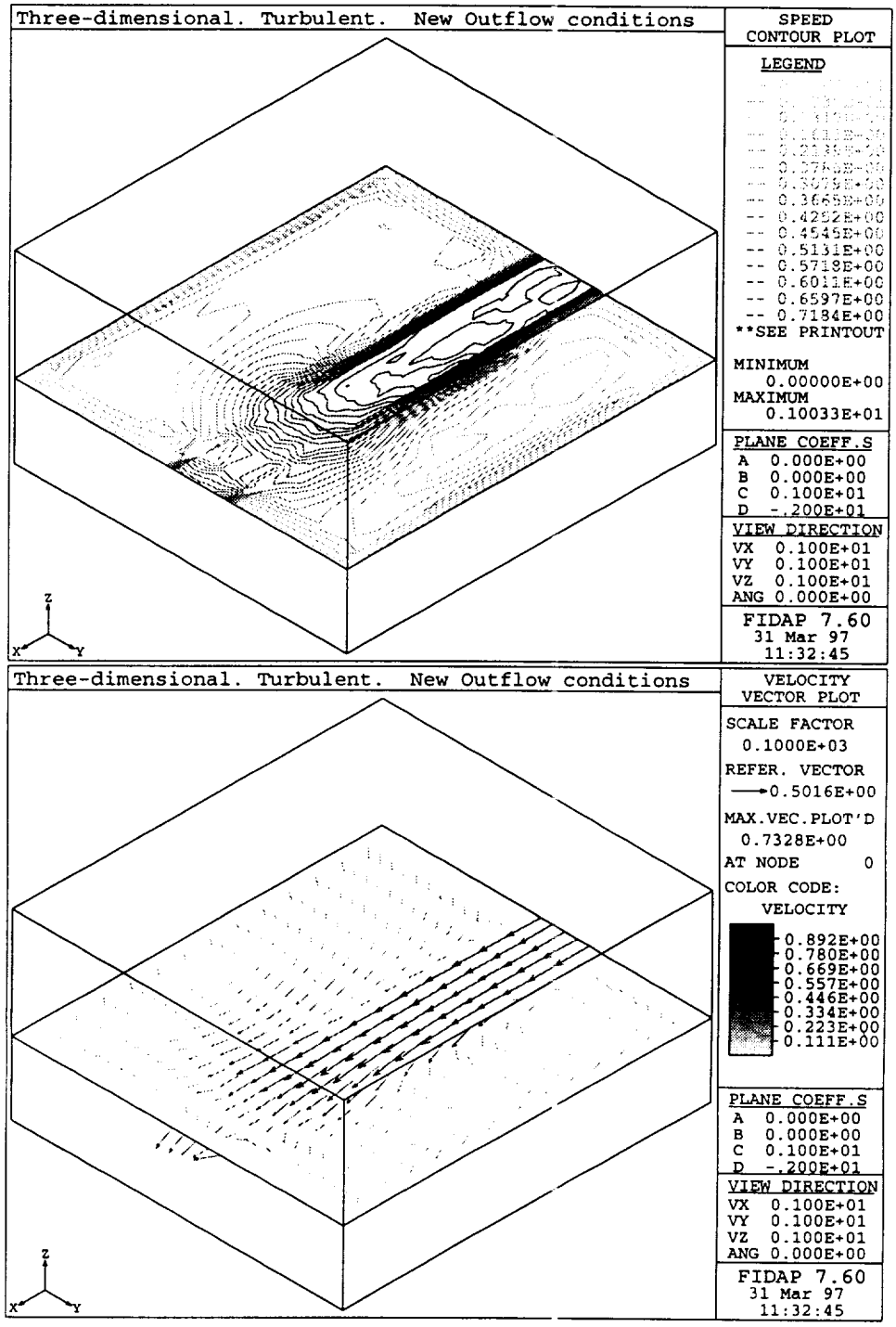

Figure 2.6: Speed Contour and Velocity Vector for the middle slice under turbulent flow conditions for a THC flow of 6.0 ms^{-1} and an IMV flow of 4.8ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

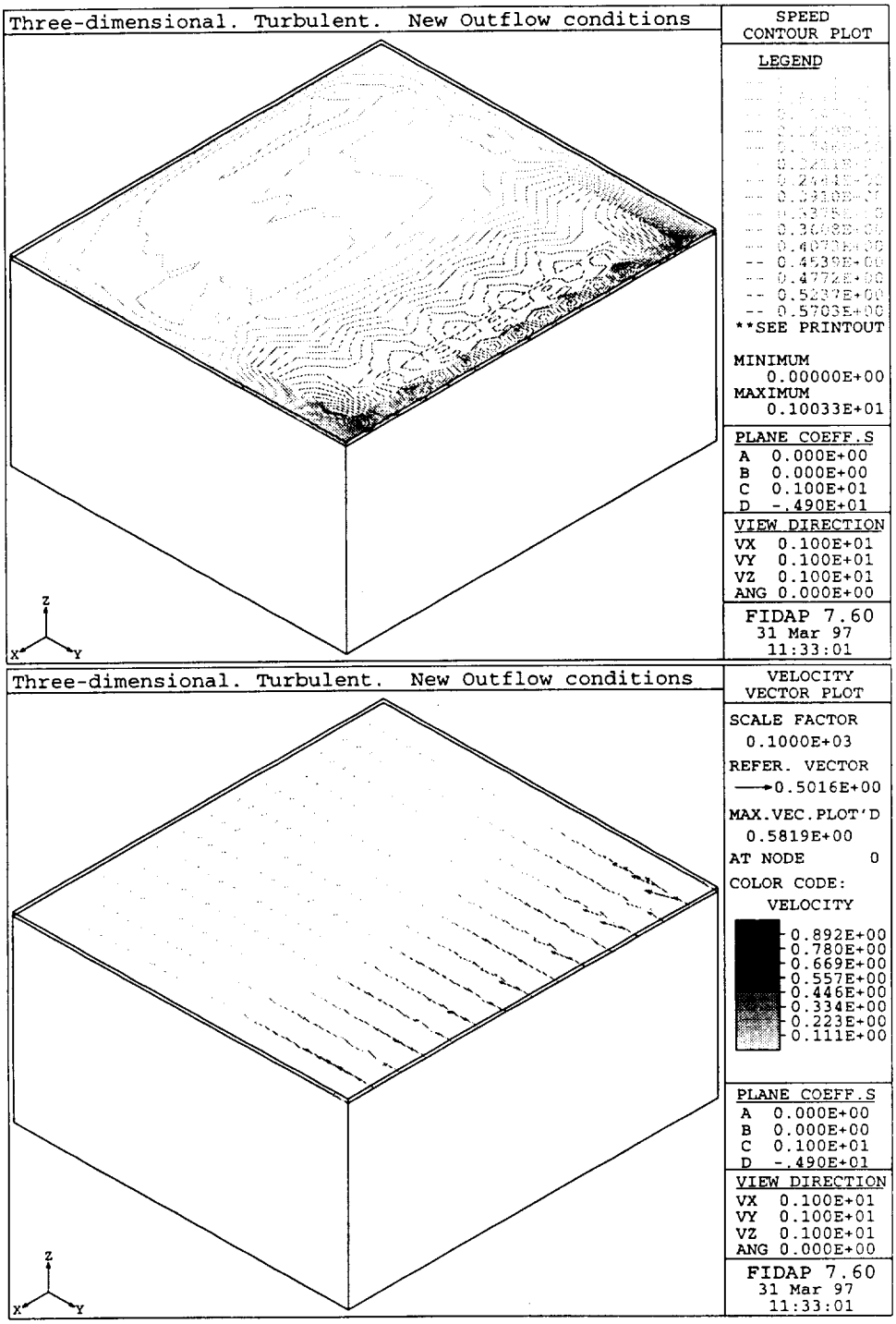

Figure 2.7: Speed Contour and Velocity Vector for the top slice under turbulent flow conditions for a THC flow of 6.0 ms^{-1} and an IMV flow of 4.8ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

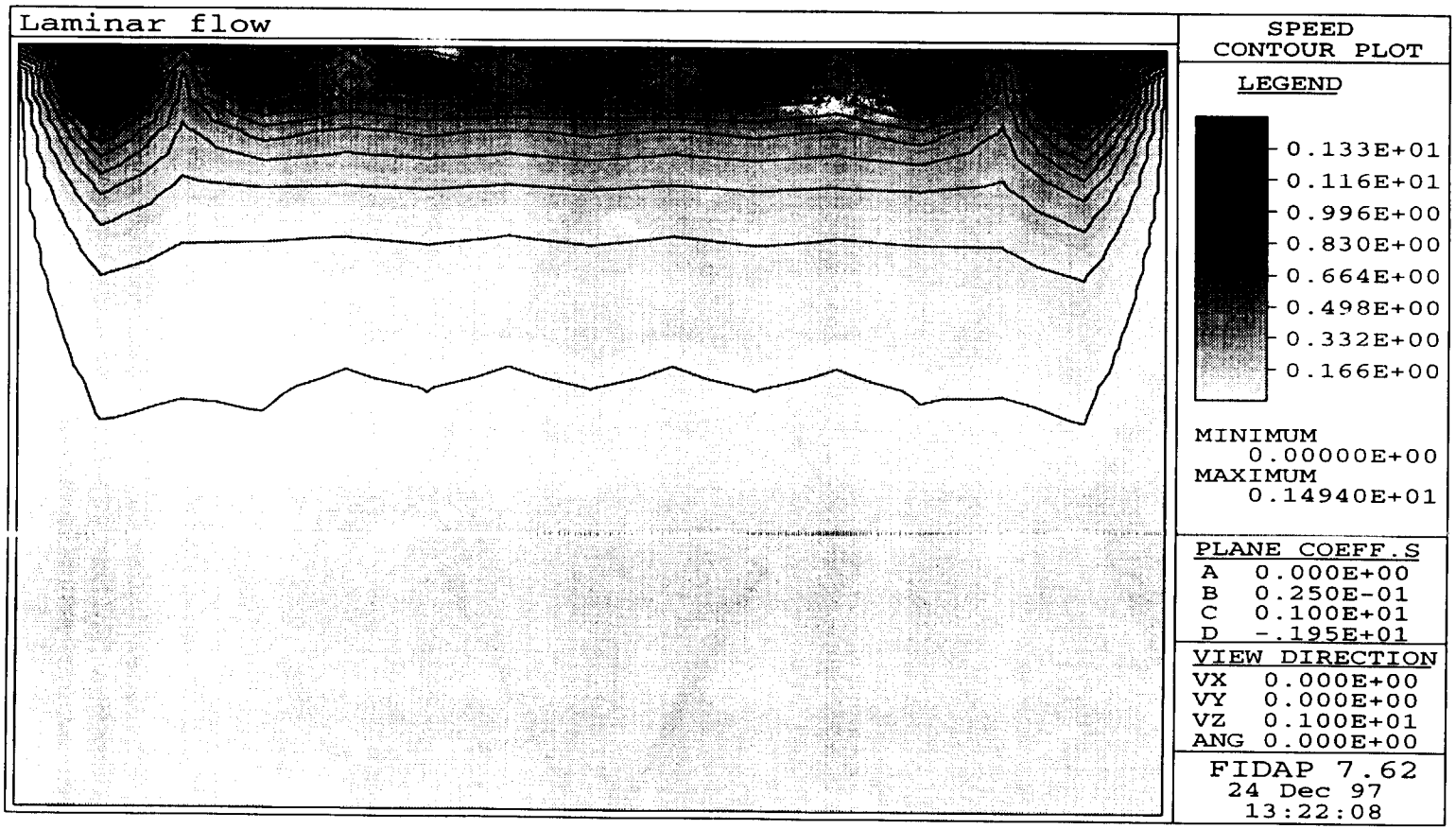

Figure 2.8: Speed Contours for the top slice (1.95m from the floor) under laminar flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.15ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

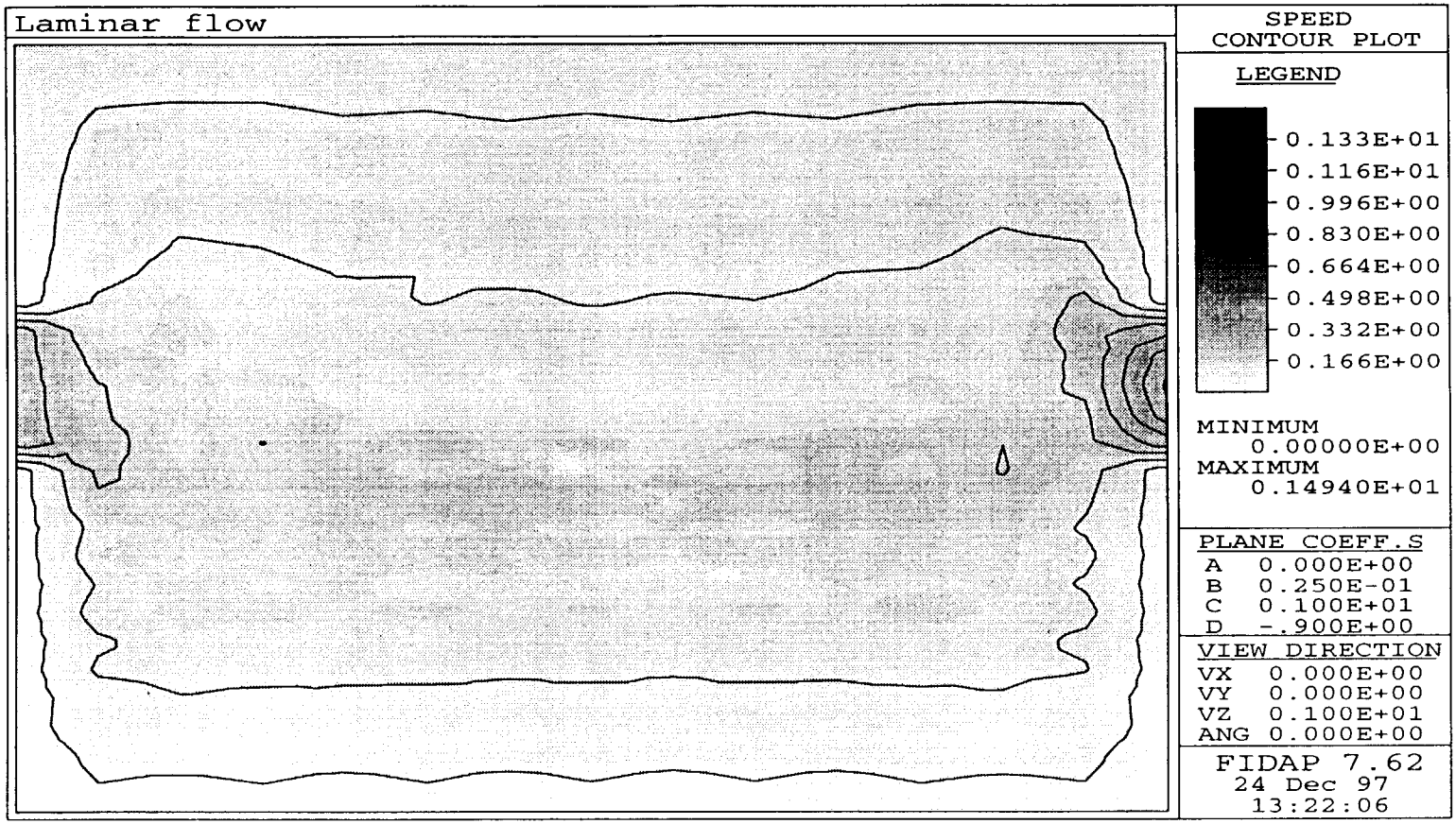

Figure 2.9: Speed Contours for the middle slice (0.9m from the floor) under laminar flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.15ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

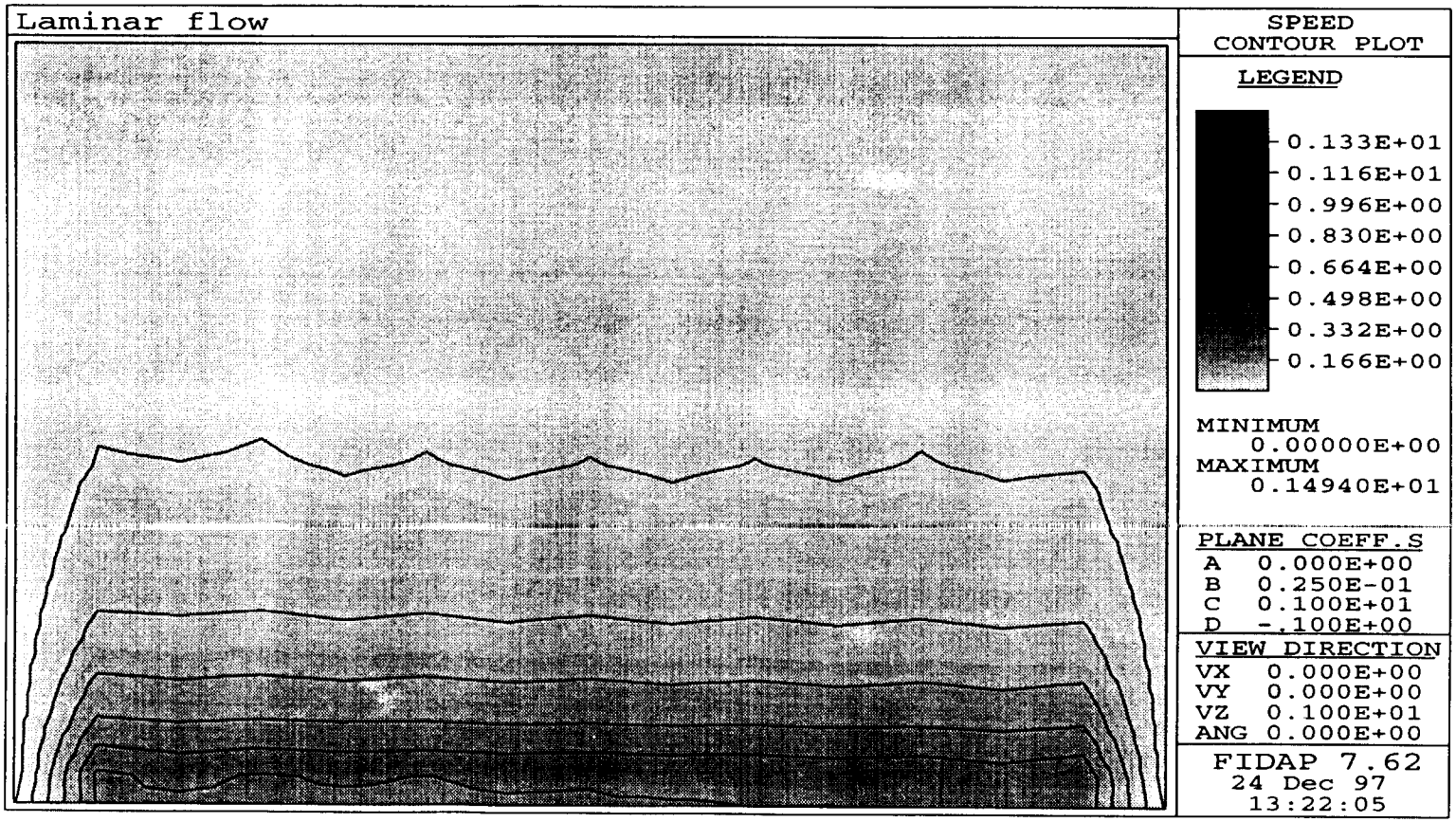

Figure 2.10: Speed Contours for the bottom slice (0.05m from the floor) under laminar flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.15ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

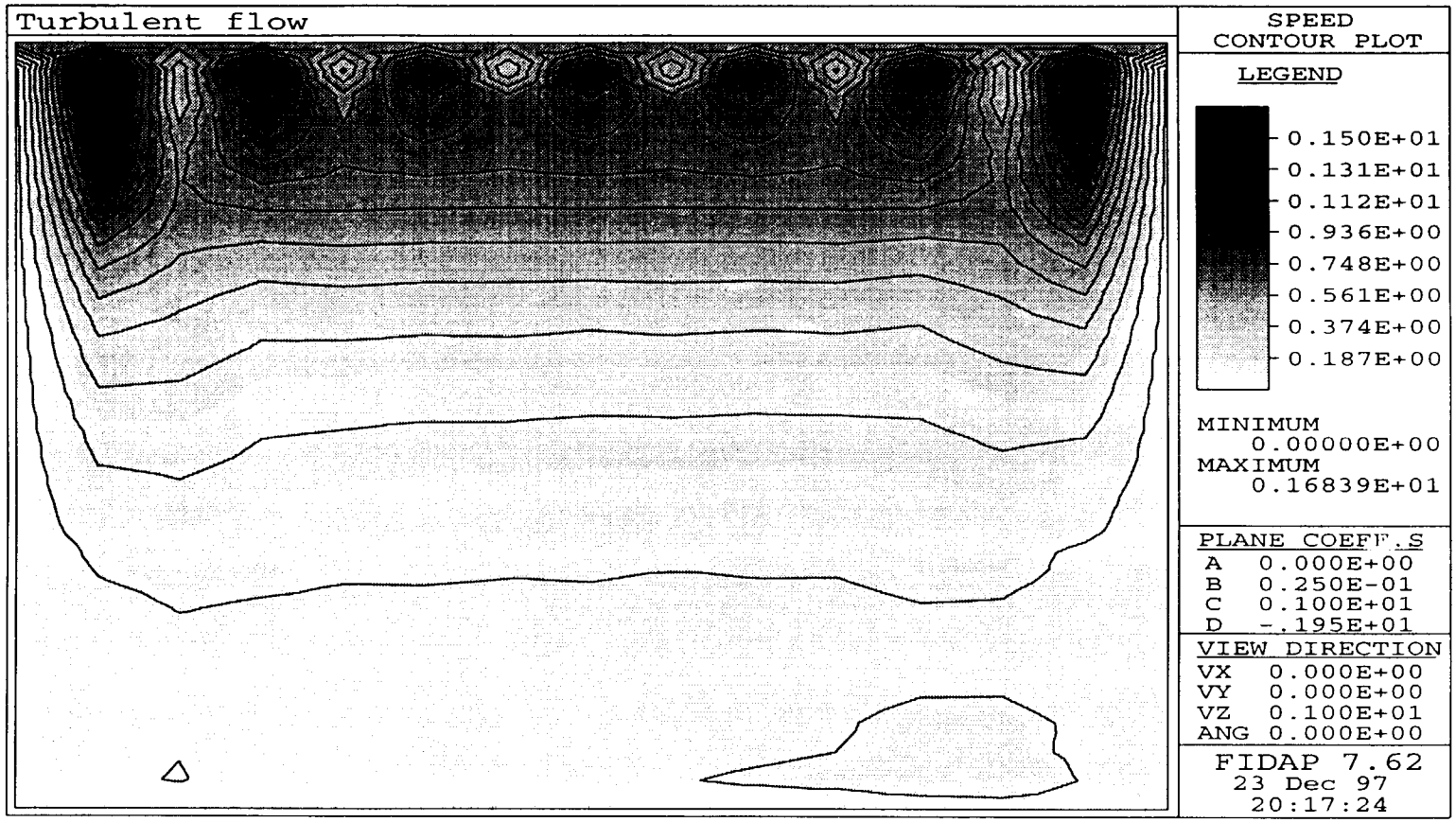

Figure 2.11: Speed Contours for the top slice - (1.95m from the floor) under turbulent flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.15ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

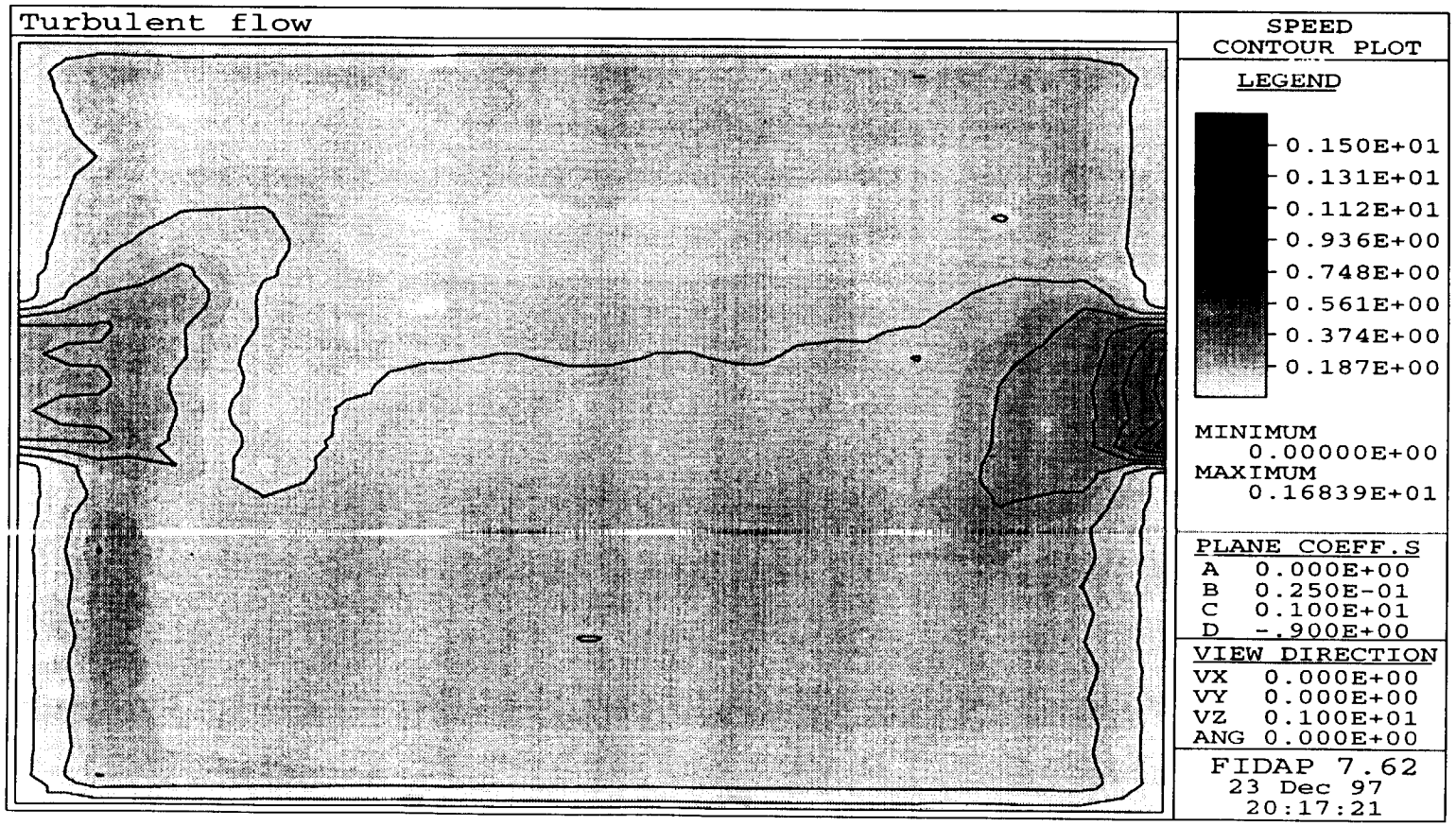

Figure 2.12: Speed Contours for the middle size (0.9m from the floor) turbulent flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.15ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

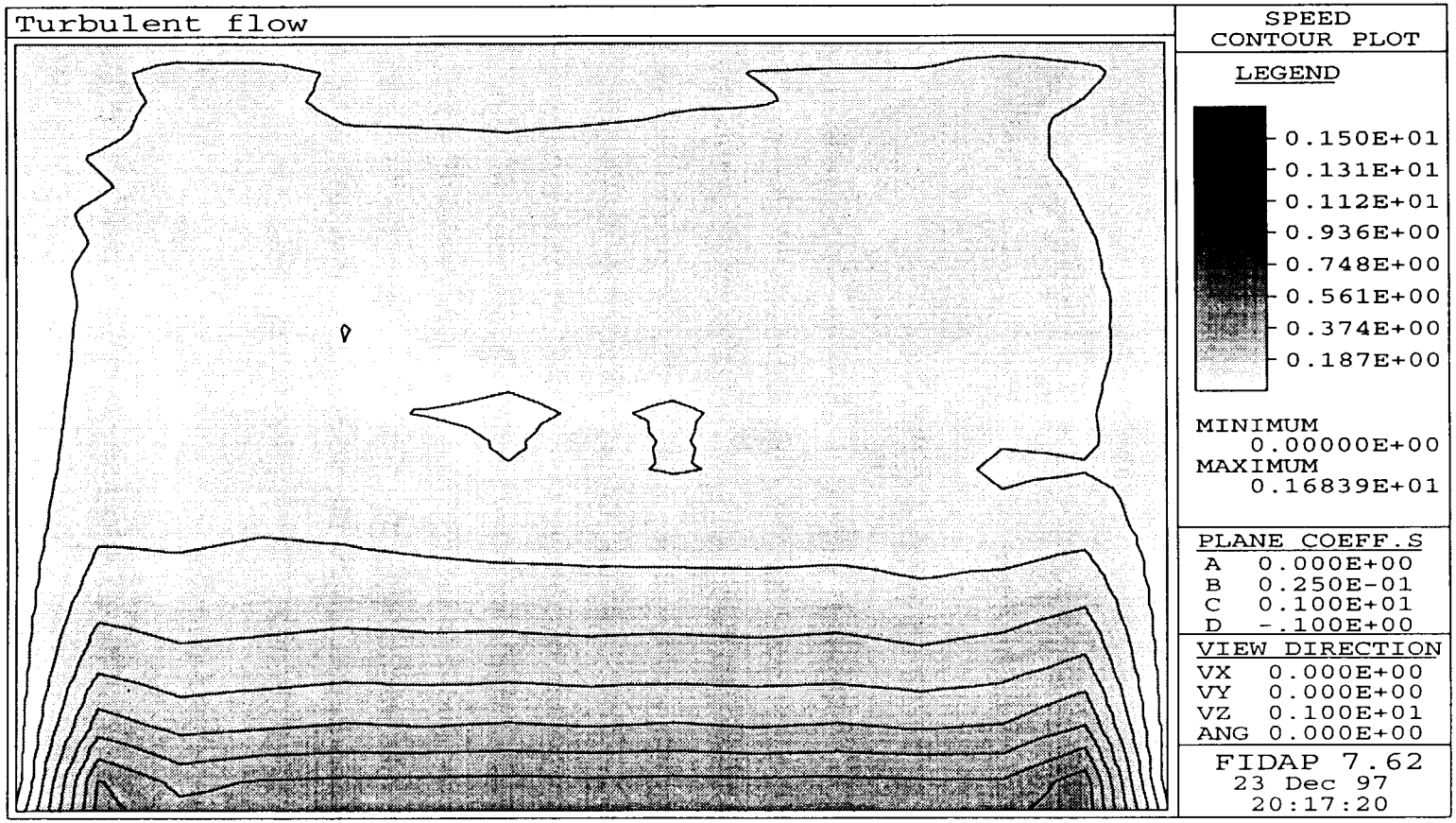

Figure 2.13: Speed Contours for the bottom slice (0.05m from the floor) under turbulent flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.15ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

Table 2.1: Normal mesh-errors for the velocity components and the pressure equations.

Degree of Freedom	RESIDUE
u	0.72617E-03
v	0.10399E-02
w	0.21082E-03
p	0.39848E-03

laminar and the turbulent profiles. We believe that the actual flow in the Space Station is mostly turbulent, and the the profiles we obtain are characteristic of low-velocity turbulence flows.

2.5 Mesh refinement studies

Numerical simulations are of course, subject to errors, and are closely related to the coarseness of the mesh used in the simulations. One common way of validating CFD simulation results is to refine the mesh being used and noting that there was no major change in the solution obtained for the same geometry, initial and boundary conditions. Tables 2.1 and 2.2 show the error residuals for each of the four equations, representing the velocities in the three co-ordinate directions and the pressure. Results are shown for both a normal mesh and for a refined mesh (double the number of mesh points). The results indicate that further refinement will not substantially change the overall flow profile. The flow profiles obtained in the previous sections suggest that a refinement of the mesh near the flow inlets would improve the accuracy of the solution near the inlets.

2.6 Complex flows

The flow in a cabin under operation is likely to be different from what we obtain by simulating an empty cabin, as we have done until now. Humans in the cabin and equipment are likely to cause more turbulence and recirculation cells. A

Table 2.2: Refined mesh-errors for the velocity components and the pressure equations.

Degree of Freedom	RESIDUE
u	0.12671E-02
v	0.14342E-02
w	0.10849E-02
p	0.71476E-03

detailed analysis of complex flows is out of scope of this work, but we attempt to obtain some preliminary results on what flow fields look like with objects in the cabin.

A set of simulations were carried out for the geometry shown in Fig. 2.14. Speed contours for three horizontal slices are shown in Figs. 2.15-2.17. The turbulence in the cabin was found to increase with the presence of an object, with more dead zones. The number of recirculation cells has increased. The velocities did not vary too much, due to the low velocity of the air flow.

2.7 Summary and Conclusions

Our objective in studying the air flow inside the cabin was to arrive at a basic understanding of cabin flows aboard the Space Station, and to obtain a few sample air flows that could be used as an input to our transport model, developed in the next Chapter. The flow profiles are very important to the transport model since most of the mass transfer in the cabin occurs as convective transport, and the accuracy of the flow field will therefore largely control the accuracy of the final transport model. Space Station flows are very poorly understood at present, and considerable further work is needed, both simulation of flows using CFD techniques, and experimental work which can validate those simulations.

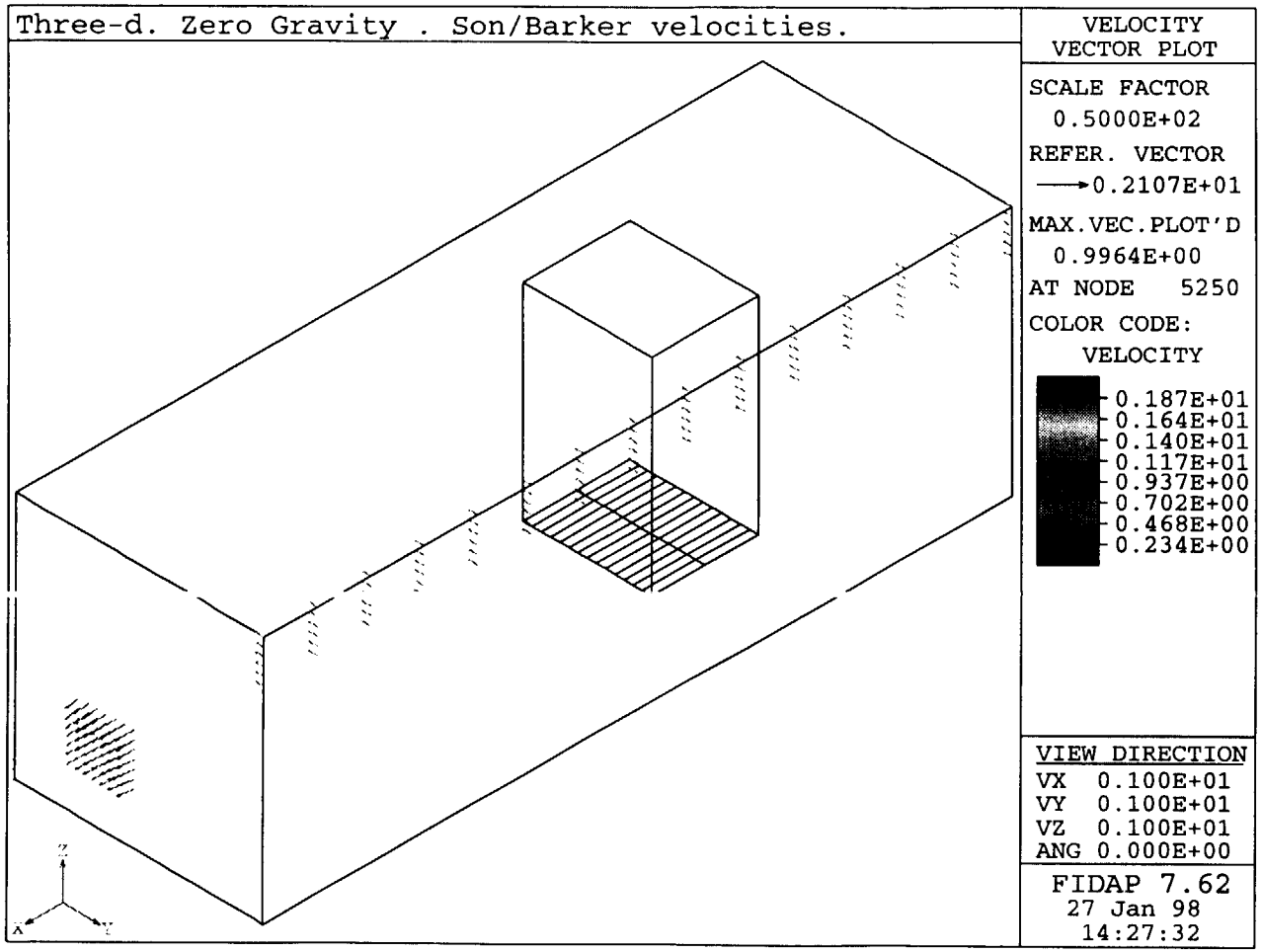

Figure 2.14: The cabin with an object inside, showing the THC inlet flow of 0.3 ms^{-1}, and the IMV outlet flow. The velocity is expressed in the units of ms^{-1}.

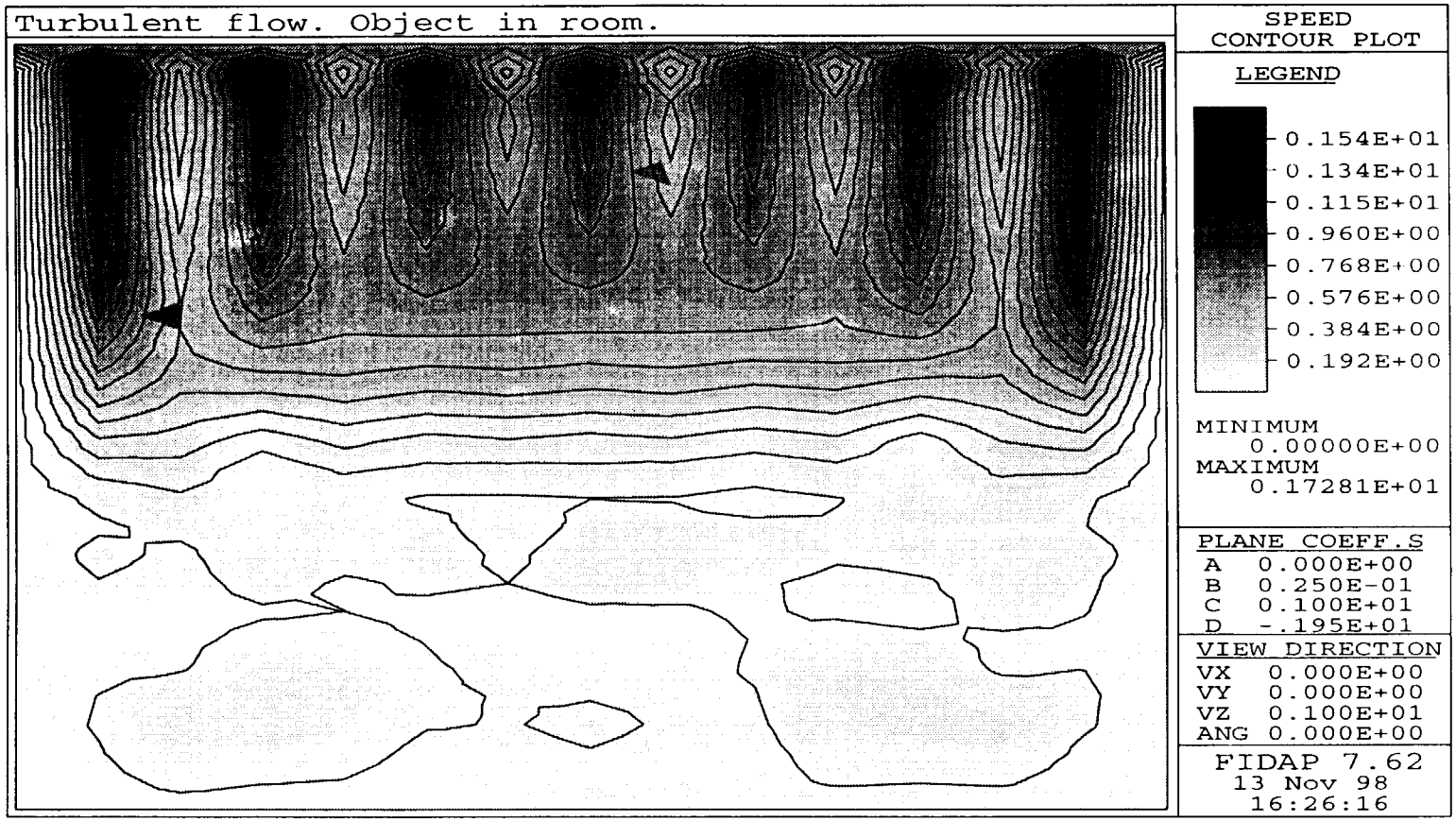

Figure 2.15: Speed Contours for the top slice - (1.95m from the floor) with an object (not shown) in the cabin under turbulent flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.5ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}.

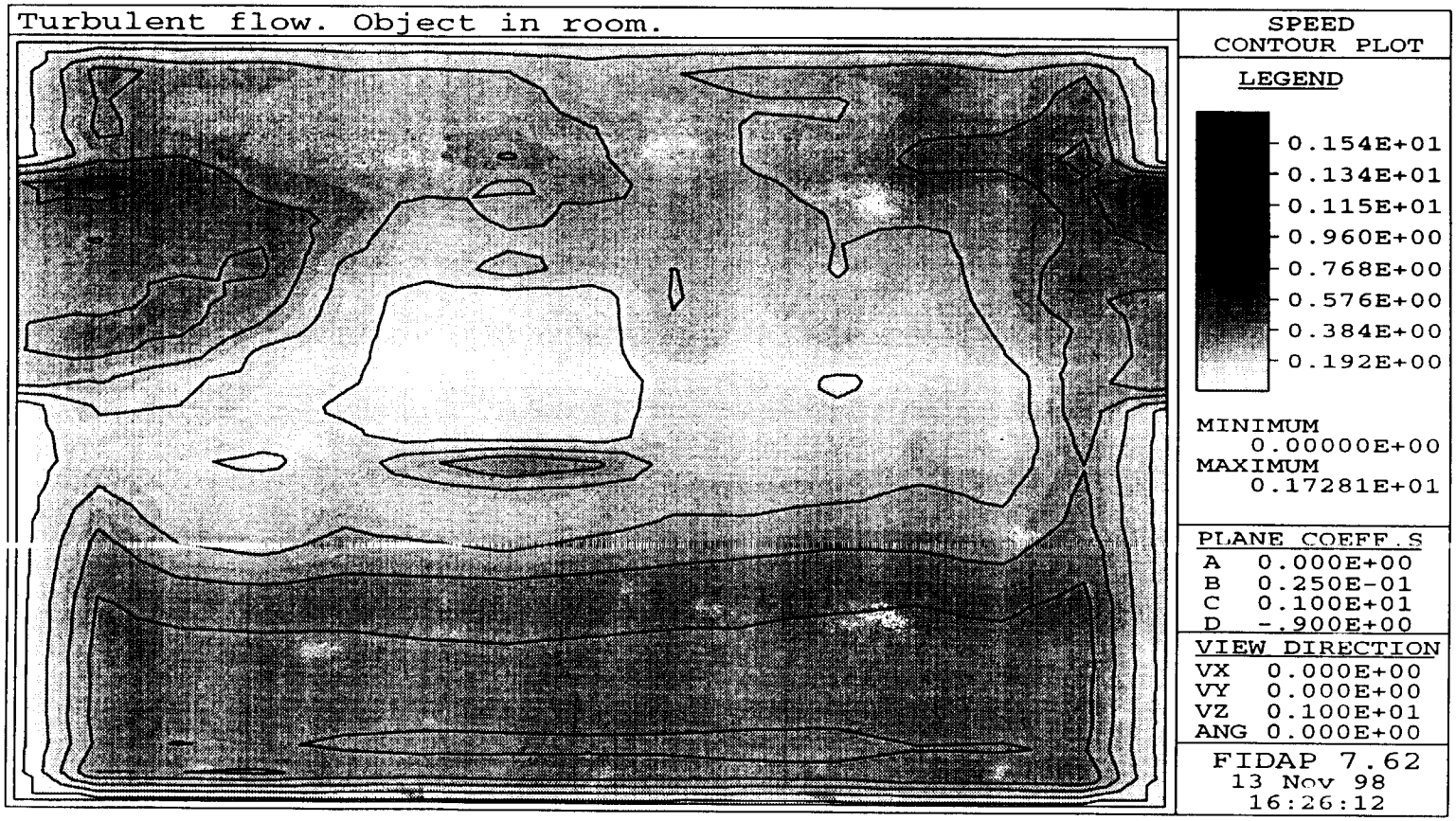

Figure 2.16: Speed Contours for the middle size (0.9m from the floor) with an object (not shown) in the cabin turbulent flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.5ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}. The zero velocity zone towards the center of the cabin marks the object.

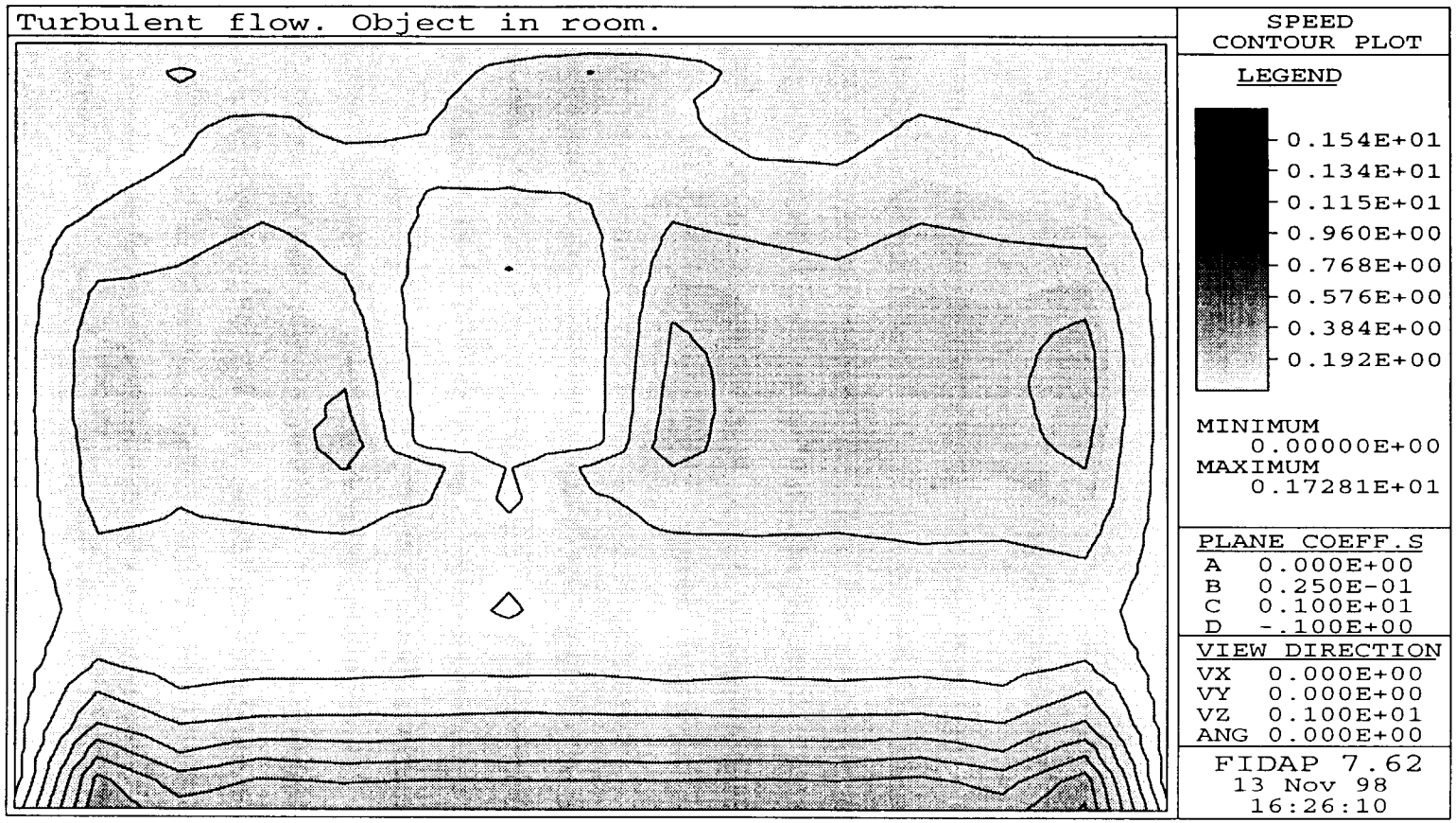

Figure 2.17: Speed Contours for the bottom slice (0.05m from the floor) with an object in the cabin under turbulent flow conditions for a THC flow of 0.3 ms^{-1} and an IMV flow of 0.5ms^{-1}. The speed and velocity are expressed in the units of ms^{-1}. The zero velocity zone towards the center of the cabin marks the object.

Chapter 3

Air contaminant dispersion modeling

3.1 Mathematical model

One of the goals of this work was the development of a complete three-dimensional transport model for the air-borne contaminants. The essential structure of the model is shown in Figure 3.1. Previous work used either the control-volume approach, in which the cabin was modeled as a collection of well-mixed reactors (National Academy of Sciences, 1981), or a space-averaged two-dimensional model (Skliar and Ramirez, 1997a).

Nazaroff reports a study of the effects of indoor air pollutants in which the indoor air was tagged with a non-reactive tracer and the decay of its concentration was monitored. They also extended this work to the study of flows between rooms using a combination of remote sensing and computed tomography techniques, which yields accurate results for the dispersion of air pollutants.

In this section, we present our model of the transport process, the discretization scheme to convert the partial differential equations to discrete representations, and discuss our solution scheme. The three-dimensional transport model developed here essentially extends the two dimensional model of Skliar (1996).

Assuming a constant density of spacecraft atmosphere and a constant molecular diffusivity, D_M, the differential mass balance of the air-borne contaminant (Bird et al., 1960) with concentration q results in the following fundamental three-

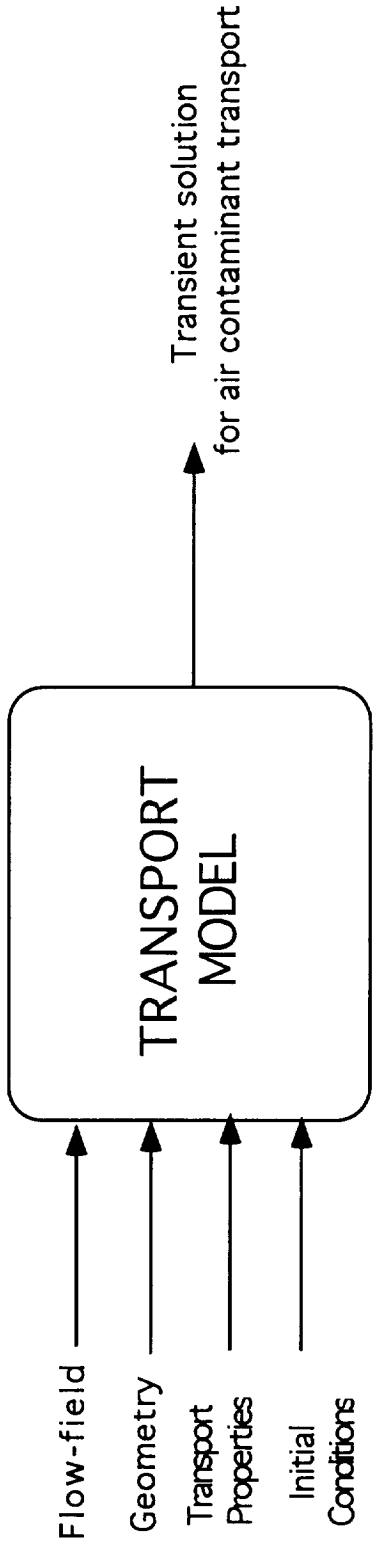

Figure 3.1: Inputs and output from the air contaminant transport model

dimensional convection-diffusion transport model

$$\frac{\partial q}{\partial t} + \mathbf{u} \cdot \nabla q = D_M \nabla^2 q + f \tag{3.1}$$

where $\mathbf{u}$ is the bulk air velocity, and f is the function that describes the capacity and location of contaminant sources and sinks. In this work, we are assuming that the contaminant is passive, i.e. it is transported with air at the same velocity in the field. In addition, we are assuming that the contaminant undergoes no chemical or physical transformations during its transport.

Equation 3.1 is applicable to both laminar and turbulent flow. However, in the case of turbulent flow, the velocity vector is extremely random, and so we resort to using the time averaged equations instead. The idea is to average the Fickian model over a time interval long enough for the integral of the instantaneous fluctuations to become zero. For the case of turbulent flow, therefore, we treat both the flow velocity and the concentration q as stochastic quantities. The transport equation, for the case of turbulent flow is written as

$$\frac{\partial \overline{q}}{\partial t} + \overline{\mathbf{u}} \cdot \nabla \overline{q} = \nabla^2 \overline{D_M q} + \overline{f} \tag{3.2}$$

where the overbar indicates that they are time-averaged quantities.

The eddy diffusivity, $\overline{D_M}$, which is the diffusivity under turbulent conditions, is a function of the flow field and is therefore not uniform throughout the geometry.

3.2 Computer Implementation of the three Dimensional Model

We solve the model Eqs. 3.1 and 3.2 above using a simple finite differencing scheme. In this section, we discuss how we discretize the equations, and then outline the solution technique used to solve the same equations. Since the flow field is expected to be turbulent in most cases, this derivation will proceed under that

assumption, and using the eddy diffusivity as the diffusivity coefficient. If laminar conditions are found to prevail, the molecular diffusivity should be used, which is independent of the flow field, and therefore is uniform throughout the cabin.

The diffusive terms are discretized using a second center difference scheme and appear as

$$\frac{\partial}{\partial x}\overline{D_M}\frac{\partial q}{\partial x} = d_{n+1,p,r}\frac{q_{n+1,p,r}-q_{n,p,r}}{\Delta x^2} - d_{n,p,r}\frac{q_{n,p,r}-q_{n-1,p,r}}{\Delta x^2} \tag{3.3}$$

$$\frac{\partial}{\partial y}\frac{\partial q}{\partial y} = d_{n,p+1,r}\frac{q_{n,p+1,r}-q_{n,p,r}}{\Delta y^2} - d_{n,p,r}\frac{q_{n,p,r}-q_{n,p-1,r}}{\Delta y^2} \tag{3.4}$$

$$\frac{\partial}{\partial z}d\frac{\partial q}{\partial z} = d_{n,p,r+1}\frac{q_{n,p,r+1}-q_{n,p,r}}{\Delta z^2} - d_{n,p,r}\frac{q_{n,p,r}-q_{n,p,r-1}}{\Delta z^2} \tag{3.5}$$

where Δx, Δy ,Δz are discretization steps along coordinates x,y, and z, d is the discrete analog of the diffusivity and the subscript is used to specify a point on the spatial mesh $\{(n,p,r) \mid n = \overline{1,N}, p = \overline{1,P}, r = \overline{1,R}\}$. N,P, and R determine the mesh size used in discretizing the cabin geometry.

The convective terms are discretized using the upwind differencing scheme in order to eliminate any possible oscillatory effects in the solution. The convective terms for the East-West, South-North, and Up-Down directions are as follows.

East-West

$$\frac{u\partial q}{\partial x} = u_{n,p,r}\frac{q^E_{n,p,r}-q^W_{n,p,r}}{\Delta x}, \tag{3.6}$$

where

$$q^E_{n,p,r} = \begin{cases} q_{n,p,r} \text{ if } u_{n,p,r} > 0, \\ q_{n-1,p,r} \text{ if } u_{n,p,r} < 0, \end{cases}$$

$$q^W_{n,p,r} = \begin{cases} q_{n-1,p,r} \text{ if } u_{n,p,r} > 0, \\ q_{n,p,r} \text{ if } u_{n,p,r} < 0. \end{cases}$$

South-North

$$\frac{v\partial q}{\partial y} = v_{n,p,r}\frac{q^S_{n,p,r}-q^N_{n,p,r}}{\Delta y}, \tag{3.7}$$

where

$$q^S_{n,p,r} = \begin{cases} q_{n,p-1,r} \text{ if } v_{n,p,r} > 0, \\ q_{n,p,r} \text{ if } v_{n,p,r} < 0, \end{cases}$$

$$q^N_{n,p,r} = \begin{cases} q_{n,p,r} \text{ if } v_{n,p,r} > 0, \\ q_{n,p-1,r} \text{ if } v_{n,p,r} < 0. \end{cases}$$

Up-Down

$$\frac{w\partial q}{\partial z} = w_{n,p,r}\frac{q^U_{n,p,r} - q^D_{n,p,r}}{\Delta z}, \tag{3.8}$$

where

$$q^U_{n,p,r} = \begin{cases} q_{n,p,r} \text{ if } w_{n,p,r} > 0, \\ q_{n,p,r-1} \text{ if } w_{n,p,r} < 0, \end{cases}$$

$$q^D_{n,p,r} = \begin{cases} q_{n,p,r-1} \text{ if } w_{n,p,r} > 0, \\ q_{n,p,r} \text{ if } w_{n,p,r} < 0. \end{cases}$$

The application of the center difference approximation of the time derivative yields the following discrete analog of the three dimensional transport model.

$$\begin{aligned}
\frac{q^{m+1}_{n,p,r} - q^m_{n,p,r}}{\Delta t} &= -\frac{1}{2}\left(u_{n,p,r}\frac{\left[q^E_{n,p,r}\right]^{m+1} - \left[q^W_{n,p,r}\right]^{m+1}}{\Delta x} + u_{n,p,r}\frac{\left[q^E_{n,p,r}\right]^{m} - \left[q^W_{n,p,r}\right]^{m}}{\Delta x} \right. \\
&+ v_{n,p,r}\frac{\left[q^S_{n,p,r}\right]^{m+1} - \left[q^N_{n,p,r}\right]^{m+1}}{\Delta y} + v_{n,p,r}\frac{\left[q^S_{n,p,r}\right]^{m} - \left[q^N_{n,p,r}\right]^{m}}{\Delta y} \\
&+ \left. w_{n,p,r}\frac{\left[q^D_{n,p,r}\right]^{m+1} - \left[q^U_{n,p,r}\right]^{m+1}}{\Delta z} + w_{n,p,r}\frac{\left[q^D_{n,p,r}\right]^{m} - \left[q^U_{n,p,r}\right]^{m}}{\Delta z} \right) \\
&- \frac{1}{2}\left[d_{n+1,p,r}\frac{q^{m+1}_{n+1,p,r} - q^{m+1}_{n,p,r}}{\Delta x^2} - d_{n,p,r}\frac{q^{m+1}_{n,p,r} - q^{m+1}_{n-1,p,r}}{\Delta x^2} \right. \\
&+ \left. d_{n+1,p,r}\frac{q^{m}_{n+1,p,r} - q^{m}_{n,p,r}}{\Delta x^2} - d_{n\,p,r}\frac{q^{m}_{n,p,r} - q^{m}_{n-1,p,r}}{\Delta x^2} \right] \\
&- \frac{1}{2}\left[d_{n,p+1,r}\frac{q^{m+1}_{n,p+1,r} - q^{m+1}_{n,p,r}}{\Delta y^2} - d_{n,p,r}\frac{q^{m+1}_{n,p,r} - q^{m+1}_{n,p-1,r}}{\Delta y^2} \right. \\
&+ \left. d_{n,p+1,r}\frac{q^{m}_{n,p+1,r} - q^{m}_{n,p,r}}{\Delta y^2} - d_{n\,p,r}\frac{q^{m}_{n,p,r} - q^{m}_{n,p-1,r}}{\Delta y^2} \right] \\
&- \frac{1}{2}\left[d_{n,p,r+1}\frac{q^{m+1}_{n,p,r+1} - q^{m+1}_{n,p,r}}{\Delta z^2} - d_{n,p,r}\frac{q^{m+1}_{n,p,r} - q^{m+1}_{n,p,r-1}}{\Delta z^2} \right.
\end{aligned}$$

$$+ \quad d_{n,p,r+1}\frac{q^m_{n,p,r+1} - q^m_{n,p,r}}{\Delta z^2} - d_{n,p,r}\frac{q^m_{n,p,r} - q^m_{n,p,r-1}}{\Delta z^2}\Bigg] + f^{m+1}_{n,p,r} \qquad (3.9)$$

where $f^{m+1}_{n,p,r}$ represents the value of the time dependent source function f evaluated at the current time step, Δt is the time discretization step, and the superscript $m = 0, 1, 2, \ldots$ is used to identify an instance $t = (m+1)\Delta t$ for which the solution of the equation is sought.

3.2.1 Numerical solution of the transport model equations

Because of their poor stability properties, explicit difference methods are rarely used to solve initial and boundary-value problems in two or more space dimensions. The solution scheme used here is the classic Alternating Direction Implicit scheme (ADI) (Douglas and Rachford, 1956; Douglas, 1962), which invokes the property of operator splitting and converts the problem into a system of three tridiagonal matrix equations, along lines parallel to the x, y and z co-ordinate directions, which can be solved using the Thomas Algorithm (Godunov, 1959). Solving the three tridiagonal equations yields a solution for the concentration at the next time step, $\mathbf{q_{m+1}}$ via the intermediate concentrations (both dummy variables), $\mathbf{q}^*$ and $\mathbf{q}^{**}$. The convective operators are discretized using an upwind first order scheme, while the diffusive terms are discretized using a second order center difference scheme. The time operator is a simple forward difference term. The error is of $O(\Delta x, \Delta y, \Delta z, \Delta t)$.

The like terms in (3.9) are collected to obtain the following equations for a single spatial mesh point (n, p, r):

$$\begin{aligned}
(-\frac{\mathbf{A}_x}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}^* &= (\frac{\mathbf{A}_x}{2} + \mathbf{A}_y + \mathbf{A}_z - \frac{\mathbf{r}}{\Delta t})\mathbf{q}_m + \mathbf{f}_m \\
(-\frac{\mathbf{A}_y}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}^{**} &= -\frac{\mathbf{A}_y}{2}\mathbf{q}_m - \frac{\mathbf{r}}{\Delta t}\mathbf{q}^* \\
(-\frac{\mathbf{A}_z}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}_{m+1} &= -\frac{\mathbf{A}_z}{2}\mathbf{q}_m - \frac{\mathbf{r}}{\Delta t}\mathbf{q}^{**}
\end{aligned} \qquad (3.10)$$

(3.11)

This system of equations is then represented as a single matrix equation in terms of the State transition matrices, $\mathbf{A}_1$ and $\mathbf{A}_2$.

$$\mathbf{A}_1\mathbf{Q}_{m+1} = \mathbf{A}_2\mathbf{Q}_m + \begin{bmatrix} \mathbf{f}_m \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{3.12}$$

where

$$\mathbf{Q}_m = \begin{bmatrix} \mathbf{q}_m^* \\ \mathbf{q}_m^{**} \\ \mathbf{q}_m \end{bmatrix} \tag{3.13}$$

$$\mathbf{A}_1 = \{\mathbf{A}_1^{ij}\} = \begin{bmatrix} (\frac{-\mathbf{A}_x}{2} - \frac{\mathbf{r}}{\Delta t}) & \mathbf{0} & \mathbf{0} \\ \frac{\mathbf{r}}{\Delta t} & (\frac{-\mathbf{A}_y}{2} - \frac{\mathbf{r}}{\Delta t}) & \mathbf{0} \\ \mathbf{0} & \frac{\mathbf{r}}{\Delta t} & (\frac{-\mathbf{A}_z}{2} - \frac{\mathbf{r}}{\Delta t}) \end{bmatrix} \tag{3.14}$$

and

$$\mathbf{A}_2 = \mathbf{A}_2^{ij} = \begin{bmatrix} 0 & 0 & (\mathbf{A}_x/2 + \mathbf{A}_y + \mathbf{A}_z - \mathbf{r}/\Delta t) \\ 0 & 0 & -\mathbf{A}_y/2 \\ 0 & 0 & -\mathbf{A}_z/2 \end{bmatrix} \tag{3.15}$$

$\mathbf{A}_x$, $\mathbf{A}_y$, and $\mathbf{A}_z$ are finite difference approximations used in the State Transition matrices representing each of the spatial directions, where, for instance, $\mathbf{A}_x$ is the approximation of

$$\frac{\partial}{\partial x}\overline{D_M}\frac{\partial}{\partial x} - \frac{\partial u}{\partial x} \tag{3.16}$$

The solution to this set of equations for the appropriate initial and boundary conditions discussed in the next section yields the concentration profiles for the cabins.

3.2.2 Boundary conditions

Boundary conditions may be of the following kinds:

Boundary Condition of the first kind:

This is also known as a *Dirichlet condition*, and specifies a given concentration along a boundary. This occurs commonly at inlet ducts where the flow is coming in at a certain concentration of contaminant.

Boundary condition of the second kind:

This is also known as a *Neumann condition*, and specifies a concentration derivative normal to the surface of the boundary, A wall, for example is represented as a Neumann condition with the derivative of the concentration being set to zero.

Boundary condition of the third kind:

Also known as a *Robbins condition*, this specifies a combination of a concentration and a flux at the boundary, and does not usually occur in cases of contaminant dispersion, although it occurs commonly in convection diffusion when applied to heat transfer problems.

Continuity boundary condition:

This is prescribed typically along interfaces, open boundaries, and at ducts linking cabins, and for no barriers to mass-transfer, specifies that the flux must be constant across a boundary. If a barrier exists, say, a membrane across which the cabin air diffuses, a resistance to mass transfer may be used to specify the boundary condition.

The different boundary conditions are used in this work in the following cases: (Skliar, 1996; Roache, 1972)

The nature of the boundary must be described mathematically to completely specify the problem. This involves describing the volume of interest, as a single volume or as a set of volumes glued to one another, and the appropriate boundary conditions. The boundary condition can be of three main types, the Dirichlet

boundary condition, where the concentration at a given boundary is specified, the Neumann boundary condition, where the mass flux at the boundary is fixed, and the mixed or Robbins condition, where a function of the flux (the derivative) and the concentration is specified. In physical terms, the boundaries are encountered in the following cases.

The wall: This is a no flux Neumann boundary condition. This also is the default boundary condition for the model.

Duct In/Out: This would happen in the case of a window of some kind, with significant convective flow in/out. The velocity at that mesh point is then specified by the free stream velocity, and the diffusion is just as before. These are inflow and outflow boundaries and the concentration is ether allowed to float (for outflow ducts) or specified as a Dirichlet condition (for inflow ducts), as far the the mesh is concerned. The value of the float is determined by the free-stream conditions outside the volume and will have to be specified from time to time.

Open Hatch: This is similar to the duct, except that flows are ignored.

Membrane wall: This boundary is treated like an interior point, with the velocity set to the flow velocity determined by the membrane, and the diffusivity is changed to the value dictated by the membrane.

Removal device wall: This is a special case, and allows for a boundary to exist within the volume, with a diffusivity different from the rest of the volume, or can act as a sink, which then would be treated a source term with a negative capacity.

The solution to any partial differential equation (PDE) can depend on the boundary conditions and the initial conditions applied to the PDE. It is therefore, not surprising that the specification of the computational boundary condition, besides affecting numerical stability, affects the accuracy of the PDE solution in a significant manner. The intermediate values, q^* and q^{**} are not necessarily approximations to the value at the end of the iteration. As a result, particularly for high

order accuracy methods, the boundary conditions for the intermediate sweeps must be obtained in terms of the boundary values at $t = nk$ and $t = (n+1)k$. For boundary conditions independent of time (eg. Dirichlet conditions), the conditions are straightforward(LeVeque, 1985)

$$q^* = q^{**} = g^{m+1} \tag{3.17}$$

where g^{m+1} is the specified concentration at the boundary. For Neumann conditions the relations are:

$$q^* = \frac{\beta}{\gamma} g^m + (1 - \gamma A_y^2)(1 - \gamma A_z^2)(\frac{\gamma}{\beta} g^{m+1} - \frac{\beta}{\gamma} g^m) \tag{3.18}$$

$$q^{**} = \frac{\beta}{\gamma} g^m + (1 - \gamma A_z^2)(\frac{\gamma}{\beta} g^{m+1} - \frac{\beta}{\gamma} g^m) \tag{3.19}$$

where g^{m+1} is the specified flux boundary condition, and $\gamma = \beta = \Delta x/2$.

The Neumann boundary also requires the calculation of the flux at the boundary. At the boundary, however, a center difference cannot be used to compute the first derivative since there are no neighbor mesh points in one of the directions (say, to the left of a point on the x boundary). This problem is circumvented by using a reflective boundary condition (where the domain is extended left of the actual boundary). One-sided differences could also be used. In this work, we use the reflective boundary condition to model the boundaries.

3.3 Model Testing

We first ran simulations of the contaminant dispersal using cases where we expected a known pattern of dispersal to see if the model was performing satisfactorily. The model was therefore tested for stagnant cabins, and for cabins with air flow, with two kinds of contaminant sources, puffs, which are instantaneous releases of contaminant, and continuous streams of contaminant flow.

Case 1: Stagnant Room – Source in the center.

In the first test case, the velocity vector was assumed to be identically zero at all the node points. This is a useful test case, since we know *a priori* what the model should predict: a gradual spread of the contaminant outwards, while preserving the symmetry of the distribution. A point source was introduced at the geometric center of the cabin, and the contaminant concentrations were observed for subsequent time steps. A 31 x 31 x 31 mesh was used to discretize a model cabin of size 7 x 7 x 7m, and Figure 3.2 shows the contours at grid point slices 1, 10, 15, and 29 (in the z-direction) for time Steps 3,7,29, and 49. Note that the solution is symmetric in all directions.

Case 2: Room with air flow – Source in the center.

Having tested a rather rudimentary test case, we now considered a case with a fictitious wind field. A wind field with $u = 0.5m/s, v = 0, w = 0$ was used at all the node points. This wind field again has the advantage that we can expect a certain pattern in the solution. Again, we use a 31 x 31 x 31 mesh, with two ducts on either side. Figure 3.3 shows the contours obtained for different time steps. Transport in the x-direction is mainly via convection, and diffusion in the y and z directions is only molecular. One can observe that the convective transport is much faster than the molecular transport. Consequently, we may conclude that the accuracy of the flow field will largely control the accuracy of the transport model.

3.4 Contamination scenarios

We now proceed to test the working of our model by simulating some test cases. In this section, we use the wind field that we obtained in the sample cases (Narayan and Ramirez, 1998b) to observe how contaminant dispersal occurs. We consider two separate cases, both of which represent actual Space Station contaminant scenarios.

Case 1: Steady-State Contamination

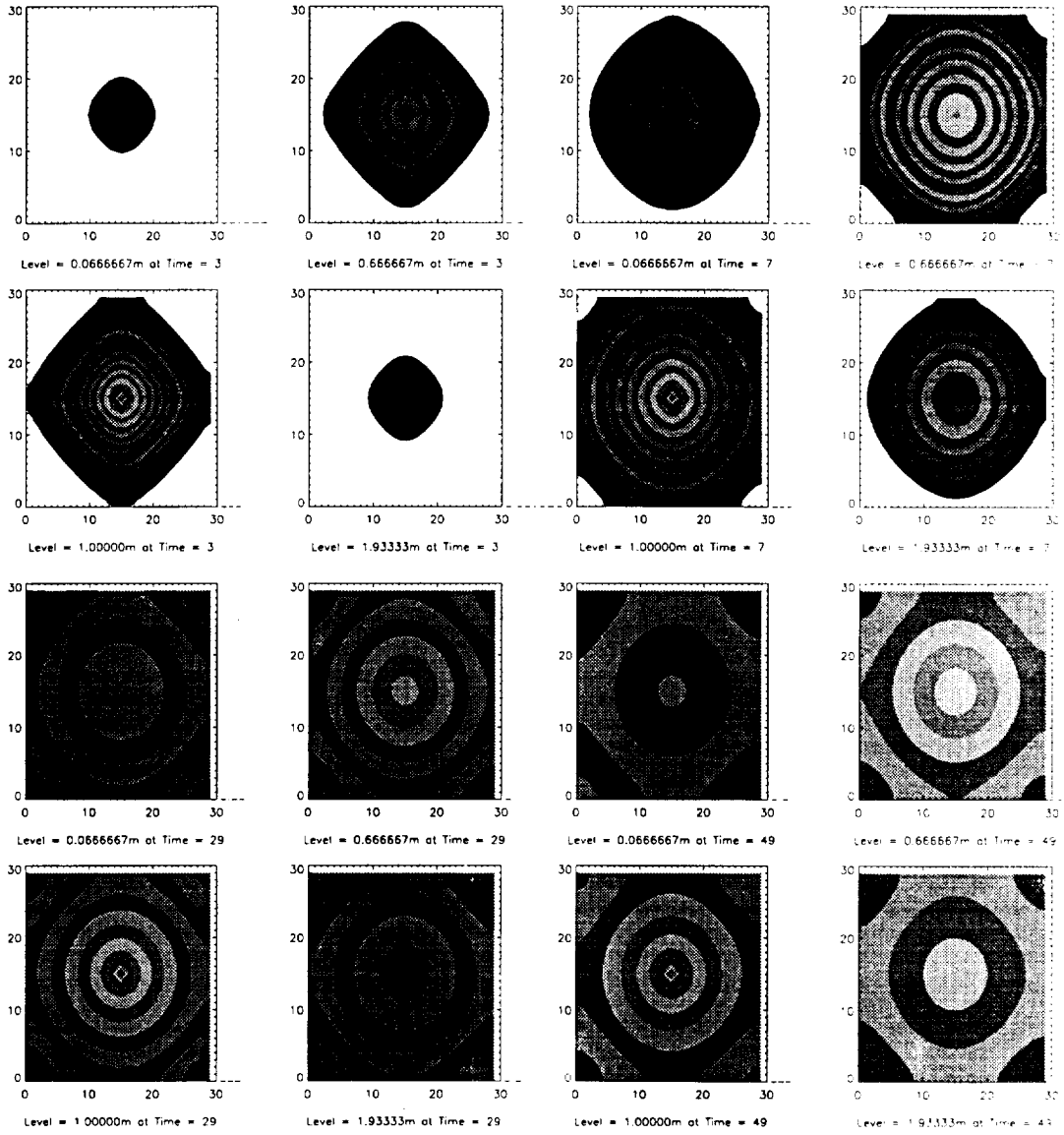

Figure 3.2: Contaminant concentration contours for a stagnant room with a continuous release of contaminant in the center of the cabin. The contours are shown for 4 horizontal slices per time step (left-right, top-down), at levels 0.07,0.7,1, and 1.9m from the floor of the cabin at time-steps 3,7,19 and 49.

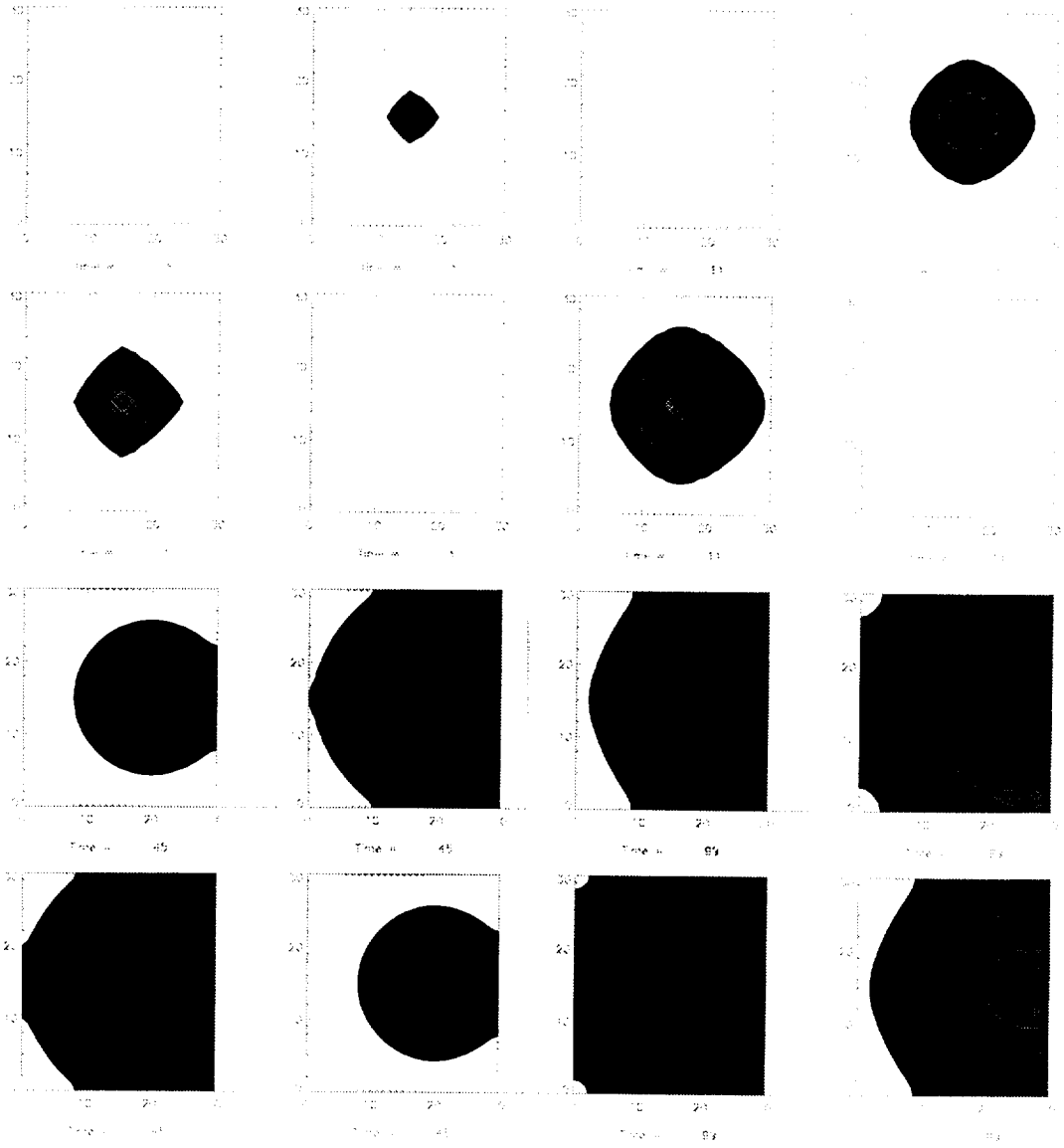

Figure 3.3: Contaminant concentration Contours for a room with flow, with a uniform velocity of 0.5 ms^{-1} shown at levels 0.07,0.7,1, and 1.9m from the floor of the cabin at time-steps 3,11,45 and 99.

This case attempts to mimic the routine operation of the Space Station module. For this specific case, we are assuming a release of 20 mg of carbon dioxide over the first two time steps(of 4.5s each). In addition, we have a steady input of carbon dioxide in the inlet streams. The cabin has a residual carbon dioxide concentration of 0.3 volume % for its initial condition. The THC air comes at a concentration of 0.5 volume % and the IMV with a concentration of 0.71 volume %. This would be realistic since the THC is treated , and we could assume that one of the other modules has a great deal of astronaut activity, and thus the high carbon dioxide level. The simulation was done for about 300 time steps (about 1350 s). By this time, we observe a steady-state concentration distribution. Figure 3.4 shows the contamination levels at four different slices in the cabin, at levels 0.067m, 0.67m, 1m, and 1.87m from the floor of the cabin after about 1300 s. The origin, (0,0,0) refers to the point in the cabin at the left bottom corner in Fig. 2.1. The surface plots of the slices closer to the middle of the room show similar profiles, and the exits and the consequent drop in concentration levels of the contaminant are clearly visible. This is due to the strong blast (relative to the rest of the cabin) of wind removing the contaminant through convective transport.

Case 2: Sudden Release of Carbon dioxide

Carbon dioxide may be used to extinguish a fire. A large release of carbon dioxide then would occur over a small time frame, and we wish to monitor how the contaminant levels gradually decrease. Figures 3.5 and 3.6 show the concentration levels 90 and 1500 s after the release. The surface plots and contours are shown at planar slices 0.067m, 0.67m, 1m, and 1.87m from the floor of the cabin. During these 1500 s, more than 70% of the released carbon dioxide has been flushed out from the room. Figure 3.5 shows the profiles 90 s after the release, which happened near the bottom left corner of the cabin (0.27m,0.14m,0.17m). In this figure, substantial amounts of the carbon dioxide are still present near the location of the occurrence, although the levels drop off near the outlets. In Fig. 3.6, which shows the profiles

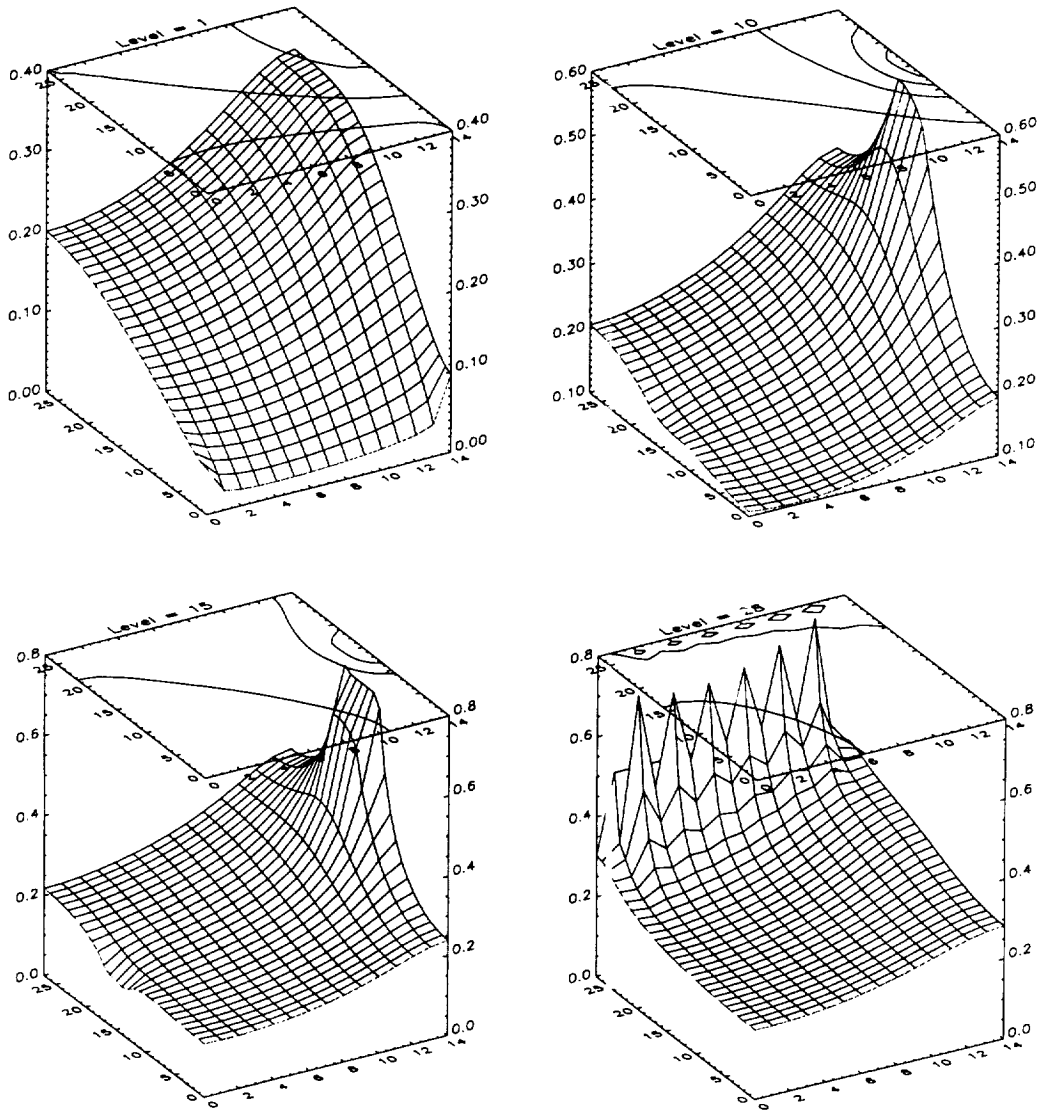

Figure 3.4: Steady State Contamination CO_2 profiles for the cabin with an initial carbon dioxide concentration of 0.3 volume %, an IMV inlet flow at 0.71 volume %, and a THC concentration of 0.5 volume %. The concentration profiles are shown at four levels, 0.067m (level = 1), 0.67m (level = 10), 1m (level = 15), and 1.87m (level = 28) above the floor of the cabin.

1500 s after the release, we observe that the carbon dioxide is almost well-mixed with only a slight bulge at the location of the release in the top slice. This accumulation of carbon dioxide near the ceiling is a consequence of the low flow region at the top, and because of the stagnation zone that exists near the ceiling. By this time, it is the flow profile which largely decides the contamination concentration profile, which is indicative of the fact that convective transport is the dominant mass transfer mechanism. For a sample average flow velocity of 1.5 m/s, the mass transfer Peclet number is 2 x 10^5, which implies that convective transport dominates molecular diffusion.

Note here that there is a significant variation in the concentrations across the room, which might mean that lumped models of the cabin would be grossly inaccurate. Also, there are regions of accumulation in the room. This could mean that SMACs could be locally violated, even though the concentration averaged over the entire cabin may be below the SMAC limit. The flow field is an important parameter in the way contaminants spread through the cabin and needs to be closely monitored.

3.5 Modeling of cleanup process

Crucial to the Advanced Environmental Monitoring process is the modeling and monitoring of the contaminant removal processes. Removal processes range from HEPA filters, molecular sieves, and adsorption packs to the use of the Sabatier and the Bosch processes for the regeneration of carbon dioxide.

The approach used in this work is that all removal devices can be modeled in two distinct ways; either as devices that remove a certain percentage of the contaminant and leave the rest in the medium, or as devices that remove almost all of the contaminant leaving only a small residual concentration behind. The air that passes through the removal device is then assumed to return to the main stream and re-enter the module. An additional parameter is the amount of contaminant that

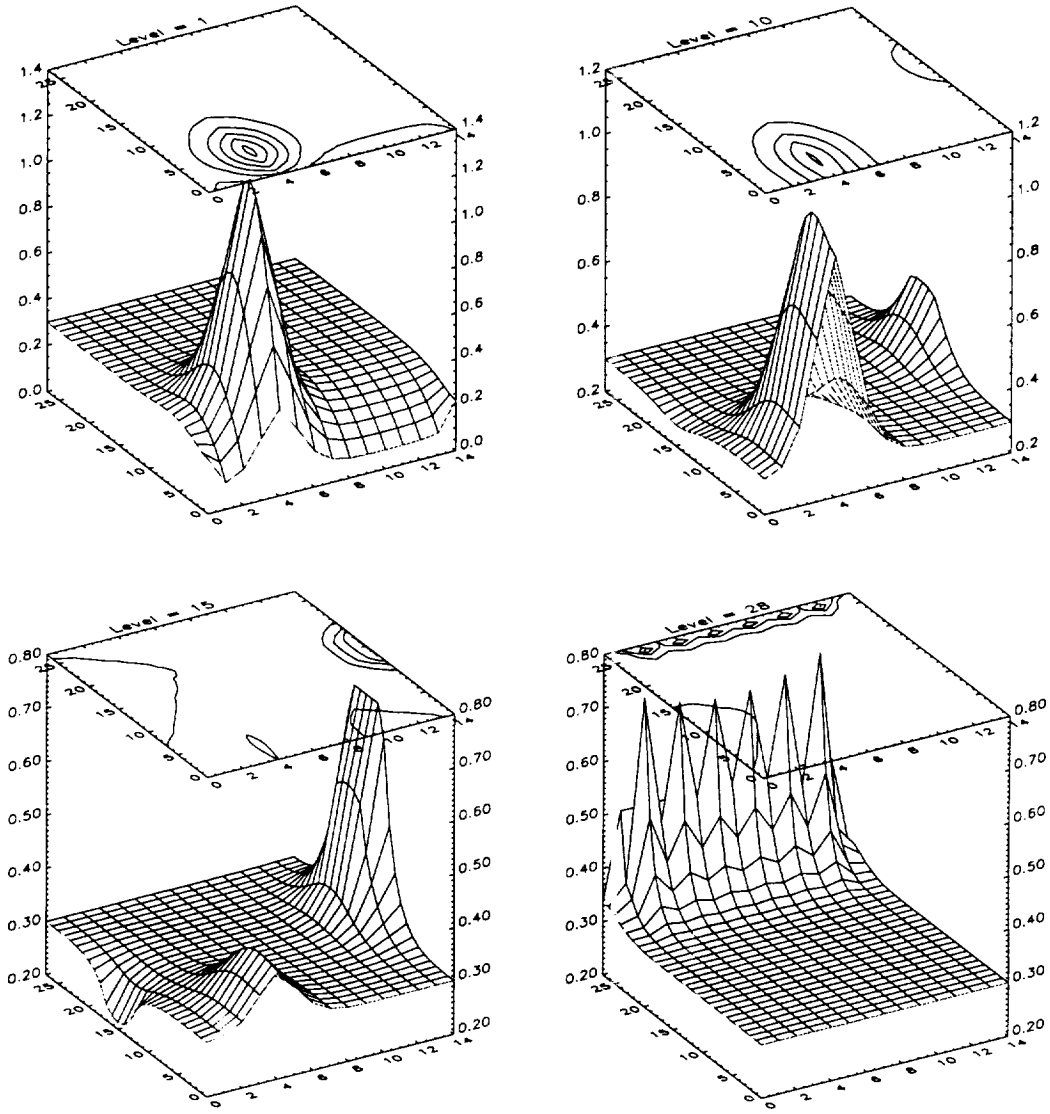

Figure 3.5: CO_2 profiles at four different elevations, about 90 seconds after a sudden release of carbon dioxide at grid location (4,5,8) with an initial carbon dioxide concentration of 0.3 volume %, an IMV inlet flow at 0.71 volume %, and a THC concentration of 0.5 volume %. The concentration profiles are shown at four levels, 0.067m (level = 1), 0.67m (level = 10), 1m (level = 15), and 1.87m (level = 28) above the floor of the cabin.

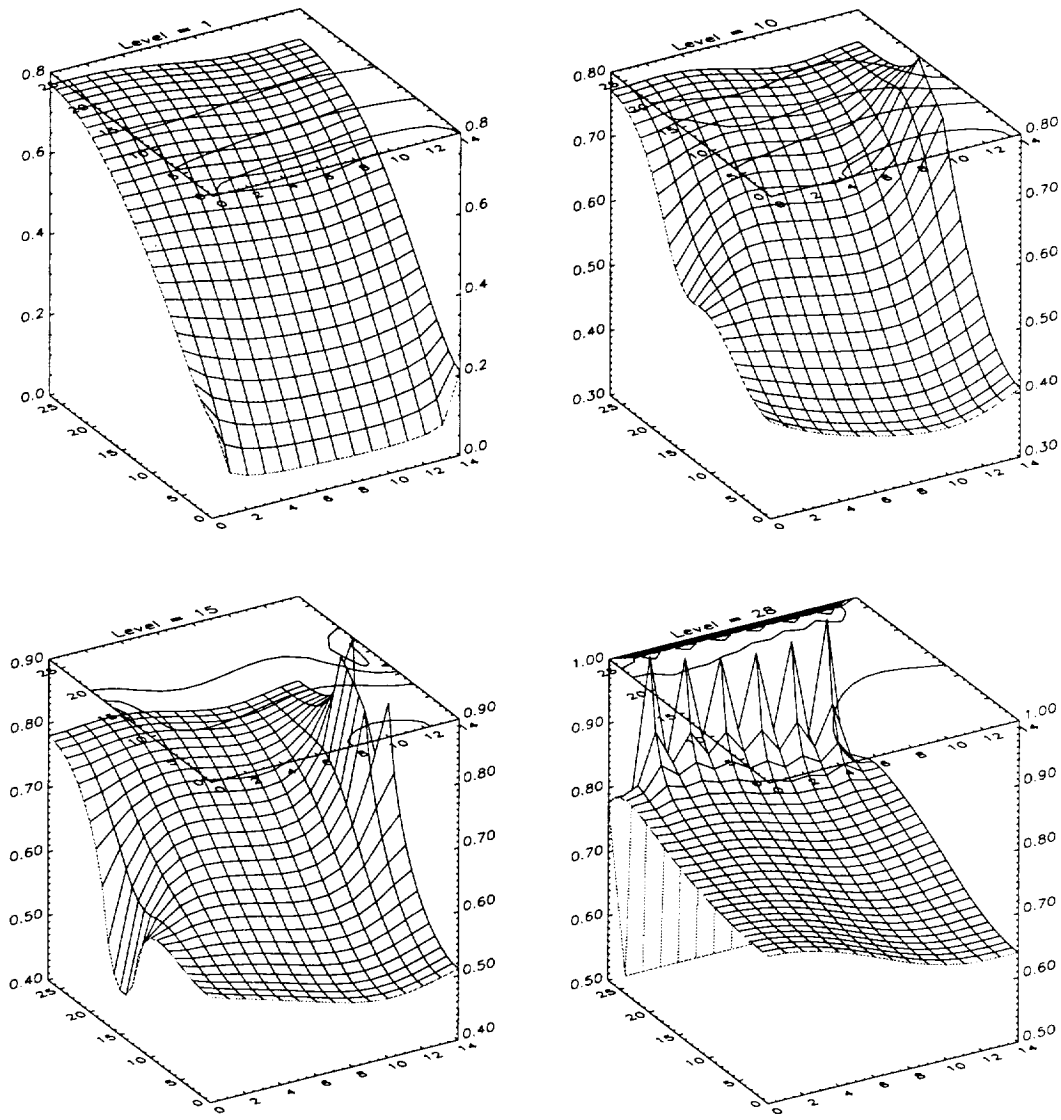

Figure 3.6: CO_2 profiles at four different elevations, about 1500 s after the sudden release of carbon dioxide at (4,5,8) with an initial carbon dioxide concentration of 0.3 volume %, an IMV inlet flow at 0.71 volume %, and a THC concentration of 0.5 volume %. The concentration profiles are shown at four levels, 0.067m (level = 1), 0.67m (level = 10), 1m (level = 15), and 1.87m (level = 28) above the floor of the cabin.

Table 3.1: Sample parameters used for the removal devices

Capacity	Boundary-type	Removal Efficiency	Name
100	1 (percentage)	0.95	CO2 scrubber
50	2 (Dirichlet)	0.0004	adsorber

the removal device can safely accumulate before it stops functioning.

In this research, the removal of all substances is monitored by a subroutine that calculates the flux of each substance at all the outlets, monitors their removal based on the input parameters, and updates the new inlet concentration boundary conditions. A sample simulation, with the sample parameters of Table 3.1 was carried out using a cabin with the removal devices in place. A source of 60 mg/L was emitted over a 10 second period, and the THC and IMV inlets had concentrations of 0.0015 mg/L and 0.002 mg/L, respectively. The concentration profiles after 450 s are shown in Figure 3.7.

This subroutine is flexible enough to model a variety of devices, including the Carbon Dioxide Removal Assembly (CDRA) that is part of the current baseline technologies in the Air Revitalization. The CDRA prototype tested(Barker et al., 1991) used a carbon dioxide removal rate that was a function of the inlet carbon dioxide partial pressure, PP_{CO_2} and was represented by the following equation:

$$Removalrate(lb/hr) = 0.1579 * PP_{CO_2}(mmHg) - 0.0348 \qquad (3.20)$$

The equation is valid for carbon dioxide partial pressures between 2.0 and 3.9 mm Hg, and can be easily included in our removal subroutine.

3.6 Summary and Conclusions

In this Chapter, the transport model was developed and implemented for air contaminant dispersal aboard spacecraft. The transport model uses the flow field

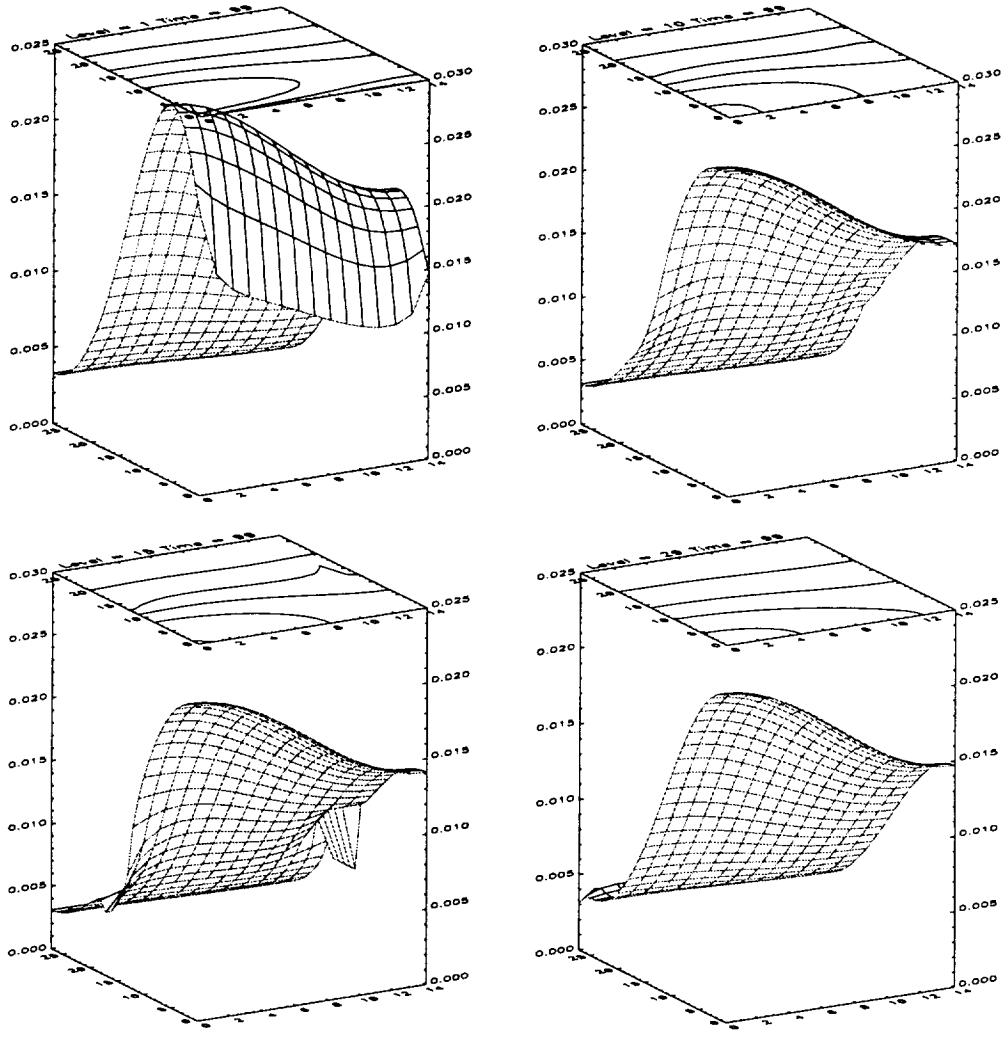

Figure 3.7: Simulation of the cabin concentrations with the removal devices in use. The parameters of the removal devices is shown in Table 3.1 and are placed at the IMV and THC outlets. The concentration surfaces are shown at four horizontal slices in the cabin, at vertical levels 1,10,15, and 29 (grid location in the z-direction out of a total of 30).

calculated using methods described in Chapter 2 as an input along with the cabin geometry and the fluid parameters, and predicts the dispersal of air contaminants as a function of time. This transport model is a three-dimensional model, an improvement over previously developed lumped and two-dimensional models, which is shown to be useful when there are wide fluctuations in the contaminant concentration in the cabin such that local violations of the SMAC level can occur even when the average across the cabin stays below the SMAC level. Convective transport was found to dominate molecular diffusion, evidenced by a mass transfer Peclet number around 2 x 10^5. The transport model accounts for sources of the contaminant, both in the inlet flows and from inside the cabin, and is also extended to model the effect of removal devices that are commonly used in the Space Station. The model is accurate for monitoring purposes, and is also computationa ly suitable for real-time application.

Chapter 4

State estimation using Implicit Kalman filtering

4.1 Why estimate the state?

State estimation is necessary while monitoring any physical process because there are always uncertainties–faults in the process under observation, errors in the mathematical model that is assumed to adequately describe the given physical process, and changes in parameters that can cause real concentrations to be different from those predicted by the model. The objective of the filtering process is to arrive at an estimate that is unbiased, i.e. has the smallest error in the least-square sense, and which gives one an accurate picture of the actual system. The cost of sensors is high, both in terms of the monetary expense, and on account of weight and electrical power issues, which necessarily restricts the number of sensors that can be carried aboard. This gives rise to the issues of placement and selection of sensors, which is an area of active research (Smith, 1996). The estimation process is very crucial to the fault detection and diagnosis process since the matrices and calculations used in fault and diagnosis procedures are used to make inferences about if and where a fault (contaminant leak) has occurred.

An identification process usually has at least three main ingredients:

- *A priori* knowledge in the form of a mathematical model about the unknown system and the noise.
- A measurement system that provides discrete or continuous measurements

of a subset of the state.

- A performance measure of the identification algorithm.

The state estimation procedure closely follows that proposed by Skliar and Ramirez (Skliar and Ramirez, 1997b). The estimation problem, formulated simply is as follows: Given a stochastic process that represents a dynamic system, we are interested in knowing the value of $x(k)$ for some fixed k, where $x(k)$ is not directly accessible to us for observation. We have a sequence of measurements that are causally related to $x(k)$ by means of a measurement system M and measurement data $z(i)$, and we wish to utilize these data to infer the value of $x(k)$. We denote the estimate of $x(k)$ by $\hat{x}(k)$ and define it to be some n-dimensional, vector-valued function ϕ_k of the measurements, viz.,

$$\hat{x}(k|j) = \phi_k[z(i), i = 1, ..., j] \tag{4.1}$$

where k refers to the time when the estimation is made, j refers to the time until which measurements are taken and used, and i is an index that refers to the measurement signal being used. Crucial to the estimation process is the definition and notion of the estimation error which is defined by the relation

$$\tilde{x}(k|j) = x(k) - \hat{x}(k|j) \tag{4.2}$$

Ideally, $\tilde{x} = 0$ and the estimate is exact. When this is not the case, we assign a penalty for the incorrect estimate. This is done through a penalty or a loss function L which has the following properties: (Meditch, 1969)

1. The loss function is a scalar-valued function of n variables.

2. $L(\tilde{x} = 0) = 0$. There is no penalty if the estimate is exact.

3. L is a non-decreasing function of the distance of the error from the origin in n-dimensional Euclidean space.

4. The loss function is symmetric about the origin in the same n-dimensional Euclidean space.

4.2 Implicit Kalman Filter

One of the classical methods of state estimation is the well established Kalman filtering algorithm (Kalman, 1960; Kalman and Bucy, 1961). The Kalman filter, which has many different implementations now, and which is widely used for the purpose of state estimation for dynamic systems that have random perturbations, is an unbiased and minimum error variance recursive algorithm to optimally estimate the unknown state of a dynamic system from noisy data taken at discrete real-time. In the Kalman filtering paradigm, the uncertainties of the model and the measurements are represented by additive stochastic white noise.

In addition, the measurements, $z(i)$ and the measurement errors $\epsilon(i)$ are assumed to possess the following properties:

(1) The measurement errors have a zero mean,

$$E(\epsilon(i)) = 0,$$

where E is the "expected value operator". A zero mean error is different from a random error in that a random error may have a non-zero expected value, which is also known as a bias. Here, we are assuming that our sensors are not biased in any particular direction.

(2) The measurement error has a constant variance, which does not change with time or with any other parameter

(3) The errors are additive, i.e,

$$z(i) = x(i) + \epsilon(i)$$

(4) The measurement errors are uncorrelated, i.e. for two measurement errors, $\epsilon(o)$ and $\epsilon(p)$,

$$cov(\epsilon(o), \epsilon(p)) = E\{[\epsilon(o) - E(\epsilon(o))][\epsilon(p) - E(\epsilon(p))]\} = 0 \; for \; o \neq p \quad (4.3)$$

The Implicit Kalman filter is one of the many alternative Kalman filters that have been developed over the years, which are theoretically equivalent to the original formulation, and which have been formulated for various desirable features (Carlson, 1990; Chin et al., 1995; Jordan, 1967) including enhanced numerical stability, computational accuracy, reduced computational requirements or for implementation using parallel computing (Jover and Kailath, 1986; Morf and Kailath, 1975; Paige and Saunders, 1977; Roy et al., 1991). Many of these variations are also discussed in textbooks (Chui and Chen, 1991) dealing with the topic of Kalman filtering. The Implicit Kalman Filter was shown to be particularly efficient in treating descriptor systems with sparse transformation matrices, an example of which is the convection-diffusion equation that forms the core of our mathematical model for air contaminant dispersion.

If we recast the model equations for the transport model developed in Chapter 3 and include the additive noise, the model can be written as a single matrix equation

$$\mathbf{A}_1\mathbf{Q}_{m+1} = \mathbf{A}_2\mathbf{Q}_m + \begin{bmatrix} \mathbf{f}_m \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{C}(m) \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} w_{1m} \quad (4.4)$$

where

$$\mathbf{Q}_m = \begin{bmatrix} \mathbf{q}_m^* \\ \mathbf{q}_m^{**} \\ \mathbf{q}_m \end{bmatrix} \quad (4.5)$$

$\mathbf{C}(m)$ represents the stochastic disturbance transition matrix that acts upon $\mathbf{w_1}$, which represents a stochastic disturbance that is an uncorrelated Gaussian white sequence with a zero mean and

$$E[\mathbf{w_{1}}_m \mathbf{w_{1}}_m{}^T] = \mathbf{Q} \tag{4.6}$$

$\mathbf{Q}$ is a diagonal matrix that represents the model noise, and is closely tied to the uncertainty surrounding the mathematical model for our physical process. A highly accurate model would have low values for its $\mathbf{Q}$, and one that does not represent the physical process too accurately would have high values for its $\mathbf{Q}$.

$$\mathbf{A}_1 = \{\mathbf{A}_1^{ij}\} = \begin{bmatrix} (\frac{-\mathbf{A}_x}{2} - \frac{\mathbf{r}}{\Delta t}) & \mathbf{0} & \mathbf{0} \\ \frac{\mathbf{r}}{\Delta t} & (\frac{-\mathbf{A}_y}{2} - \frac{\mathbf{r}}{\Delta t}) & \mathbf{0} \\ \mathbf{0} & \frac{\mathbf{r}}{\Delta t} & (\frac{-\mathbf{A}_z}{2} - \frac{\mathbf{r}}{\Delta t}) \end{bmatrix} \tag{4.7}$$

and

$$\mathbf{A}_2 = \mathbf{A}_2^{ij} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & (\mathbf{A}_x/2 + \mathbf{A}_y + \mathbf{A}_z - \mathbf{r}/\Delta t) \\ \mathbf{0} & \mathbf{0} & -\mathbf{A}_y/2 \\ \mathbf{0} & \mathbf{0} & -\mathbf{A}_z/2 \end{bmatrix} \tag{4.8}$$

and

$$\mathbf{z}_{m+1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{H}(m+1) \end{bmatrix} \mathbf{Q}_{m+1} + \mathbf{v}_{m+1} \tag{4.9}$$

$\mathbf{v}_{m+1}$ is an uncorrelated Gaussian white noise sequence that represents the measurement noise with a covariance represented by $\mathbf{R}$. The $\mathbf{R}$ matrix is a diagonal matrix that contains information about the uncertainty surrounding the measurements. Lower values for $\mathbf{R}$ indicate more accurate sensors. The estimation of the contaminant concentration is determined from the sequential solution to the following tridiagonal equations:

$$(-\frac{\mathbf{A}_x}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}^* = (\frac{\mathbf{A}_x}{2} + \mathbf{A}_y + \mathbf{A}_z - \frac{\mathbf{r}}{\Delta t})\mathbf{q}_m + \mathbf{f}_m \quad (4.10)$$
$$+\mathbf{L}_1[\mathbf{z} - \mathbf{H}_1\hat{\mathbf{y}}_{m+1|m}]$$

$$(-\frac{\mathbf{A}_y}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}^{**} = -\frac{\mathbf{A}_y}{2}\mathbf{q}_m - \frac{\mathbf{r}}{\Delta t}\mathbf{q}^* + \mathbf{L}_2[\mathbf{z} - \mathbf{H}_1\hat{\mathbf{y}}_{m+1|m}] \quad (4.11)$$

$$(-\frac{\mathbf{A}_z}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}_{m+1} = -\frac{\mathbf{A}_z}{2}\mathbf{q}_m - \frac{\mathbf{r}}{\Delta t}\mathbf{q}^{**} + \mathbf{L}_3[\mathbf{z} - \mathbf{H}_1\hat{\mathbf{y}}_{m+1|m}] \quad (4.12)$$

where the predicted estimation of the auxiliary variable $y = A_2Q$ is given by the following equation:

$$\hat{\mathbf{y}}_{m+1|m} = \begin{bmatrix} \mathbf{A}_2^{13} \\ \mathbf{A}_2^{23} \\ \mathbf{A}_2^{33} \end{bmatrix} \hat{\mathbf{q}}_{m|m} + \begin{bmatrix} \mathbf{f}_m \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \quad (4.13)$$

The superscripts in the matrices $\mathbf{A}_2^{13}$, $\mathbf{A}_2^{23}$, and $\mathbf{A}_2^{33}$ refer to the row and column partitions of the $\mathbf{A_2}$ State transition matrix.

Note that the Eq. 4.10-4.12 are identical in structure to the model equations in 3.11, the only difference being the addition of the last term, which is the effect of the filter on the equations.

The modified measurement matrix $\mathbf{H}_1 = \{\mathbf{H}_{1j}\}$ is calculated using the following equation:

$$\begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} & \mathbf{H}_{13} \end{bmatrix} \mathbf{A}_1 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{H} \end{bmatrix}, \quad (4.14)$$

which is equivalent to solving the following equations:

$$\mathbf{H}_{13}\mathbf{A}_1^{33} = \mathbf{H}, \quad (4.15)$$
$$\mathbf{H}_{12}\mathbf{A}_1^{22} = -\mathbf{H}_{13}\mathbf{A}_1^{32},$$

$$\mathbf{H}_{11}\mathbf{A}_1^{11} = -\mathbf{H}_{12}\mathbf{A}_1^{21}.$$

The Implicit Kalman Gain, a matrix that multiples the estimated error residual is given by the equation

$$\mathbf{L}_{m+1} = \mathbf{P}^{y}_{m+1|m}\mathbf{H}_1^T[\mathbf{H}_1\mathbf{P}^{y}_{m+1|m}\mathbf{H}_1^T + \mathbf{R}(m+1)]^{-1} \tag{4.16}$$

where $\mathbf{P}^{y}_{m+1|m}$ is the predicted error covariance matrix given by the equation

$$\mathbf{P}^{y}_{m+1|m} = \begin{bmatrix} \mathbf{A}_2^{13}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{13^T} & \mathbf{A}_2^{13}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{23^T} & \mathbf{A}_2^{13}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{33^T} \\ \quad +\mathbf{C}\mathbf{Q}\mathbf{C}^T & & \\ \mathbf{A}_2^{23}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{13^T} & \mathbf{A}_2^{23}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{23^T} & \mathbf{A}_2^{23}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{33^T} \\ \mathbf{A}_2^{33}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{13^T} & \mathbf{A}_2^{33}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{23^T} & \mathbf{A}_2^{33}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{33^T} \end{bmatrix} \tag{4.17}$$

The actual estimation error covariance matrix, $\mathbf{P}^{y}{}_{m+1|m+1}$ is then calculated from the predicted error covariance matrix using the relation

$$\mathbf{P}^{y}{}_{m+1|m+1} = [\mathbf{I} - \mathbf{L}_{m+1}\mathbf{H}_1]\mathbf{P}^{y}{}_{m+1|m}, \tag{4.18}$$

The error covariance matrices are measures of the uncertainty inherent in the computed quantities.

The error covariance in the state, $\mathbf{P}^{\mathbf{Q}}_{m+1|m+1}$ which is the uncertainty associated with the computed concentration is determined from the equation

$$\mathbf{P}^{y}{}_{m+1|m+1} = \mathbf{A}_1\mathbf{P}^{\mathbf{Q}}_{m+1|m+1}\mathbf{A}_1^T \tag{4.19}$$

The solution to the last equation is reduced to a sequential solution of the following six tridiagonal equations:

$$
\begin{aligned}
\mathbf{A}_1^{11}[\mathbf{P}^Q_{m+1|m+1}]^{11}\mathbf{A}_1^{11^T} &= [\mathbf{P}^y_{m+1|m+1}]^{11}, \qquad (4.20)\\
\mathbf{A}_1^{11}[\mathbf{P}^Q_{m+1|m+1}]^{12}\mathbf{A}_1^{22^T} &= [\mathbf{P}^y_{m+1|m+1}]^{12} - \mathbf{A}_1^{11}[\mathbf{P}^Q_{m+1|m+1}]^{11}\mathbf{A}_1^{21^T},\\
\mathbf{A}_1^{11}[\mathbf{P}^Q_{m+1|m+1}]^{13}\mathbf{A}_1^{33^T} &= [\mathbf{P}^y_{m+1|m+1}]^{13} - \mathbf{A}_1^{11}[\mathbf{P}^Q_{m+1|m+1}]^{12}\mathbf{A}_1^{32^T},\\
\mathbf{A}_1^{22}[\mathbf{P}^Q_{m+1|m+1}]^{22}\mathbf{A}_1^{22^T} &= [\mathbf{P}^y_{m+1|m+1}]^{22} - \mathbf{A}_1^{21}[\mathbf{P}^Q_{m+1|m+1}]^{11}\mathbf{A}_1^{21^T} -\\
&\quad \mathbf{A}_1^{21}[\mathbf{P}^Q_{m+1|m+1}]^{12}\mathbf{A}_1^{22^T} - [\mathbf{A}_1^{21}[\mathbf{P}^Q_{m+1|m+1}]^{12}\mathbf{A}_1^{22^T}]^T,\\
\mathbf{A}_1^{22}[\mathbf{P}^Q_{m+1|m+1}]^{23}\mathbf{A}_1^{33^T} &= [\mathbf{P}^y_{m+1|m+1}]^{23} - \mathbf{A}_1^{21}[\mathbf{P}^Q_{m+1|m+1}]^{12}\mathbf{A}_1^{32^T} -\\
&\quad \mathbf{A}_1^{22}[\mathbf{P}^Q_{m+1|m+1}]^{22}\mathbf{A}_1^{32^T} - \mathbf{A}_1^{21}[\mathbf{P}^Q_{m+1|m+1}]^{13}\mathbf{A}_1^{33^T},\\
\mathbf{A}_1^{33}[\mathbf{P}^Q_{m+1|m+1}]^{33}\mathbf{A}_1^{33^T} &= [\mathbf{P}^y_{m+1|m+1}]^{33} - \mathbf{A}_1^{32}[\mathbf{P}^Q_{m+1|m+1}]^{22}\mathbf{A}_1^{32^T} -\\
&\quad \mathbf{A}_1^{32}[\mathbf{P}^Q_{m+1|m+1}]^{23}\mathbf{A}_1^{33^T} - [\mathbf{A}_1^{32}[\mathbf{P}^Q_{m+1|m+1}]^{23}\mathbf{A}_1^{33^T}]^T
\end{aligned}
$$

where $[\mathbf{P}^Q_{m+1|m+1}]^{33} = \mathbf{P}^q_{m+1|m+1}$.

We can now formulate the algorithm of the estimation of the contaminant concentration $\mathbf{q}_{m+1}$ based on the measurement data and the transport model.

1. Compute the predicted estimate of the concentration, $\mathbf{y}_{m+1|m}$ by propagating the concentration at the previous time step, $\hat{\mathbf{q}}_{m|m}$ according to Eq. 4.13.

2. Successively solve three tridiagonal matrix equations for the modified measurement matrix $\mathbf{H}_1$.

3. Solve the tridiagonal model equations with the new perturbation to obtain the optimal estimate $\mathbf{q}_{m+1|m+1}$.

The calculation of the gain $\mathbf{L}_{m+1}$ of the Implicit Kalman Filter follows the following algorithm:

1. Calculate $\mathbf{P}^y_{m+1|m}$ according to Eq. 4.17.

2. Calculate the Implicit Kalman filter gain from Eq. 4.16.

3. Calculate $\mathbf{P}^y_{m+1|m+1}$ according to Eq. 4.18.

4. The final step is to initiate the gain calculation for the next time step, for which one sequentially solves Eq. 4.21. This is by far the most time consuming step. The updating step can also be implemented through the use of square-root filtering schemes, which can reduce the computation time, and also provide some added stability to the filter (Skliar and Ramirez, 1997b).

4.3 Filter Implementation and Testing

A "true" test of the filter can only occur in an experimental setting, with a physical cabin and measurements. In the absence of that, we tested the filter using the results of the model itself. We added a random Gaussian noise to the contaminant concentrations from the model, as we would expect to get in a real setting and then checked to see if the filter was able to track the contaminant concentrations with sufficient accuracy.

The filter is a computationally intensive program, and we found that it was taking around 30 seconds for each time step on a DEC-Alpha Station 250 4/266. We therefore ran the model and the filter with a 30s time step. We are using a set of five sensors to estimate the concentration in the whole room. In order to test the filtering process, we have crowded all of the sensors in one corner of the cabin, to see how robust the filtering process is. Table 4.1 shows the position of the sensors. The co-ordinates refer to the location of the sensor grid points within a rectangular geometry of size 6 x 2 x 2m which used a grid of dimensions 15 x 28 x 30. For the filter to be useful, good estimates of the model uncertainty and the measurement uncertainty are needed. This estimation is not a trivial process, and will require among other things, experience in running the filter for specific systems and conditions. The model uncertainty, Q_{diag} can be closely tied to the amount of turbulence in the system and the error in the numerical solution to the convection-diffusion equation. The measurement uncertainty, $\mathbf{R}_{diag}$ should reflect the uncertainty present in each sensor, which is quantifiable by test experiments.

Table 4.1: Sensor location and the associated measurement noise

Sensor	Co-ordinates	R
1	3,2,3	0.001
2	1,1,3	0.001
3	2,2,2	0.001
4	3,1,2	0.0019
5	2,4,3	0.0019

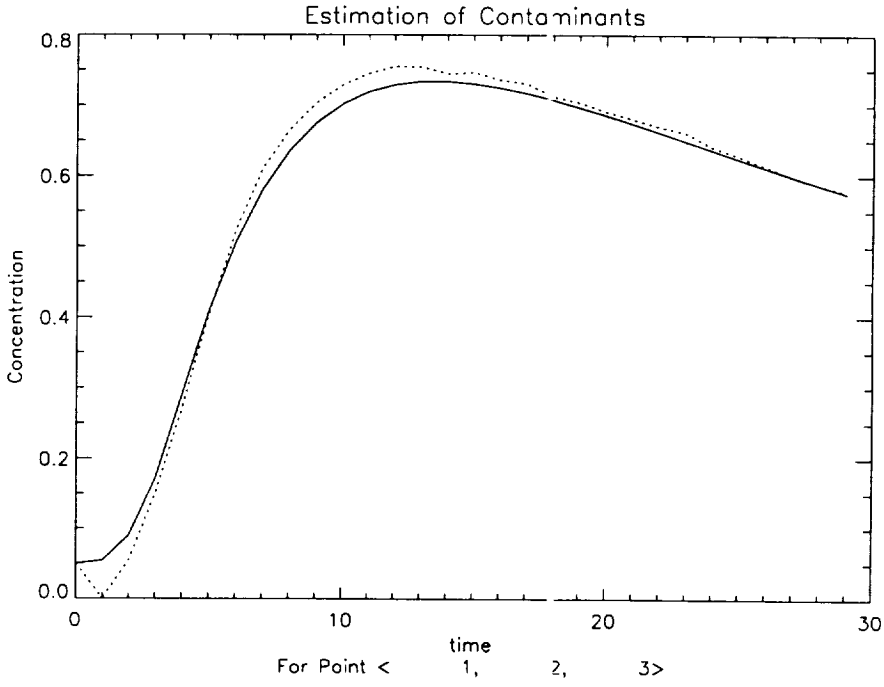

Figure 4.1: Filter Performance–Tracking an arbitrary point in the cabin. The real concentration at point (1,2,3) is given by the solid line, and the estimated concentration from the filter is given by the dotted line.

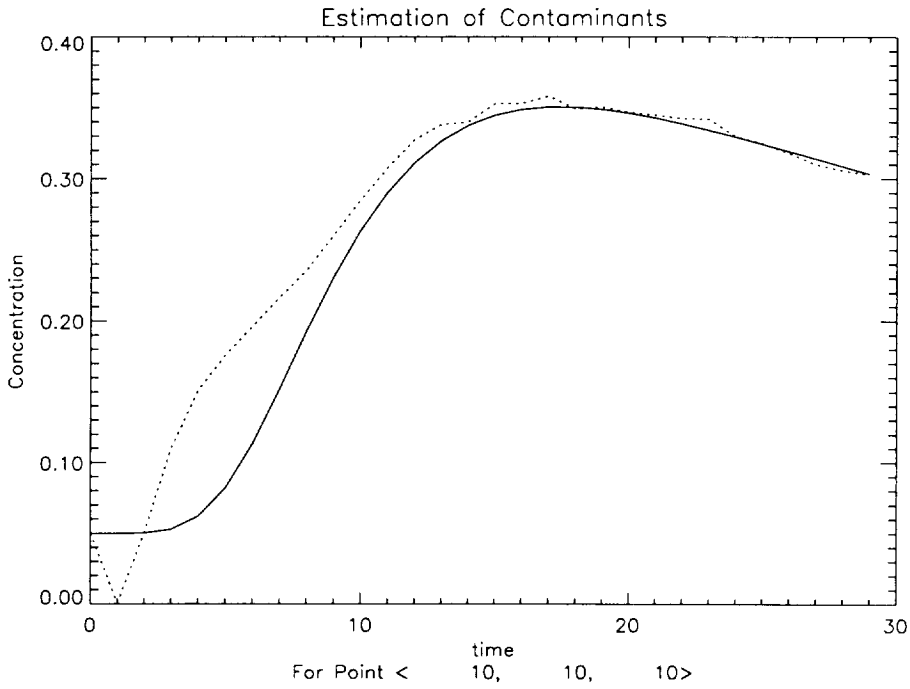

Figure 4.2: Filter Performance at (10,10,10). The real concentration is given by the solid line, and the estimated concentration from the filter is given by the dotted line.

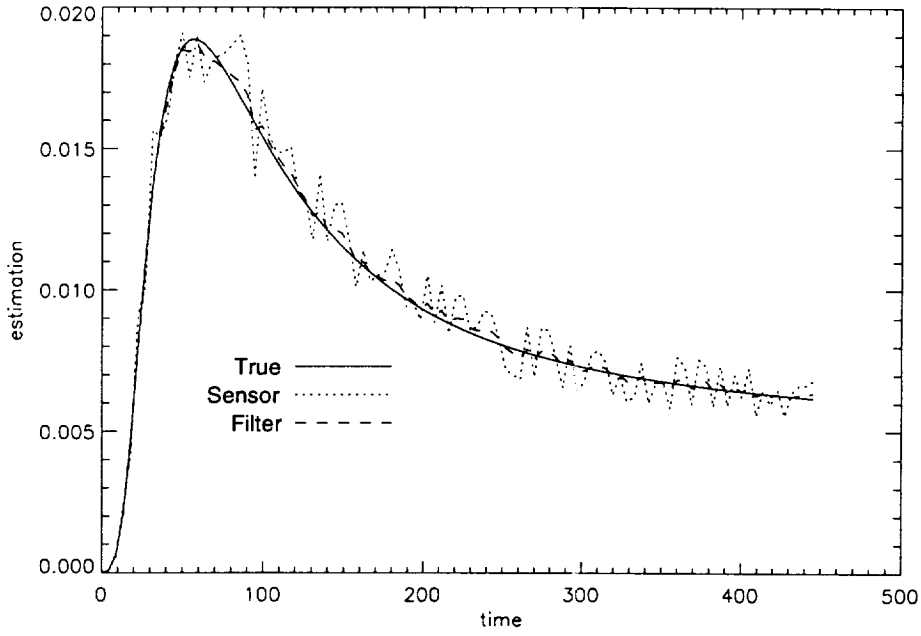

Figure 4.3: Filter Performance–Tracking at Sensor #1, located at (3,2,3). The solid, dotted, and dashed lines indicate the true concentration, the measured concentrations, and the optimal estimate of the concentrations using the Implicit Kalman Filter, respectively.

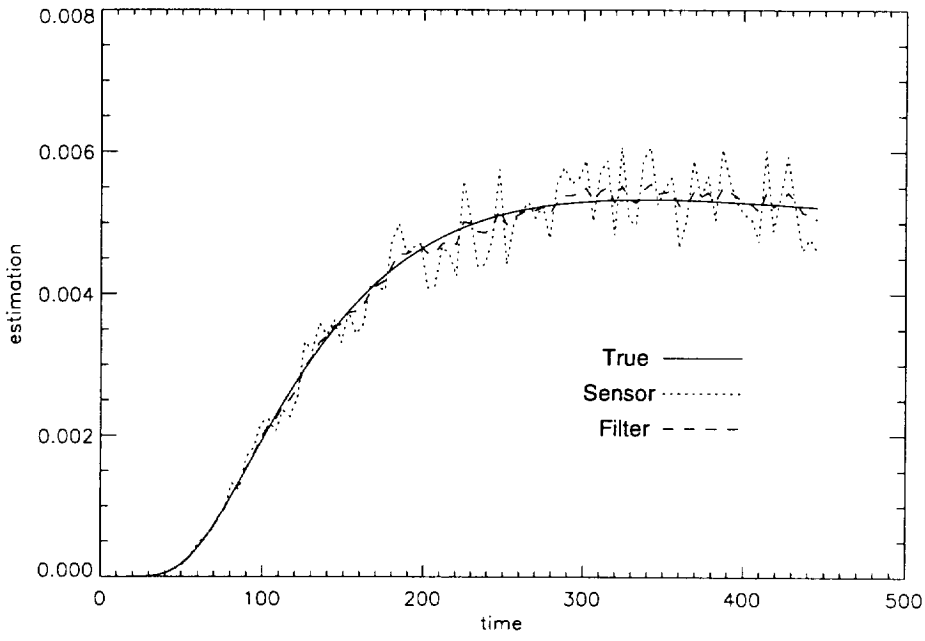

Figure 4.4: Filter Performance–Tracking at Sensor #2, located at (1,1,3). The solid, dotted, and dashed lines indicate the true concentration, the measured concentrations, and the optimal estimate of the concentrations using the Implicit Kalman Filter, respectively.

The graphs show sample results of the contaminant concentration estimation for 30 time steps. Figures 4.1 and 4.2 show the tracking of two arbitrary points in the domain, while Fig. 4.3-4.4 show how the filter functions at the sensor locations in the presence of measurement noise. The tracking is fairly accurate and is mathematically consistent. The error for points in the domain that are distant from the sensors is naturally higher. The filter slightly lags the "true" values. That, too, should be expected, since the filter does not "know" about new source emissions.

Figure 4.5 shows the error at four different sensor locations as a function of time. The dashed line, which is the bound of the error is determined by the estimation error covariance at that location. In this example, we have chosen 3σ as the bound, where σ represents the standard deviation of the expected estimation error, based on the fact that 99 % of the estimates would fall within this bound. It is expected that the error stays within the bounds that are indicated by the uncertainty, which changes from its initial value until it reaches a steady state value. Figure 4.6 shows how the uncertainty varies with time.

4.4 Uncertainty

4.4.1 Modeling uncertainty

The main source of uncertainty in the system being studied is that of the wind velocities. The flow field is by far the most important parameter in the dispersal of the contaminants. In addition, there are the other usual uncertainties inherent in the physical modeling of any system, errors in the measurement, and the presence of faults. There are two reasons why there are uncertainties in the flow field. Firstly, turbulence is stochastic in nature, and the flow field obtained, by definition, is approximate since it is a time averaged quantity. Graphs showing the actual measured velocities indicate this quite clearly (Zhang et al., 1992; White, 1974). The other reason is that numerical procedures essentially yield approximate results. The

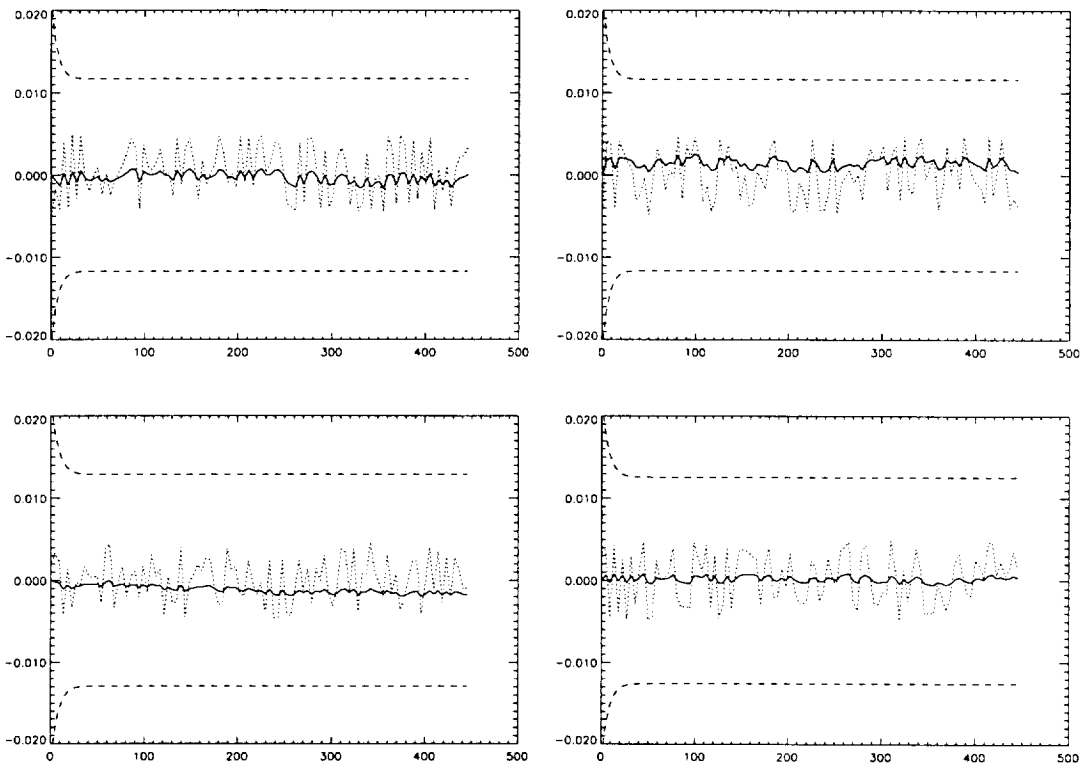

Figure 4.5: Filtering results showing the noisy measurements and the filtered estimates; the ordinate represents the measurement noise (dotted line) and the estimation error (solid line) as a function of time. The dashed line represents the error bounds, and is equal to 3σ, where σ represents the standard deviation of the expected estimation error

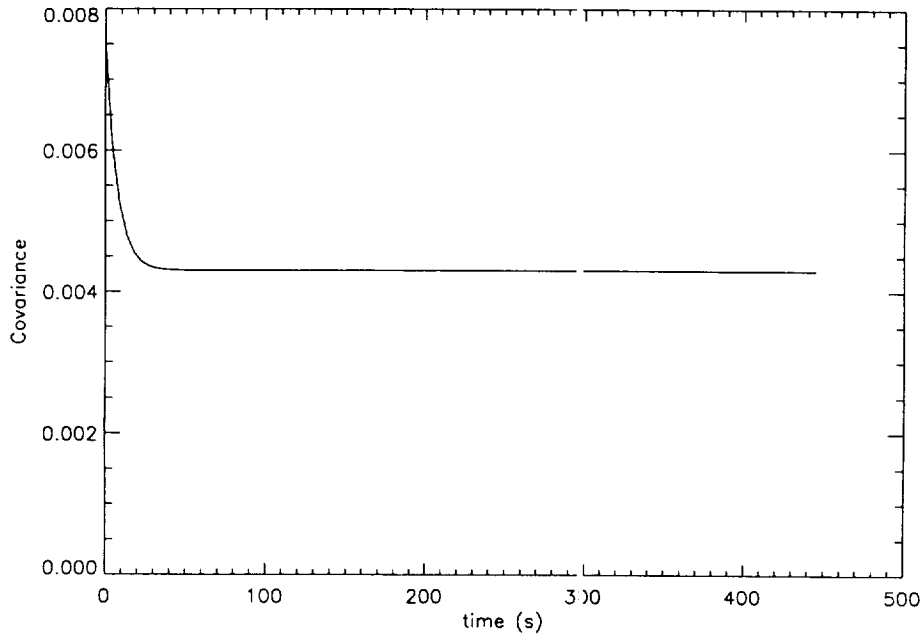

Figure 4.6: Filter uncertainty – Covariance of the estimation error at Sensor #3 as a function of time.

residual norms of the errors are a measure of how approximate the numbers are. The objective in studying the uncertainty is to structure the uncertainty in order to produce a measure of how reliable our estimate of the state is and how robust the procedure is. It is crucial that the diagnosis be accurate, that no fault go undetected; at the same time, any false alarms must be avoided since that would mean the loss of precious crew time. Most fault detection procedures will ultimately detect any given fault; the objective is to detect the fault as early as possible using a procedure that is known not to yield too many (or any) false alarms. The structuring of the uncertainty therefore, is a prerequisite to the fine tuning of the fault detection algorithm.

An analysis of the uncertainty would require an experimental set-up, and data in order to evaluate the model results. In the absence of that, one could artificially change some parameters, and then evaluate the performance of the filter and the model. We next consider the effect of a change in the inlet velocity on the model results.

4.4.2 Effect of randomness in inlet velocity

The dispersion model assumes the existence of a steady state flow profile. Here, we examine the effects of a variation in the inlet velocity on the estimation error.

For the turbulent flow field simulated in Chapter 2, the THC inlet velocity is specified to be 0.5 m/s. Now, assume that the velocity decreases by 15 %.

The steady-state Navier-Stokes equations are next solved for the new boundary conditions. A study of the actual velocities shows that some internal velocities decrease by more than 15 %, depending on the location, not a surprising result since the Navier-Stokes equations are non-linear, and the turbulent energy itself causes substantial noise in the system.

The new velocities are used in conjunction with the mathematical model

Table 4.2: Sensor location and the associated measurement noise

Sensor	Co-ordinates	R
0	10 10 10	0.0001
1	8 2 23	0.001
2	3 5 7	0.0001
3	6 6 6	0.0023
4	4 11 22	0.019
5	8 14 18	0.0019

to generate the "true" concentrations for the given concentration initial and boundary conditions. The filter, however, is not updated with these velocities, since our objective is to evaluate its performance when conditions change unknown to the model. Monitoring the inlet flow velocity would take care this change and update the model, a procedure that is discussed in the next section. In the first case, we use the uncertainties from Table 4.2.

The tracking at Sensor locations 0 and 4 is shown in Figs. 4.7 and 4.8. The filter uncertainties in this case balance the model uncertainties, and the filter consistently over-predicts the concentrations at both locations. The main reason for this is the fact that the reduced inlet velocities causes a reduced flux of the contaminant into the chamber, since the inlet concentrations remains unchanged. This 15 % reduction in the flux persists throughout the duration of the experiment, and the filter, constrained by the mass balance over predicts the concentration.

The residual error curves for Sensors 0 and 1, shown in Figs. 4.9 and 4.10 show the error that is negative (with the estimated concentrations being lower than the actual) and increases with time. The errors increase at different rates, and the rates depend on the local velocity and location with respect to the ducts.

The Euclidean norm of the prediction error, shown in Fig. 4.11 rises dramatically, and clearly exceeds the error bounds, and the fault in the system is quite apparent.

This case dealt with the situation in which the model and sensor uncertain-

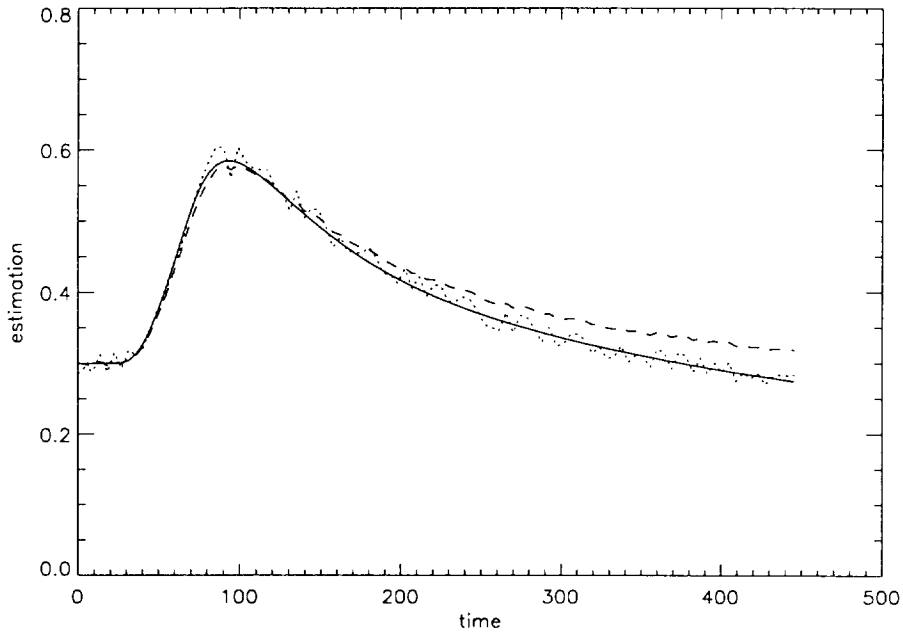

Figure 4.7: Tracking of the contaminant concentration at sensor location #0 with a 15 % reduction in the inlet velocity. The filtered estimate (dashed line) over-predicts the real concentration (solid line). The measurements are shown by the dotted line.

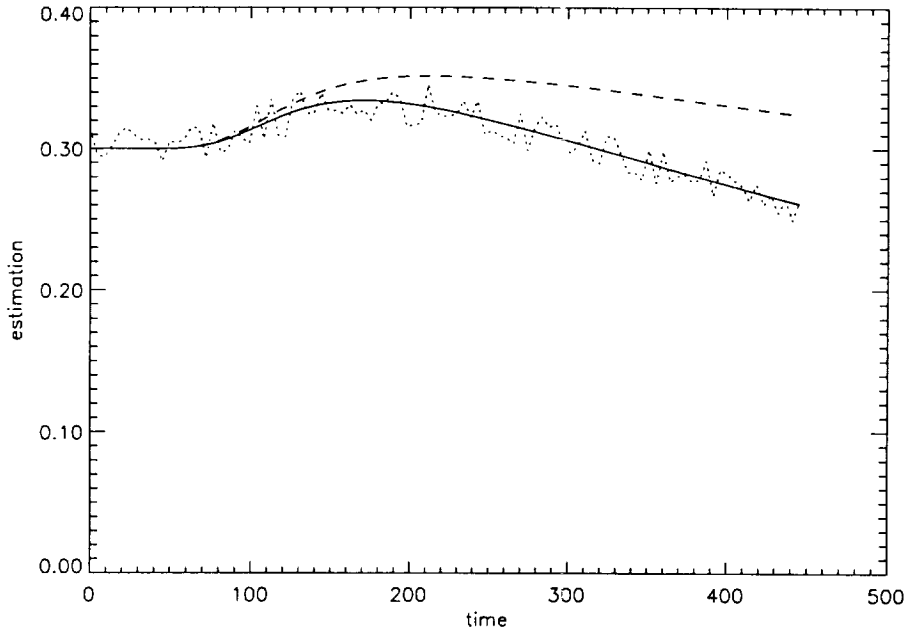

Figure 4.8: Tracking at location 4 with a 15 % reduction in the inlet velocity. The filtered estimate (dashed line) over-predicts the real concentration (solid line). The measurements are shown by the dotted line.

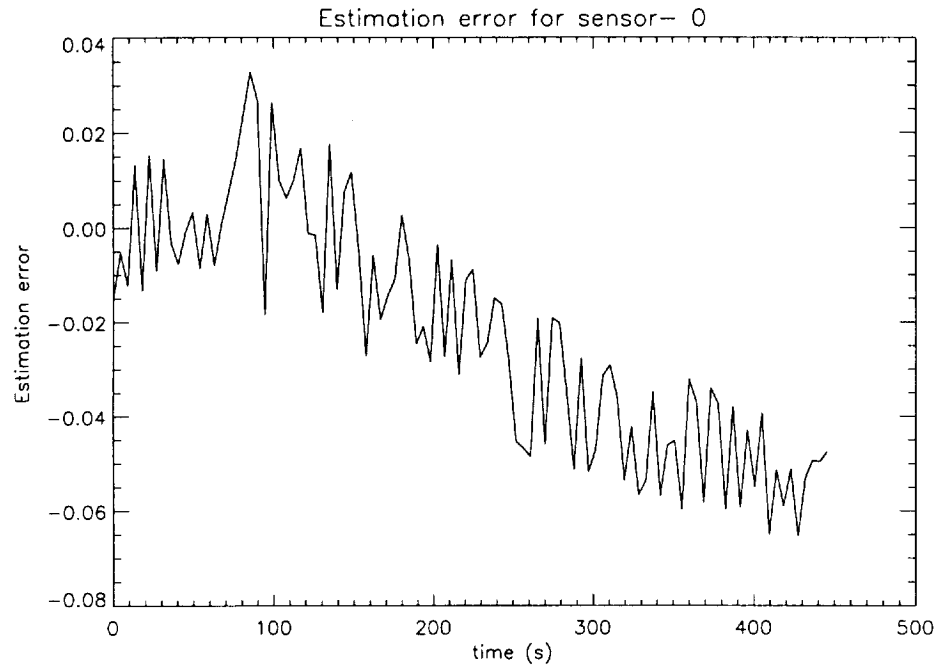

Figure 4.9: Residual error at sensor location 0 with a 15 % reduction in the inlet velocity. The magnitude of the error is increasing with time.

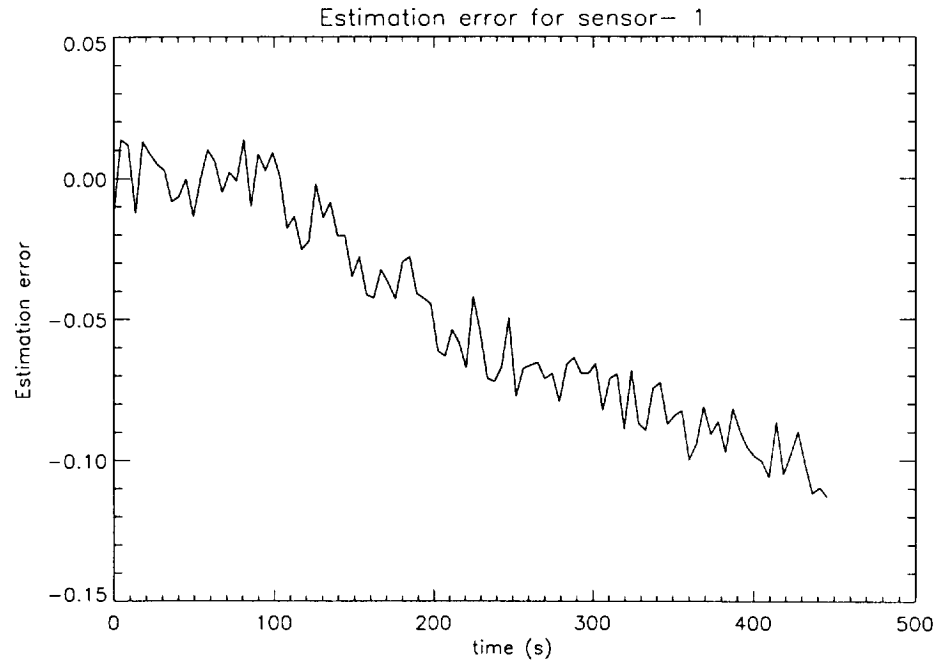

Figure 4.10: Residual error at sensor location 1 with a 15 % reduction in the inlet velocity. The magnitude of the residual (estimation) error increases with time.

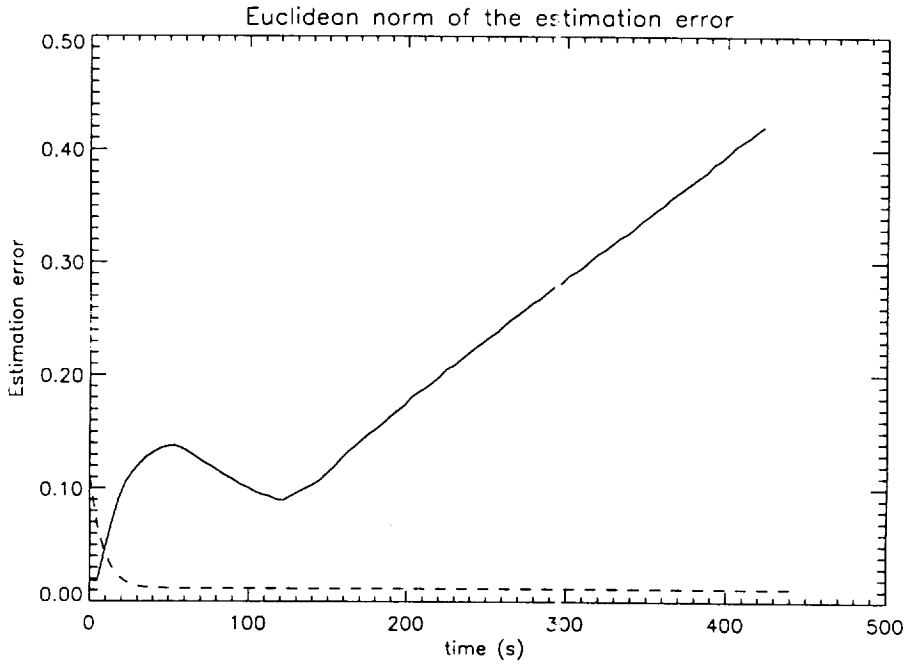

Figure 4.11: The Euclidean norm of the error with a 15 % reduction in the inlet velocity is shown increasing with time. The estimation error clearly exceeds the error bounds (dashed line), indicating the presence of a fault.

Table 4.3: Sensor location and the associated measurement noise- very low sensor uncertainty

Sensor	Co-ordinates	R
0	10 10 10	0.000001
1	8 2 23	0.00001
2	3 5 7	0.000001
3	6 6 6	0.000023
4	4 11 22	0.00019
5	8 14 18	0.000019

ties were weighted equally. If we expect extensive noise in the system, for example, if a 15 % change in the inlet velocity were common, one would then have to give the measurements a higher certainty level in order to obtain better tracking performance.

For this situation, we repeat the previous case, but with the lower measurement uncertainties listed in Table 4.3. The sensor uncertainties, R_{diag} have been reduced to about one hundredth of their usual values. The results are shown graphically in Figs. 4.12 through 4.16. One can clearly observe that the tracking using these very low uncertainties is an improved over the previous case. In this scenario, the model is relatively unimportant, and the measurements become paramount. This all comes at a cost, of course. With these very low sensor uncertainties the ability to filter out the noise is impaired.

4.4.3 Double Filter

Since one of the major sources of uncertainty and error in this model is the velocity, a variant of the Implicit Kalman filter such as a double filter could be used for improved performance. The double filter is a Kalman filter, but the difference between the measured and model velocity is also used to update the final state of the system. It is expected that this will increase the sensitivity of the filter and consequently increase accuracy. The filter will therefore have a set of concentration sensors, measurements from which are used to update the right side of the state

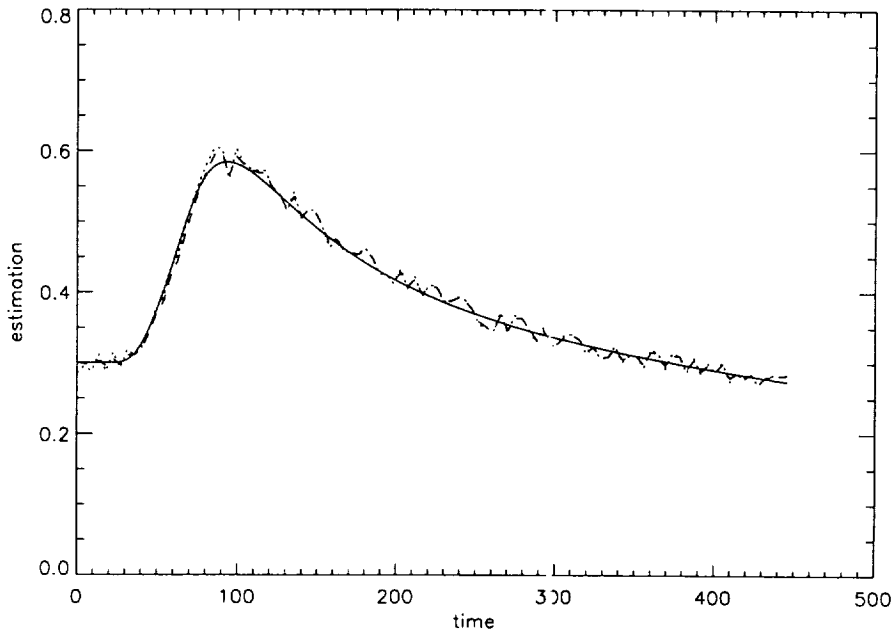

Figure 4.12: Tracking at sensor location 0 with a 15 % reduction in the inlet velocity for very low measurement uncertainties. The filter (dashed line) now tracks the measurements (dotted line) very closely.

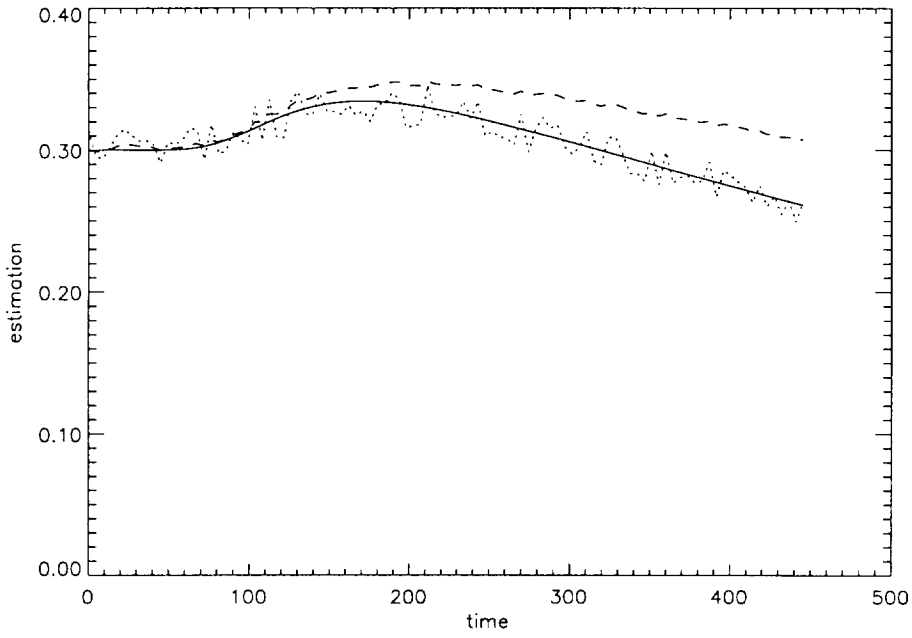

Figure 4.13: Tracking at sensor location 4 with a 15 % reduction in the inlet velocity for very low measurement uncertainties. The filter (dashed line) still over-predicts the concentration, but the tracking is slightly better than in Fig. 4.8.

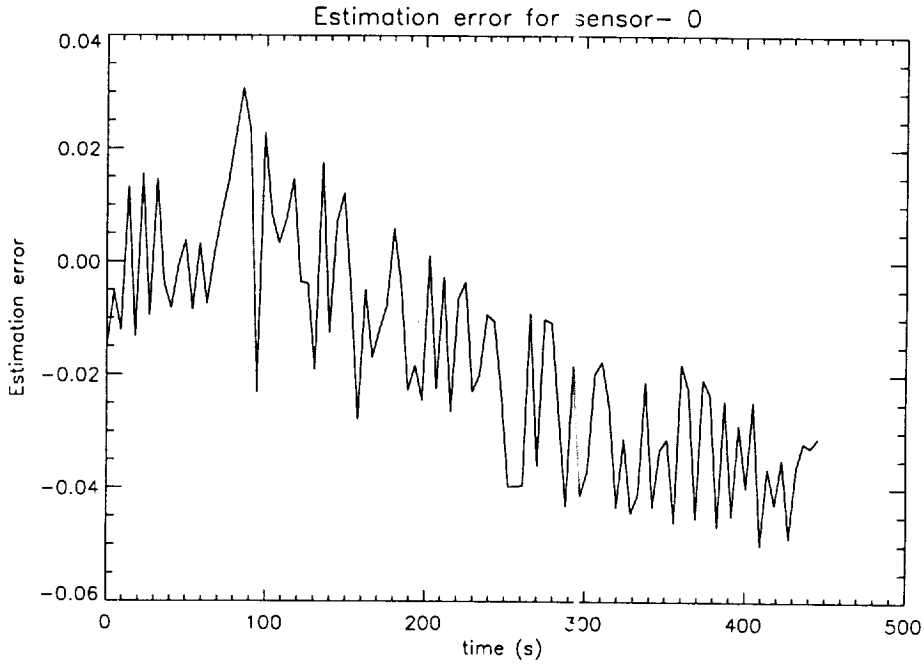

Figure 4.14: Residual error at sensor location 0 with a 15 % reduction in the inlet velocity for very low measurement uncertainties. The presence of the residual error indicates the presence of a fault, but the error in this case is lower in magnitude than with higher sensor uncertainties.

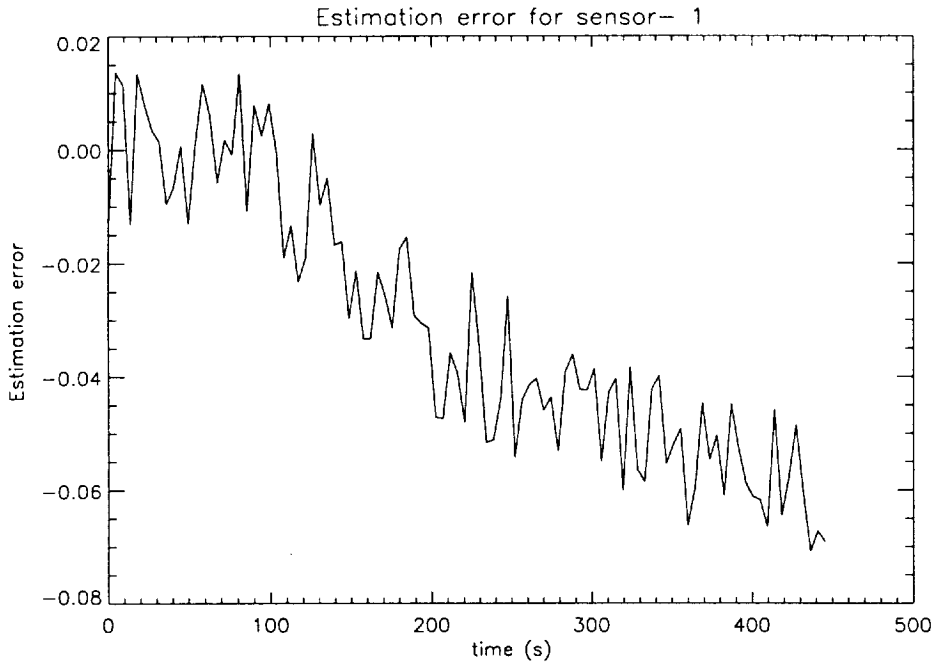

Figure 4.15: Residual error at sensor location 1 with a 15 % reduction in the inlet velocity for very low measurement uncertainties. The presence of the residual error indicates the presence of a fault, but the error in this case is lower in magnitude than with higher sensor uncertainties.

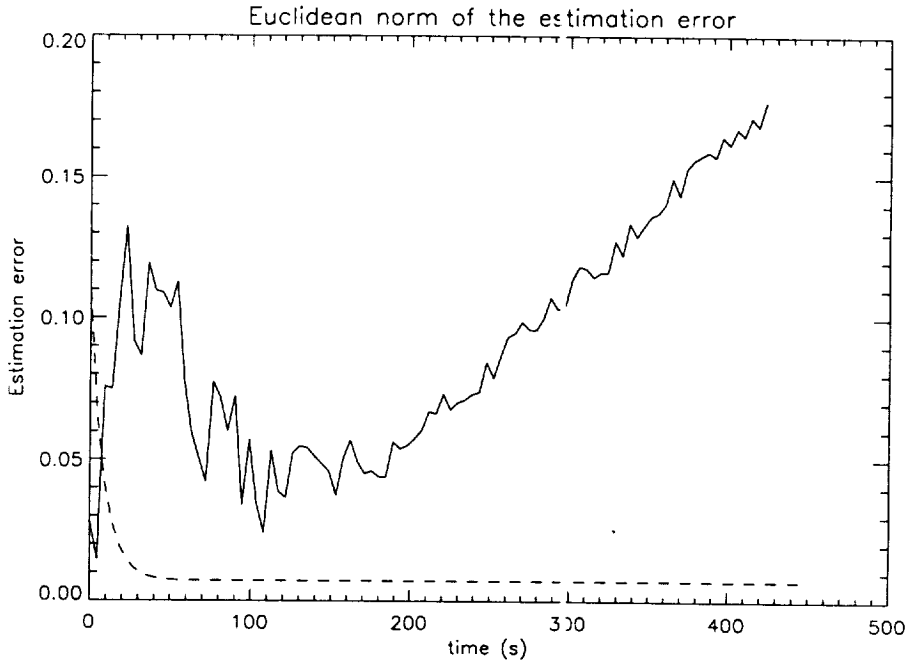

Figure 4.16: Euclidean norm of the error with a 15 % reduction in the inlet velocity for very low measurement uncertainties. The presence of a fault in the system is evidenced by the fact that the residual error (solid line) clearly exceeds the error bounds (dashed line).

estimation equations, and a set of velocity sensors, which will perturb the left-hand side of the state estimation equations. The new state estimation equations, will therefore be of the form,

$$\begin{aligned}
(-\alpha_x/2 - 1/\Delta t)q^* &= (\alpha_x/2 + \alpha_y + \alpha_z - 1/\Delta t)q_m + f_m + L_1[z - H_1\hat{y}_{m+1|m}] \\
(-\alpha_y/2 - 1/\Delta t)q^{**} &= -\alpha_y/2 q_m - 1/\Delta t q^* + L_2[z - H_1\hat{y}_{m+1|m}], \\
(\frac{-\alpha_z}{2} - 1/\Delta t)q_{m+1} &= \frac{-\alpha_z}{2} q_m - 1/\Delta t q^{**} + L_3[z - H_1\hat{y}_{m+1|m}],
\end{aligned} \tag{4.21}$$

where α_x, α_y, and α_z represent the filtered left hand side terms. The time step used in updating the velocity filter could be much larger than the time step used in the rest of the filter, in order to minimize the computations needed.

Implementation of this double filter was out of scope for this work, but can be easily implemented and tested.

4.5 Summary and Conclusions

A state estimation procedure was implemented using the Implicit Kalman Filter, which provides accurate estimates of the contaminant concentrations at all points in the cabin, using the transport model developed in Chapter 3 in conjunction with a measurement system. The filter is an effective tool for rejecting sensor noise, and provides smooth estimates of the state of the system in real time. The performance of the filter in the presence of a major disturbance was studied, which showed that a proper choice of the model and measurement noise covariance matrices can lead to good tracking behavior in the presence of noise. Parameter estimation, and its use in combination with the state estimation procedure could lead to further improvements in overall estimation performance.

Chapter 5

Fault detection in distributed parameter systems

5.1 Introduction

In this chapter, we discuss the use of the Implicit Kalman Filter in the implementation of a Fault Detection algorithm. Fault detection is the procedure which alerts the user to a malfunction in the system. Fault detection is the first step in the comprehensive Fault Detection and Isolation (FDI) problem. Basseville, in his discussion of current methods in FDI (Basseville, 1997), mentions that FDI is split into two steps; the *generation of residuals*, which are ideally zero under fault-free conditions, minimally sensitive to noises and disturbances, while being maximally sensitive to faults, and *residual evaluation*, which concerns the design of decision rules based on these residuals. A detailed account of fault detection algorithms can be found elsewhere (Basseville and Nikiforov, 1993). Methods of fault detection are classified into methods that are *Model based* or those that are *Statistically based.* Statistical methods are the only options when detailed model information is not available and consists of continuously examining the statistical properties of the measurement data, and noting any substantial deviations from a pre-determined threshold band. Model based methods, on the other hand, use the knowledge about the system and infer unmeasurable characteristics of the system from the measurable using this knowledge.

The fault detection algorithm consists of a sensor fault test procedure and

a process fault detection procedure. The sensor fault testing procedure forms the inner shell of the algorithm, since the sensor readings will have to be validated before further processing. Then, the readings are evaluated for a possible process fault. Under normal conditions, both the tests would be negative, and the filter would continue onto its next time step. For the purposes of air contaminant monitoring in spacecraft, we envision two kinds of faults; *Instrument or sensor faults* and *Process faults.* A Sensor fault, as its name suggests, implies that one or more of the sensors is not functioning. The fault could be in the form of a total malfunction, where the sensor readings are totally random with no physical basis, or could be manifest as a bias in some direction.

A process fault may be present as an unknown source problem, or as a violation of a safety requirement. While we expect that faults will be infrequent, a fault detection procedure is crucial since it is under conditions of a fault that the utility of the system is realized.

5.2 Sensor Fault

One method for detecting sensor faults is that of hardware redundancy (Emami-Nacini et al., 1986), in which sensors are used at each location, and an agreement between all three sensor readings is necessary for that sensor reading to be accepted as valid. If one of the three sensors shows a deviation statistically significant from the other two, then that sensor is considered to be faulty. Hardware redundancy is costly because of its need for triple the amount of hardware especially in a space environment, where weight are power requirements are concerns. Hardware redundancy, though might become feasible for on-board applications if the cost and weight of the sensors drop enough to offset the cost of the extra computational requirement due to analytical redundancy.

Another approach is called analytical redundancy. Analytical (or functional) redundancy is a model-based fault detection procedure where a single set of

measurements is used in conjunction with a model and a detection algorithm in order to detect and distinguish between faults. Here, the set of sensors is separated into two. In our example, Sensors 1-3 (Table 4.1) are part of Filter bank 1; Sensors 4-6 part of Filter bank 2, and the complete set of sensors 1-6 form part of Filter Bank 0. The three filters are run simultaneously, one with the complete set of sensors, and one with each individual set of sensors. Three residuals characterize these banks of sensors,

$$\mathbf{r_0} = \mathbf{z}_{m+1} - \mathbf{H_1}\hat{\mathbf{q}}_{m+1|m+1} \tag{5.1}$$

$$\mathbf{r_1} = \mathbf{z}^1_{m+1} - \mathbf{H_1}^1\hat{\mathbf{q}}^1_{m+1|m+1} \tag{5.2}$$

$$\mathbf{r_2} = \mathbf{z}^2_{m+1} - \mathbf{H_1}^2\hat{\mathbf{q}}^2_{m+1|m+1} \tag{5.3}$$

where the superscripts refer to the appropriate sensor banks. When all three banks show a similar residual, the sensor system is working normally. When two of the banks deviate, then one of the sensors in that bank is malfunctioning, and the other bank alone should be used in the filtering process. Mathematically, this amounts to checking if

$$\|\mathbf{r_0}\|_2 = \|\mathbf{r_1}\|_2 = \|\mathbf{r_2}\|_2 \tag{5.4}$$

within bounds of error at each time step. $\|\mathbf{r}\|_2$ represents the Euclidean norm of the quantity, which represents the distance of the vector from the origin.

Figure 5.1 shows the residuals calculated for three different banks in the absence of a fault, while Fig. 5.2 shows them in the presence of a fault in Bank 0 and Bank 1. In Fig. 5.1, the residuals of the estimation error in all three banks is around 0.015 throughout the time period under consideration. In Fig. 5.2, however, while Bank 2 maintains its residual of 0.015, Banks 0 and 1 have residuals around 1.5, which are significantly higher. Bank 2 is therefore functioning normally, and its

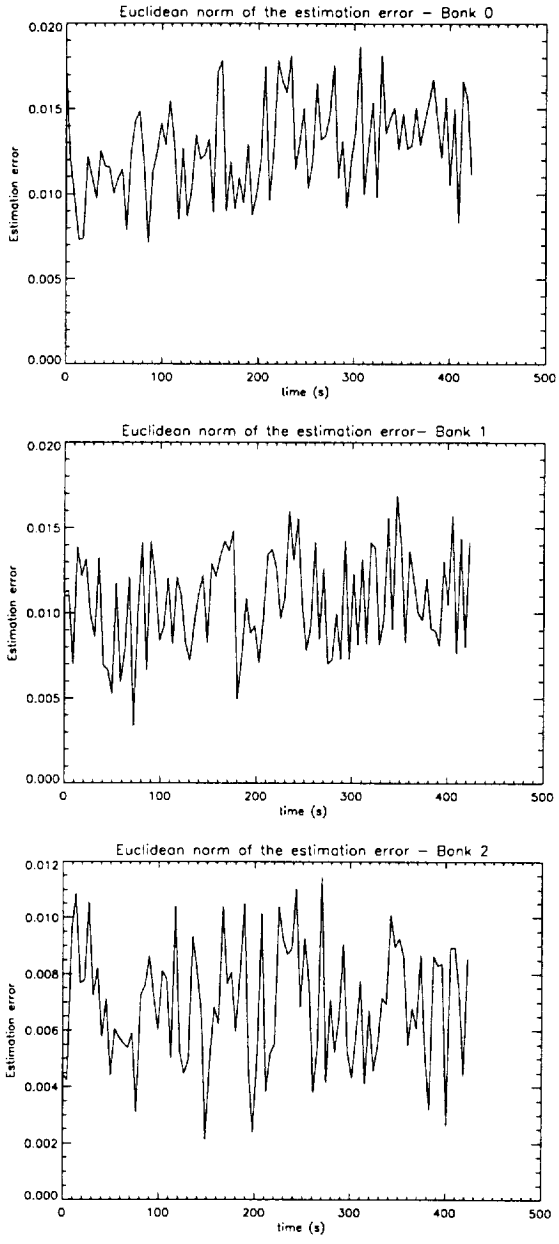

Figure 5.1: Residuals of the three Filter banks under normal operation. Note that all three banks posses estimation errors that are similar in magnitude.

estimates alone should be used to monitor the contaminants until the faulty sensor is identified and rectified.

This method requires high computational power but does not need an interruption of the operation once a fault is detected, since a smooth transition to the functioning bank of sensors can be made. Alternatively, one can just monitor each individual sensor for its residual and determine a malfunction. Figure 5.3 shows the residual error for three sample sensors in the case of a sensor fault. It is quite apparent from the three sensor error graphs that Sensor 2 is malfunctioning, and that the others are functioning normally. The working of the filter can be seen. Since the filter does not "know" that a fault has occurred, the filter is starting to respond, and even the non-faulty sensors are showing errors that are almost exceeding the bounds.

This procedure required less computational power, but the filtering process after the detection of the fault becomes complicated, because of the process of taking out the malfunctioning sensor's readings from the existing filter.

5.3 Unknown source

The problem of an unknown source could be something as minor as higher than expected carbon dioxide levels because of increased activity in the cabin or could be a major leak. An unknown source will cause the model to vary considerably from the measurements and cause large estimation errors. Monitoring the estimation error is the key to identifying the presence of unknown source substances and contaminants. The way this is detected is through having safe bounds for the residual error, and if the residual error exceeds the bounds, then the diagnosis for an unknown source is initiated. Figure 5.4 shows a sample result where the error bound is exceeded owing to an unknown source. The case of an unknown source is discussed in greater detail in the next section, in a simulated Space Station scenario.

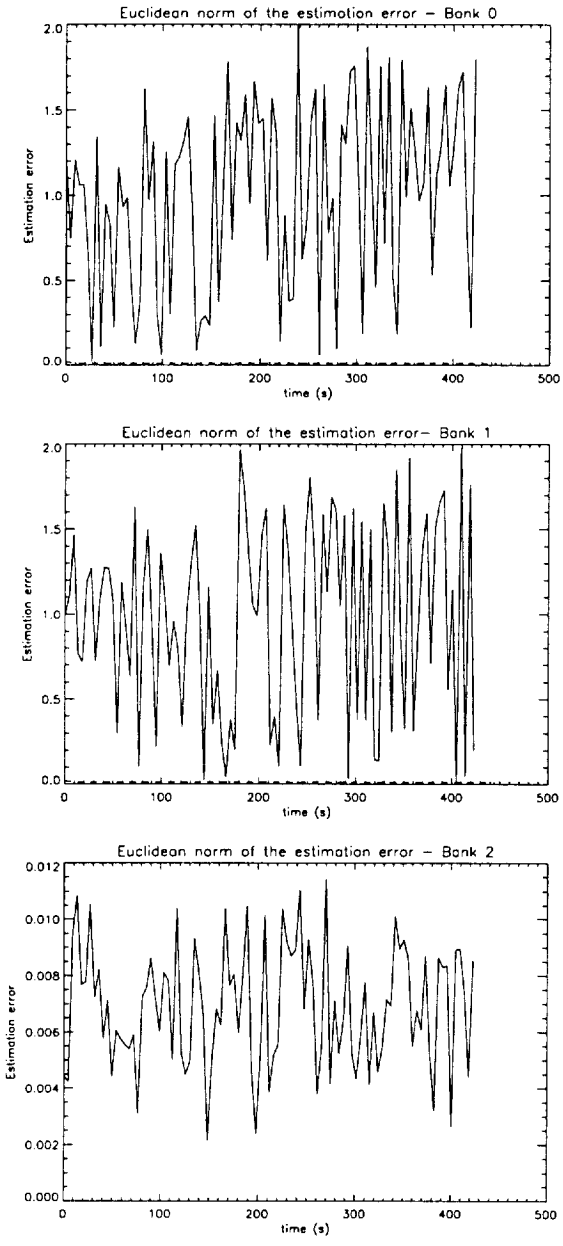

Figure 5.2: Residuals of the three Filter Banks with a fault in Sensor #2. Note the huge change in the estimation error in the faulty banks (0 and 1) from normal operation (Bank 2).

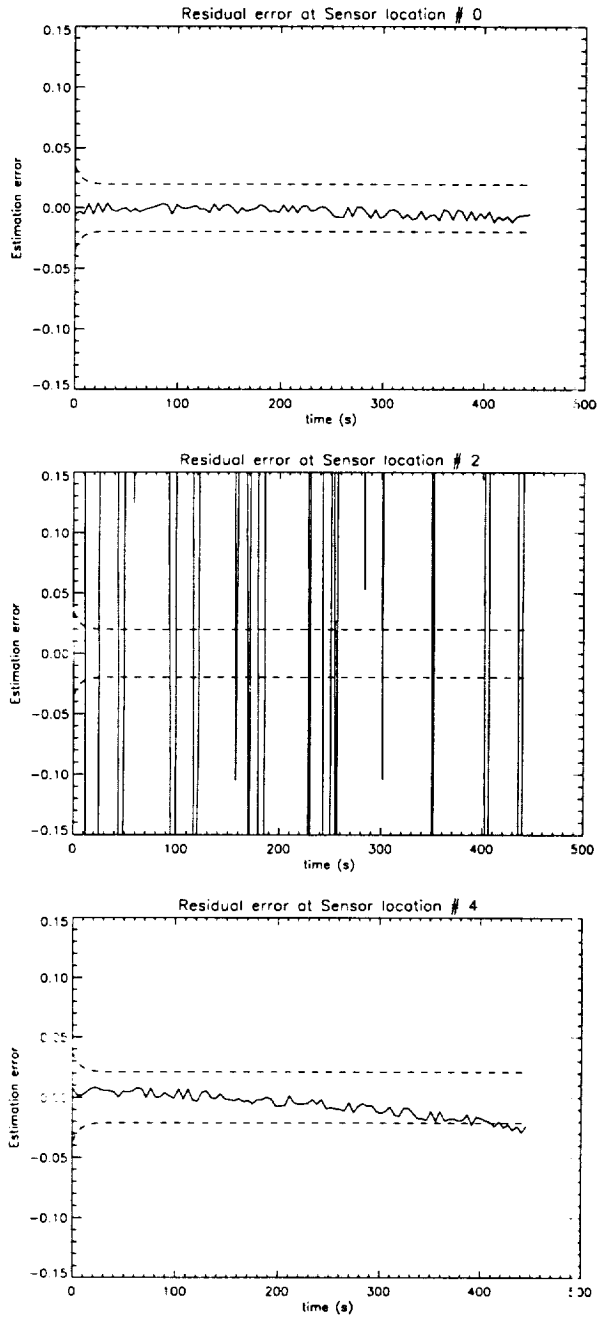

Figure 5.3: Residual error in three sample sensors with a fault in Sensor #2. Sensor #2 has residuals that are clearly outside the error bounds, indicative of a sensor fault. The fault in Sensor #2 is beginning to affect the other sensors as well.

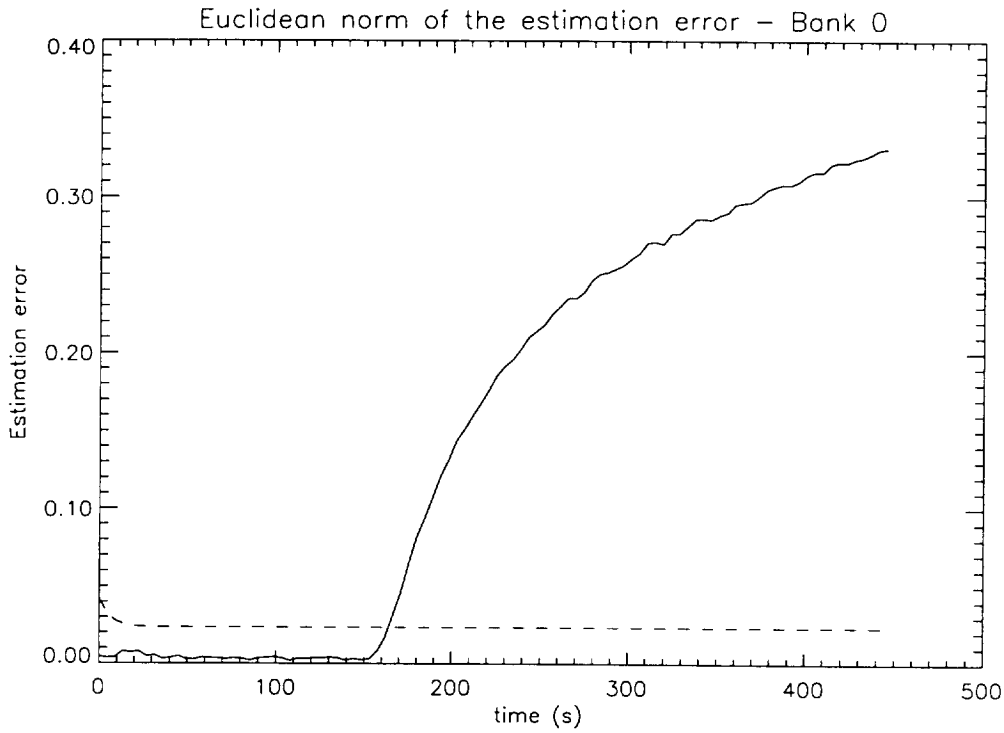

Figure 5.4: Error residual with an unknown source at t=150 s

5.4 Operation fault

An operation fault is a rather serious situation that arises when the space operation itself is in identifiable danger because the space atmosphere is seriously contaminated. One way of recognizing a fault in the system would be by studying the state of the system, which in this case would be the concentrations of each contaminant at every grid point. A fault can be posed using the norms of these data, the way it is commonly done in linear system theory. For purposes of contaminant monitoring aboard the Space Station, a fault can occur both in the l_2 and the l_1 sense. A fault in the l_1 sense occurs for the case of substances that cannot exceed a certain SMAC, and which is an acute toxic. For such substances, a fault occurs if

$$\| l_1(q) \| > q_c \tag{5.5}$$

where q_c is the appropriate SMAC for that contaminant, and l_1 and l_2 are the 1 and 2 norms of the concentration vector. A fault in the l_2 sense occurs when the process fault affects the health of the cabin in an overall sense. For example, a fault in the oxygen system occurs if

$$\| l_2(q) \| < q_c$$

and a fault in the carbon dioxide system occurs if

$$\| l_2(q) \| > q_c$$

Situations could vary from high CO_2 levels that have exceeded the long term or short term SMACS, or with toxic releases detected at levels that are known to be harmful to humans aboard. The situation could further be subdivided into either a local or a global fault, and the detection algorithm is able to distinguish among these.

5.5 Scenarios

5.5.1 CO_2 operation

We will discuss the utility of the monitoring system in two specific scenarios, both relating to the CO_2 levels in the Space Station (Narayan and Ramirez, 1998a). According to NASA standards, CO_2 must be present in the cabin at levels between 0.3 - 0.80 volume %. The highest level it can reach is 1.3 volume %, and when that level is exceeded, the mission is called off.

In the first scenario, we consider a case in which there is a leak of carbon dioxide from a carbon dioxide storage system. This means that CO_2 is constantly being added to the system and is likely to accumulate until some action is taken. Since this is unknown to the model, a good test of the filter would be to see how quickly this is detected. In this circumstance, it would be useful to monitor the levels and raise an alarm if the emergency levels are exceeded. Figure 5.5 shows how this situation leads to an emergency situation, when no mitigating actions are undertaken. The CO_2 concentration violates the SMAC at time, t = 320 s.

In the second scenario, we simulate a fire in one section of the cabin, which is extinguished by a CO_2 extinguisher. (Halon cannot be used aboard the Space Station.) The CO_2 level consequently will immediately rise in the vicinity of the fire (both due to combustion product and due to the use of the extinguisher), and we wish to monitor how the level declines, and when the cabin becomes habitable again. This sort of a simulation would be invaluable in cases where there are multiple modules, and activity can be curtailed in the module under scrutiny until the levels are safe again. The release occurs at time t = 20 s, and continues for 40 s. Figure 5.6 shows how the concentration of CO_2 changes with time for Sensor location #0. The SMAC is locally violated at t = 70 s, and the cabin is safe for habitation again at t = 200 s. The utility of the three-dimensional model lies in the fact that even local violations of safety standards can be detected, at both sensor and non-sensor

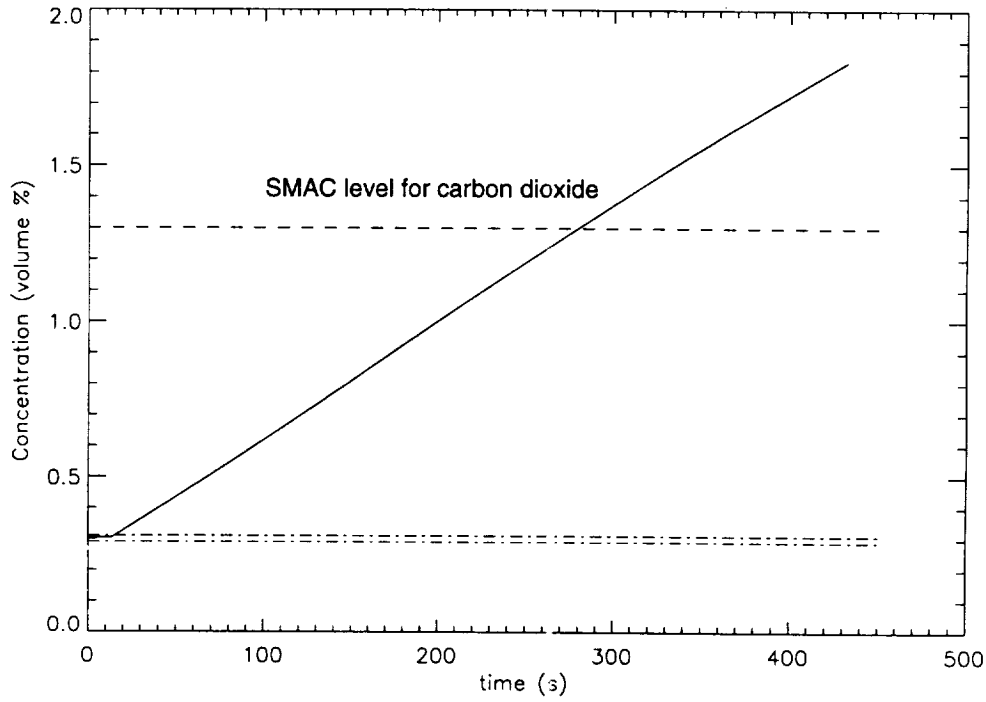

Figure 5.5: Average concentrations in the cabin in the presence of a continuous leak of CO_2. The dotted-dashed line indicates the error bounds outside which a fault is declared, and the dashed line is the SMAC level for CO_2.

locations even when the average concentration of the contaminant in the cabin lies below the SMAC.

5.5.2 Other contaminants

The monitoring of acute toxics would follow a slightly different procedure. For one, they are not normally present in the cabin atmosphere, so there are no sensor readings. Secondly, there are many specific toxics, each with different SMACS, and different sensors. Also, the tolerance for these substances will be tighter, so the procedure needs to be extra-sensitive.

For this purpose, the best procedure would be to have a backup filter ready, and initialized, which can be activated as soon as any of the sensors register a reading for the toxic. The procedure can then access a central database for the SMAC for the substance, and begin to operate on the filter measurements. Since it is likely to be an unknown source, a diagnosis will have to be performed at the very beginning itself. A sample result is shown in Fig. 5.7. Sensor readings are identically zero, and the residual and the filter both show a zero reading. A source is introduced at time = 100 s , unknown to the filter. The filter responds almost immediately, and the fault is quite apparent within 10 s of the release.

5.6 Summary and Conclusions

Fault detection algorithms have been implemented using the error residuals from the Implicit Kalman filter. The algorithm is able to detect and distinguish between sensor and process faults. The principle of Analytical redundancy using parallel banks of filters is used to detect sensor faults, while process faults are detected when the residual estimation error from the filter exceeds pre-determined bounds. The filter is able to detect faults very quickly, which would be critical during space missions.

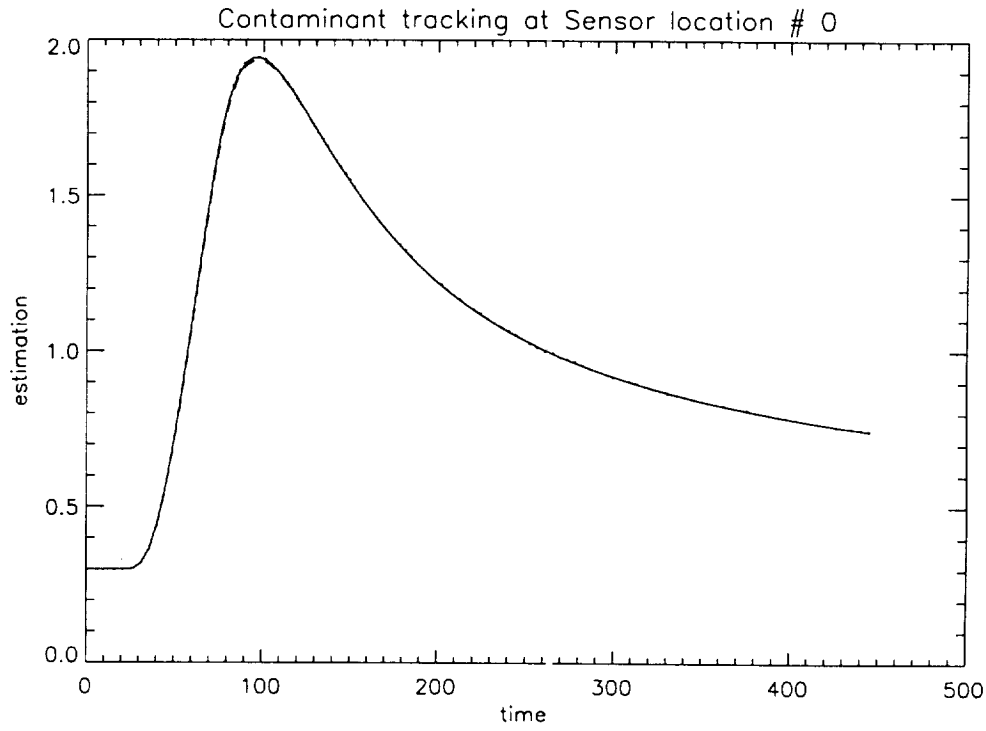

Figure 5.6: Local concentration at a sample location in the cabin. The solid line indicated the true concentration at the Sensor location, and the dotted line, almost indistinguishable from the solid line is the filtered estimate of the concentration using the Implicit Kalman Filter

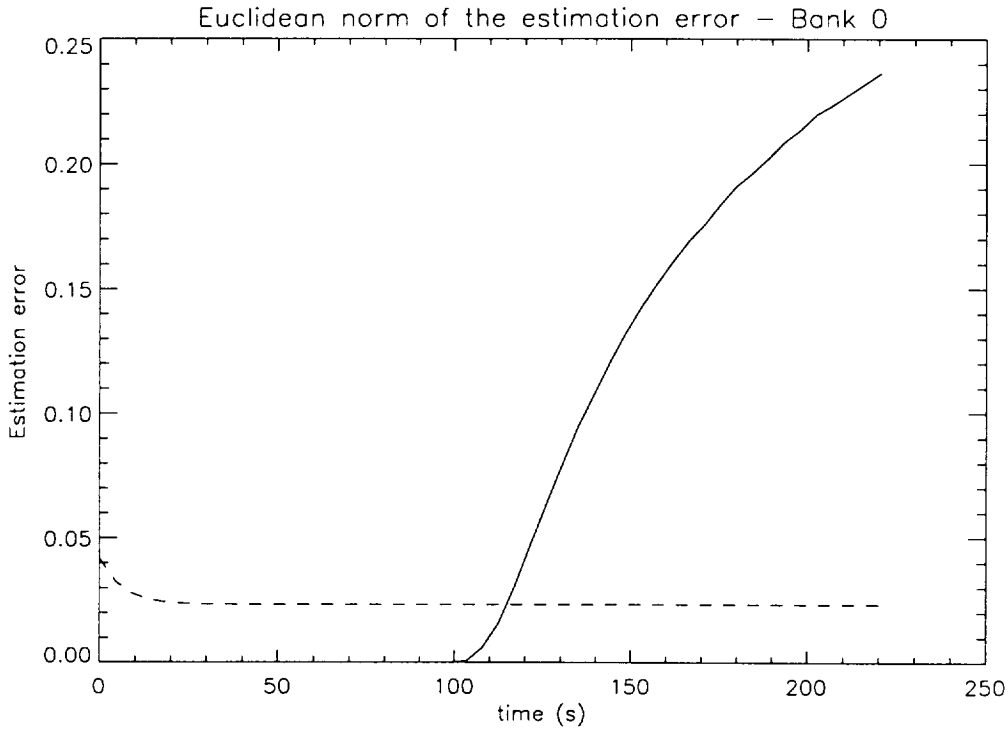

Figure 5.7: Filter response for an acute toxic release at t = 100 s. The filter responds rapidly, and clearly signals a fault with the residual error exceeding the error bound in about 15 s after the release.

Chapter 6

Source identification - solving the inverse problem

6.1 Introduction

The final portion of this work is devoted to the inverse problem – that of identifying the unknown source that is causing a fault that has been detected by the detection algorithm. A rich body of literature exists in the realm of inverse problems (Alifanov, 1994; Kurpisz and Nowak, 1995), although much of the work has remained theoretical (Kirsch, 1996) with a few practical solutions. Part of the reason for this is the relative intractability of inverse problems, beyond simple cases with restrictive assumptions.

One characteristic of an inverse problem is the unique manner in which data errors affect the error in the solution. The classic example given by Hadamard (Hadamard, 1923) was that of finding a solution u to the Laplace equation

$$\Delta u(x,y) := \frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial y^2} = 0 in \Re X[0,\infty) \tag{6.1}$$

that satisfies the conditions

$$u(x,0) = f(x), \frac{\partial}{\partial y} u(x,0) = g(x), x \in \Re, \tag{6.2}$$

where f and g are given functions. The unique solution for

$$f(x) = 0$$

and

$$g(x) = \frac{1}{n} sin(nx)$$

is given by

$$u(x,y) = \frac{1}{n^2} sin(nx) sinh(ny), x \in \Re, y \geq 0. \tag{6.3}$$

With this solution, we therefore have

$$sup\{|f(x) + g(x)|\} = \frac{1}{n} \rightarrow 0, n \rightarrow \infty, \tag{6.4}$$

but then, for the error,

$$sup|u(x,y)| = \frac{1}{n^2} sinh(ny) \rightarrow \infty, n \rightarrow \infty \tag{6.5}$$

for all $y > 0$. So, even though the error in the data tends to zero, the error in the solution u tends to infinity.

Thus, a zero error in data tends to result in an infinite error in the solution u.

6.2 Literature survey of solution methods

Skliar (Skliar, 1996) used a one-shot optimization solution to estimate the location and capacity of a source, once it was detected. While this a relatively quick operation, and computationally non-intensive, it is prone to very high errors, especially in the presence of measurement noise.

A study very similar in scope to ours was carried out (Richards et al., 1997a; Richards et al., 1997b) for application to a fire detection problem. They proposed a method for detecting, locating and sizing accidental fires in warehouses, based on the solution to an inverse heat transfer problem. They use a forward solution database, and minimize the least square error between estimated and measured times of activation of sensors that have been installed on the ceiling of the warehouses. On

closer examination, however, their problem turns out to be simpler since they assume a quiescent room with no air motion, and that the heat transfer occurs primarily through a buoyant plume of combustion gases that rises to the ceiling. Theirs is a one-shot solution mechanism.

The estimation of multiple unknown sources is further complicated, and not many solutions exist. Cheng-Hung Huang and Jan-Yue Wu (Huang and Wu, 1994) solved the two-dimensional inverse problem for two heat sources, but then, they assumed the boundary at which the sources were acting, therefore converting the problem to an inverse boundary problem.

Mass transport inverse problems are common in geology where measurements at the earth's surface are used to infer properties of processes occurring deep inside the earth. A general model has been developed (Talenti and Tonani, 1995) for gas-emitting geological systems, where the bulk gas velocity at the surface is used to locate the strength and location of the gas source. Another application for inverse problems that dealt with mass transfer was developed by Australian researchers (Newsam and Enting, 1988; Enting and Newsam, 1990; Enting, 1993), who considered the problem of estimating surface sources of carbon dioxide and other trace contaminants from surface concentration data. They used a three-dimensional diffusion model for their transport process and analytically solved the equation to account for the influence of various factors on the ability to invert measurement data to obtain source estimations.

An elegant mathematical formulation for the determination of the source term was developed (Nanda and Das, 1996) for special cases of the heat conduction equation, which however assumes a specific mathematical form of the source function.

6.2.1 Extended Implicit Kalman Filter

Once a fault is detected, the next step is the identification of the source of the fault, namely its location and capacity. The 2-D model of Skliar used a one-shot

source identification, which is susceptible to high errors, especially in a system with so much noise (Skliar, 1996). Source identification problems of this type fall into the realm of inverse problems (Alifanov, 1994). An inverse problem , simply defined, is one where the cause is discovered from a known result. They arise in electrodynamics, geophysics, astrophysics and many other fields (Kurpisz and Nowak, 1995). Inverse problems are ill-posed, which would mean that small data errors can lead to serious errors. Many fault diagnosis methods have been developed over the years, most of them applying to lumped systems. In general, they can be classified into pattern recognition (e.g. fault dictionaries), logic-based/information flow graphs (i.e., fault trees, signed directed graphs), and estimation/analytical redundancy methods. The reader is referred to one of many survey articles that address these issues (Basseville, 1988; Frank, 1990; Gertler, 1991; Korbicz et al., 1991). While many tested techniques exist for lumped systems, distributed systems prove harder to solve because of the indirect relationship between the measurements and the model variables, and due to the large size of the model matrices.

A multi step identification is proposed here, and the source identification will carried out over the time range between the time that a fault is suspected and the time that it is finally isolated. During the diagnostic process, the main filter will continue to run, but with a larger time step. At the end of the diagnostic process, the detected source term will be incorporated into the main mathematical model.

The Implicit Kalman Filter, developed in Chapter 4, can be extended in order to estimate the unknown source. Extended Kalman filters are modifications to the Kalman filter (Halme and Seikainaho, 1986; Himmelblau, 1986) which can then be used to estimate both the state and parameters of systems. This is done through augmenting the state by adding the unknown source vector, $\mathbf{f_u}$ to the state. Using a vector representation for the unknown source allows us to generalize the formulation to include unknown sources that are distributed spatially and multiple unknown sources.

The new augmented state is now

$$\mathbf{Q_m^a} = \begin{bmatrix} \mathbf{Q_m} \\ \mathbf{f_{um}} \end{bmatrix} \tag{6.6}$$

where $\mathbf{Q_m}$, the augmented state is given by Equation 4.5, and $\mathbf{f_u}$, the unknown source vector of dimension n is given by the equation

$$\mathbf{f_u} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ f_{ui} \\ 0 \\ 0 \end{bmatrix} \tag{6.7}$$

Here, f_{ui} represents the capacity of a single unknown source. The vector, $\mathbf{f_u}$ can be modified to handle a single source distributed spatially, or multiple sources.

In addition, we assume that the source term is relatively unchanging, and that it satisfies the equation

$$\frac{d\mathbf{f_u}}{dt} = 0 + \mathbf{w_{um}} \tag{6.8}$$

where $\mathbf{w_{um}}$ is the uncertainty associated with the unknown source.

Integrating and discretizing the equation, we obtain the following equation for the description of the new state.

$$\mathbf{f_{um+1}} = \mathbf{f_{um}} + \mathbf{k_m w_{um}} \tag{6.9}$$

where $\mathbf{k_m}$ is a matrix reflecting the integration time step. The noise in the unknown

source is characterized by its covariance, $\mathbf{S_m}$, which is a user-defined constant that depends on how much the unknown source capacity can vary. High values for $\mathbf{S_m}$ reflect a high uncertainty in the unknown source capacity, and will result in a filter that responds quickly to the residual errors, whereas low values for $\mathbf{S_m}$ reflect a fairly constant unknown source capacity and will result in a filter that responds slower to the residual errors, but one that will provide a smoother solution for the unknown source estimation.

The state equations, once the unknown source is included now read

$$\mathbf{A_1}^a \mathbf{Q}^a_{m+1} = \mathbf{A_2}^a \mathbf{Q}^a_m + \begin{bmatrix} \mathbf{f_m} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{C_m} & \mathbf{k_m} \end{bmatrix} \begin{bmatrix} \mathbf{w_m} \\ \mathbf{w_{um}} \end{bmatrix} \tag{6.10}$$

where

$$\mathbf{A_1}^a = \begin{bmatrix} (\frac{-\mathbf{A_x}}{2} - \frac{\mathbf{r}}{\Delta t}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{\mathbf{r}}{\Delta t} & (\frac{-\mathbf{A_y}}{2} - \frac{\mathbf{r}}{\Delta t}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\mathbf{r}}{\Delta t} & (\frac{-\mathbf{A_z}}{2} - \frac{\mathbf{r}}{\Delta t}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W} \end{bmatrix} \tag{6.11}$$

and

$$\mathbf{A_2}^a = \begin{bmatrix} \mathbf{0} & \mathbf{0} & (\mathbf{A_x}/2 + \mathbf{A_y} + \mathbf{A_z} - \mathbf{r}/\Delta t) & \mathbf{W} \\ \mathbf{0} & \mathbf{0} & -\mathbf{A_y}/2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{A_z}/2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W} \end{bmatrix} \tag{6.12}$$

$\mathbf{W}$ is an n x n diagonal matrix, where n is the number of grid points used for the

discretization, and has the structure,

$$\mathbf{W} = \begin{bmatrix} 0 & & & & & & \\ & 0 & & & & & \\ & & \ddots & & \mathbf{0} & & \\ & & & 0 & & & \\ & & & & 1 & & \\ & & \mathbf{0} & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \tag{6.13}$$

The matrix is extremely sparse, with all zeros except for source location, where there is a unit term. This unit term picks off the source capacity scalar term on expansion to yield Eq. 6.9. For multiple sources or single sources that are distributed spatially, there will be as many non-zero terms as there are sources, placed at the appropriate grid points.

The measurement matrix remains unchanged since we have no measurements that have a direct bearing on the unknown source, but the matrix has to be augmented.

$$\mathbf{z}_{m+1} = \begin{bmatrix} \mathbf{H_1}(m+1) & 0 \end{bmatrix} \mathbf{Q}^a_{m+1} + \mathbf{v}_{m+1} \tag{6.14}$$

The filter changes to account for the new state and the estimation algorithm consists of an Implicit and an explicit part, the Implicit part used in the estimation of the state, and a parallel explicit portion used in the estimation of the strength of the unknown source. The gain terms serve to multiply the residual errors generated in order to update the state estimation for the next time step.

Filter(Implicit):

$$
\begin{aligned}
(-\frac{\mathbf{A_x}}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}^* &= (\frac{\mathbf{A_x}}{2} + \mathbf{A_y} + \mathbf{A_z} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}_m + \mathbf{f_m} + \mathbf{f}_{um} + \\
&\quad \mathbf{L_1}[\mathbf{z} - \mathbf{H_1}\hat{\mathbf{y}}_{m+1|m}] \\
(-\frac{\mathbf{A_y}}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}^{**} &= -\frac{\mathbf{A_y}}{2}\mathbf{q}_m - \frac{\mathbf{r}}{\Delta t}\mathbf{q}^* + \mathbf{L_2}[\mathbf{z} - \mathbf{H_1}\hat{\mathbf{y}}_{m+1|m}] \qquad (6.15) \\
(-\frac{\mathbf{A_z}}{2} - \frac{\mathbf{r}}{\Delta t})\mathbf{q}_{m+1} &= -\frac{\mathbf{A_z}}{2}\mathbf{q}_m - \frac{\mathbf{r}}{\Delta t}\mathbf{q}^{**} + \mathbf{L_3}[\mathbf{z} - \mathbf{H_1}\hat{\mathbf{y}}_{m+1|m}]
\end{aligned}
$$

Filter(Explicit):

$$\mathbf{f_{um+1}} = \mathbf{f_{um}} + \mathbf{L_4}[\mathbf{z} - \mathbf{H_1}\hat{\mathbf{y}}_{\mathbf{m+1|m}}] \qquad (6.16)$$

The predicted estimated value for the auxiliary variable $\mathbf{y^a}$ is given by the following equation

$$\hat{\mathbf{y}}^{\mathbf{a}}_{\mathbf{m+1|m}} = \begin{bmatrix} \mathbf{A_2} \\ \mathbf{W} \end{bmatrix} \hat{\mathbf{q}}^{a}_{m|m} + \begin{bmatrix} \mathbf{f}_m + \mathbf{f}_{um} \\ \mathbf{0} \end{bmatrix} \qquad (6.17)$$

where $\mathbf{y^a}$ is the new augmented estimate and is a column vector of size 4n.

The augmented Implicit Kalman Gain matrix is a 4n by m matrix, with four partitions (one for each co-ordinate direction and one for the unknown source) and has the structure

$$\mathbf{L_{m+1}}^{a} = \begin{bmatrix} \mathbf{L_1^T} & \mathbf{L_2^T} & \mathbf{L_3^T} & \mathbf{L_4^T} \end{bmatrix}^T \qquad (6.18)$$

and is obtained from the equation

$$\mathbf{L^a_{m+1}} = \mathbf{P^y_{m+1|m}}\,^{\mathbf{a}}\mathbf{H_1^{T^a}}[\mathbf{H_1}\,^{\mathbf{a}}\mathbf{P^y_{m+1|m}}\mathbf{H_1^{T^a}} + \mathbf{R(m+1)}]^{-1} \qquad (6.19)$$

where

$$\mathbf{P}^{y}_{m+1|m}{}^{a} = \tag{6.20}$$

$$\begin{bmatrix} \mathbf{A}_2^{13}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{13T} + & \mathbf{A}_2^{13}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{23T} & \mathbf{A}_2^{13}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{33T} & \mathbf{P}^{u}_{m|m} + \mathbf{CQS}^{T} \\ +\mathbf{P}^{u}_{m|m} + \mathbf{CQC}^{T} & & & \\ \mathbf{A}_2^{23}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{13T} & \mathbf{A}_2^{23}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{23T} & \mathbf{A}_2^{23}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{33T} & \mathbf{0} \\ \mathbf{A}_2^{33}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{13T} & \mathbf{A}_2^{33}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{23T} & \mathbf{A}_2^{33}\mathbf{P}^{q}_{m|m}\mathbf{A}_2^{33T} & \mathbf{0} \\ \mathbf{P}^{u}_{m|m} + \mathbf{SQC}^{T} & \mathbf{0} & \mathbf{0} & \mathbf{P}^{u}_{m|m} + \mathbf{SQS}^{T} \end{bmatrix}$$

$\mathbf{P}^{y}_{m+1|m}{}^{a}$ is the augmented predicted error covariance matrix, and is of dimension 4n by 4n. $\mathbf{P}^{u}_{m|m}$ is the uncertainty matrix associated with the unknown source, and in general is an n by n diagonal matrix, with the uncertainty due to each unknown source location as the diagonal elements, though it reduces to a sparse matrix with just one diagonal element, for a fault caused by an emission at one spatial location. $\mathbf{P}^{q}_{m|m}$ is the covariance of the estimated error, and defined by Eq. 4.21. $\mathbf{C}$ and $\mathbf{Q}$ represent the stochastic model disturbance transition matrix and the model covariance matrix, respectively. The superscripts in the $\mathbf{A_2}$ matrix refer to the appropriate partitions of the $\mathbf{A_2}^{a}$ matrix defined earlier.

The gain matrix partitions, $\mathbf{L_2}$ and $\mathbf{L_3}$ remain unchanged from the simple implicit filter, and only the $\mathbf{L_1}$ partition changes, along with the introduction of the new gain partition for the unknown source, $\mathbf{L_4}$. The gain matrix partition, $\mathbf{L_1}$ now becomes a function of the uncertainty term of the unknown source, $\mathbf{P}^{u}_{m|m}$, which is then used in the correction term for the first Implicit Kalman Filter equation in Eq. 4.13.

Example:

For a system with six sensors and one unknown source, which is the most commonly expected fault, the gain matrix partition for the unknown source term has the structure

$$\mathbf{L_4} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \mathbf{l}_{4(i)} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{6.21}$$

where i refers to the location of the unknown source in the cabin. $\mathbf{l}_{4(i)}$ is a row vector of dimension m, and is evaluated by simplifying Eq. 6.19 for one unknown source. In general, $\mathbf{L_4}$ has as many non-zero rows as there are unknown sources.

$$\mathbf{l}_{4(i)} =$$

$$(p^u_{m|m} + sqc) \begin{bmatrix} h_{11(1)} & h_{11(2)} & h_{11(3)} & h_{11(4)} & h_{11(5)} & h_{11(6)} \end{bmatrix} \times$$

$$\left\{ \begin{bmatrix} h_{11(1)} \\ h_{11(2)} \\ h_{11(3)} \\ h_{11(4)} \\ h_{11(5)} \\ h_{11(6)} \end{bmatrix} (p^u_{m|m} + sqc) \begin{bmatrix} h_{11(1)} & h_{11(2)} & h_{11(3)} & h_{11(4)} & h_{11(5)} & h_{11(6)} \end{bmatrix} + \right. \tag{6.22}$$

$$\left. \begin{bmatrix} R_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & R_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & R_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & R_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & R_6 \end{bmatrix} \right\}^{-1}$$

The lower case symbols, p, s, q, and c represent the scalar terms for the uncertainty

associated with the unknown source, the transformation factor representing the integration term in the unknown source terms, and the model uncertainty for that grid location.

The measurement term for the grid point where the unknown source is located for sensor $s.$ is $h_{11(s)}$. The computation therefore yields $\mathbf{l}_{4(i)}$, which is a row vector of dimension 6, since we have six sensors in this case. Each of these six gain terms multiplies the appropriate sensor residual from $\mathbf{z} - \mathbf{H_1 y}$, which is then used to update our estimate for the unknown source location via Eq. 6.16.

The gain term, in turn is used to propagate the uncertainty for the next time step, which is governed by the equation

$$\mathbf{P}^{ya}{}_{m+1|m+1} = [\mathbf{I} - \mathbf{L}^{a}_{m+1}\mathbf{H}^{a}_{1}]\mathbf{P}^{ya}{}_{m+1|m} \tag{6.23}$$

and

$$\mathbf{P}^{ya}{}_{m+1|m+1} = \mathbf{A}^{a}_{1}\mathbf{P}^{Q}{}_{m+1|m+1}\mathbf{A_1}^{Ta}, \tag{6.24}$$

6.2.2 Initial guesses through sensitivity matrices

The Extended Implicit filter, like all Kalman filters requires an initial guess for $\mathbf{f}_{um}$, which is not a trivial problem. While the filter performs well in filtering out noise, and adjusting for model errors, its performance in estimating an unknown source depends crucially on a good initial guess for the location of the unknown source emission.

For the purpose of this derivation, we assume that there is only a single localized source term that is causing the fault. The solution can be extended to multiple and distributed sources, but it would be more complicated since the single source will be broken down into a combination of linearly independent single sources.

In this work, we use pre-calculated sensitivity coefficients for the purpose.

Sensitivity coefficients have been widely used in estimating solutions to inverse problems, especially in areas of heat conduction, geology, and tomography (Alifanov, 1994). Some detailed analysis of sensitivity coefficients and their properties and use are also available (Beck et al., 1985; Kurpisz and Nowak, 1995). The sensitivity coefficient can be defined as

$$\mathbf{Z} = \frac{\partial \mathbf{q}}{\partial \mathbf{f_u}} \tag{6.25}$$

where $\mathbf{f_u}$ represents the source term. The sensitivity coefficient term represents the sensitivity of the concentration $\mathbf{q}$ at each mesh point with respect to the release of a source at every mesh point. $\mathbf{Z}$ is calculated by solving the basic model equation, Eq. (3.1). While $\mathbf{Z}$ is defined over the entire domain, the only points of interest will be the sensor locations since those are only points about which we have direct information about the concentration that we can use in the event of a fault. $\mathbf{Z}$ is therefore partitioned into a usable and non-usable part,

$$\mathbf{Z} = \left[\begin{array}{cc} \mathbf{Z}_{sensors} & \mathbf{Z}_{non-sensors} \end{array} \right] \tag{6.26}$$

This pre-calculated sensitivity matrix can be computed for different times, in order to provide a window of time over which the fault diagnosis can be conducted, depending on how soon the fault is detected.

The first step, therefore, is to calculate $\mathbf{Z}$. In order to do this , we multiply Eq. 3.1 throughout by $\frac{\partial}{\partial \mathbf{f_u}}$, which leads to the equation

$$\frac{\partial}{\partial \mathbf{f_u}} \left[\frac{\partial \mathbf{q}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{q} = D_M \nabla^2 \mathbf{q} + \mathbf{f_u} \right] \tag{6.27}$$

Since $\mathbf{f_u}$ is independent of the co-ordinate axes, we can rewrite Eq. 6.27 as

$$\frac{\partial}{\partial t} \frac{\partial \mathbf{q}}{\partial \mathbf{f_u}} + \mathbf{u} \cdot \nabla \left(\frac{\partial \mathbf{q}}{\partial \mathbf{f_u}} \right) = D_M \nabla^2 \left(\frac{\partial \mathbf{q}}{\partial \mathbf{f_u}} \right) + \frac{\partial \mathbf{f_u}}{\partial \mathbf{f_u}} \tag{6.28}$$

Replacing $\mathbf{Z}$ for $\frac{\partial \mathbf{q}}{\partial \mathbf{f_u}}$ from the definition for the sensitivity coefficient, we

get the direct well-posed problem equation

$$\frac{\partial \mathbf{Z}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{Z} = D_M \nabla^2 \mathbf{Z} + \mathbf{I} \tag{6.29}$$

subject to the appropriate boundary and initial conditions. The identity matrix, I essentially is a unit source that is sequentially placed at every grid location in the cabin in order to measure its effect on all the other grid locations. The boundary conditions and the initial conditions will also have to be divided throughout by $\frac{\partial}{\partial \mathrm{fu}}$ to obtain the initial and boundary conditions for the sensitivity problem. The **Z** matrix is then partitioned to obtain $\mathbf{Z}_{sensors}$. Since the structure of the equation is unchanged from the original model equation, the same algorithms and numerical techniques can be used in computing the sensitivity matrix. For each time step, the computation of the sensitivity matrix takes about 180 CPU minutes on the DEC-Alpha 500 AU, and therefore, the computations have to be performed off- line and before the time that the diagnosis can take place.

The solution to Eq. 6.29, $\mathbf{Z}(t)$, is a function of time, where the time refers to the time elapsed since the fault occurred. For the purpose of using the $\mathbf{Z}_{sensors}$ partition in calculating an initial guess for the capacity and location of an unknown source, we use

$$\mathbf{Z}_{critical} = \mathbf{Z}_{sensors}|_{t=t_{detect}} \tag{6.30}$$

where t_{detect} refers to the time when the fault is detected. This yields a $\mathbf{Z}_{critical}$ matrix of dimensions n x m, where n is the number of grid points, and m is the number of sensors. In other words, the $\mathbf{Z}_{critical}$ matrix contains the response observed at each sensor location for a unit perturbation at every location in the cabin. In the event of a fault, this response is used in conjunction with the observed measurement response in order to estimate the perturbation that caused the fault.

The solution proposed here consists of two parts; a first off-line part that in-

volved pre-calculating sensitivity matrices, and which is a computationally intensive process; and a final real-time computationally non-intensive portion that actually computes a first guess for the location and the capacity of the unknown source emission.

The strategy of the method is as follows:

- The ill-posed problem is made well-posed by making assumptions about the problem in the areas of ill-posedness, in this case the unknown source term
- The well-posed direct problem is solved for this assumed value(s).
- The measured quantities are noted for the ill-posed problem using the sensor system
- The calculated values for the assumed problem are compared with the measured values from the sensor system, and the assumed input data are modified to ensure a matching of these two quantities. Although this method is likely to yield a correct solution, the nature of an ill-posed problem could mean that there are multiple solutions, with the method being able to identify only one of them.

Once the sensitivity matrix has been calculated, we now demonstrate how it is used in the fault diagnosis process.

In the event of a fault being diagnosed, an error m-tuple is generated using the prediction errors from the Implicit Kalman Filter.

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{m-1} \\ e_m \end{bmatrix} \tag{6.31}$$

where e_1, e_2, e_{m-1} and e_m represent the prediction errors for Sensors 1,2,m-1 and m, and represent the error between the projected estimate and the measurement signal,

$$e_i = z_i - \hat{y}_{im+1|m} \tag{6.32}$$

Next, the capacity of the unknown source is calculated, sequentially assuming that the emission occurred at each location in the cabin,

$$\mathbf{Cap}_i = \frac{\mathbf{Z_i}}{e_i} \tag{6.33}$$

This calculation results in n projected capacity m-tuples, with the assumption that the source is in each of n locations, where n refers to the number of grid points being used. The algorithm being used for this estimation (Fig. 6.1) is given below.

(1) Calculate capacities using Eq. 6.33.

(2) Scale capacities within each m-tuple using the maximum and minimum capacities within the m-tuple.

(3) Compute the standard deviation of the calculated capacities for each m-tuple, for both of the scaled versions.

(4) Pick the point with the least standard deviation. If the same point is picked using both the minimum and maximum scalings, that point is likely to be the source location.

(5) If minimum and maximum scalings yield different locations, use the location which yields the lower standard deviation as the starting guess for the location of the unknown source.

(6) The initial guess for the capacity is calculated from Step 1.

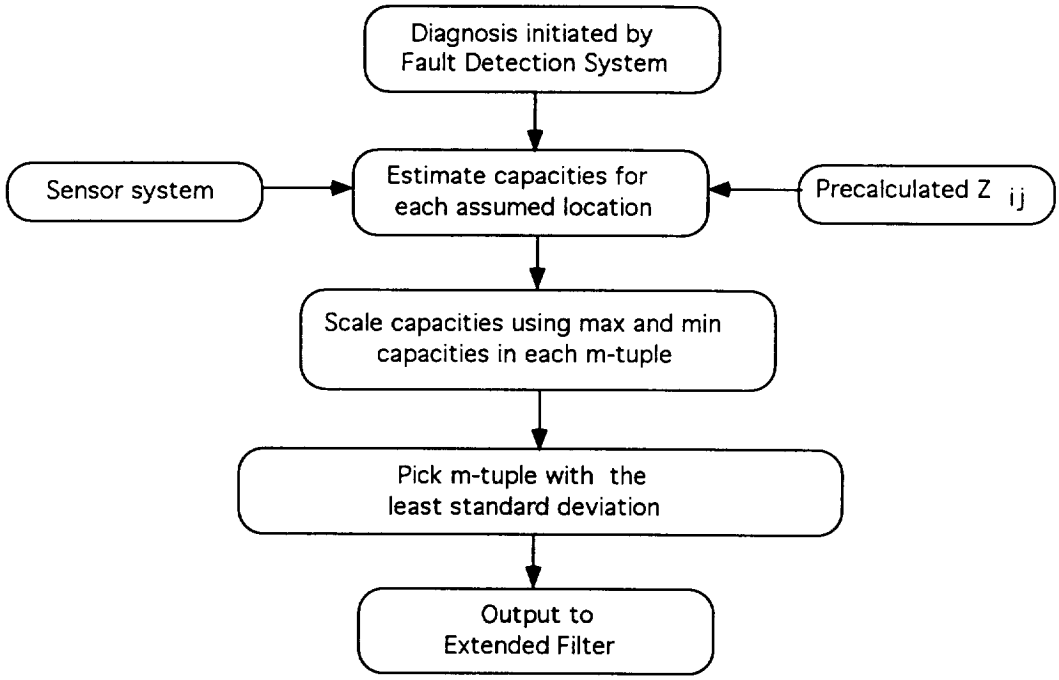

Figure 6.1: Sensitivity analysis algorithm that yields a guess for the location and capacity of an unknown source

Table 6.1: Measurements after 1 time step

Source strength	Measurement
0	0.3000001
1	0.3005528
50	0.3276336
100	0.3552671

The final estimate is based on the principle that for the correct assumption of the source, the m-tuple of the projected capacities would show the minimum standard deviation, and would match the previously generated m-tuple of sensitivity coefficients for the location, to a multiplicative real constant.

This is an area of work that is amenable to treatment using Artificial Intelligence or Knowledge Based systems, where the algorithm is trained to pick up patterns and suggest intelligent solutions based on learned past experience.

6.2.3 Sensitivity experiments

Tables 6.1 through 6.5 are the concentration measurements at Sensor location #6 (8,14,18) for a unit source release at location (8,15,12). This location happens to be at a sensor location closest to this source, and consequently has a high response. The data are shown to indicate how the sensitivity indices helps scaling.

The experiment was repeated for a sensor location farther from the source emission site, location #3 at (3,5,7). As expected, the sensitivity coefficients are smaller in magnitude. Tables 6.6 through 6.10 show the concentration measurements

Table 6.2: Measurement at Sensor location #6 after 2 time steps

Source strength	Measurement
0	0.3000004
1	0.3013975
50	0.3698579
51	0.3712551
100	0.4397156

Table 6.3: Measurement at Sensor location #6 after 10 time steps

Source strength	Measurement
0	0.3000708
1	0.3147669
50	1.034878
51	1.049574
100	1.769685

Table 6.4: Measurement at Sensor location #6 after 100 time steps

Source strength	Measurement
0	0.2533103
1	0.2858888
50	1.882236
51	1.914815
100	3.511162

Table 6.5: Measurement at Sensor location #6 after 300 time steps

Source strength	Measurement
0	0.121
1	0.1617357
50	2.115451
51	2.155323
100	4.109037
1000	39.99361

Table 6.6: Measurements at #3 after 1 time step

Source strength	Measurement
0	0.2999996
1	0.2999996
50	0.2999996
51	0.2999996
100	0.2999996

at (3,5,7) for different times.

One interesting result is that the source is not detectable by this sensor during the first two time steps, and even at 10 times steps, the effect is very slight.

The distance between the source emission location and the sensor location is efficiently captured in the sensitivity coefficient, which is computed earlier. The use of the sensitivity matrix in estimating the detectability of a fault is discussed later.

6.2.4 Sensitivity analysis-Results

In a real setting, the sensitivity analysis would be tested by introducing a fault, taking sensor readings, and checking to see if the sensitivity analysis will estimate the source correctly from the readings. In the absence of an experimental setting, we introduce the unknown fault, run the model, and use the model calculated concentrations at sensor locations as sensor readings. In order to simulate a real physical setting, we also add a Gaussian noise to the sensor readings to generate pseudo-experimental values for testing the Sensitivity analysis.

For our test case, we first pre-calculate the sensitivity matrix for the sensor

Table 6.7: Measurement at Sensor location #3 after 2 time steps

Source strength	Measurement
0	0.2999778
1	0.2999778
50	0.2999778
51	0.2999778
100	0.2999778

Table 6.8: Measurement at Sensor location #3 after 10 time steps

Source strength	Measurement
0	0.2826567
1	0.2826571
50	0.2826755
51	0.2826759
100	0.2826943

Table 6.9: Measurement at Sensor location #3 after 100 time steps

Source strength	Measurement
0	0.1181842
1	0.1196233
50	0.1901395
51	0.1915785
100	0.2620947

Table 6.10: Measurement at Sensor location #3 after 300 time steps

Source strength	Measurement
0	0.044009563
1	0.048030302
50	0.2450462
51	0.2490669
100	0.4460829

Table 6.11: Sensor location and the associated measurement noise

Sensor	Co-ordinates	R
1	10 10 10	0.0001
2	8 2 23	0.001
3	3 5 7	0.0001
4	6 6 6	0.0023
5	4 11 22	0.019
6	8 14 18	0.0019

array shown in Table 6.11. The base-line concentrations are taken to be the steady state concentration field for an initial CO_2 concentration of 0.6 volume %.

A source of strength 500 mg/m^3 is then introduced into the model, acting at location (8,6,5), and the pseudo-measurements that are generated from the model are used for the sensitivity analysis. A pseudo-random vector with six elements with a bound of 5 % of the measurement was generated and added to each measurement in order to mimic a sensor system with 5 % measurement noise.

The sensitivity analysis reports location (8,6,6) as the one with the least standard deviation (σ = 23.67), and the capacities as calculated by each of the sensors are reported in Table 6.12, the mean of which will serve as our initial guess for the Extended Implicit Kalman Filter.

While the sensitivity analysis' initial estimate is quite satisfactory for the purpose of the filter, we also note that the next highest standard deviation (σ = 200.86) was reported for the correct location (8,6,5), whose guess capacities are reported in Table 6.13. One observes that Sensors 1-4 have excellent estimates for the capacity, but the correct location is not reported by the sensitivity analysis since Sensors 5 and 6 report capacities that increase the standard deviation of that particular m-tuple.

Eq. 6.29 yields $\mathbf{Z}_{sensors}$ that can be evaluated for different times. This can provide a window of time over which the fault diagnosis can be conducted, depending

Table 6.12: Estimated capacities for assumed source at (8,6,6)

Sensor	Capacities (mg/m^3)
1	158.2
2	140.7
3	154.7
4	158.6
5	103.8
6	115.3

Table 6.13: Estimated capacities for assumed source at (8,6,5)

Sensor	Capacities (mg/m^3)
1	501.9
2	512.5
3	512.5
4	512.5
5	105.5
6	137.4

on how soon the fault is detected. In our examples, we have only used $\mathbf{Z}_{critical}$, and this could cause large errors if the fault were not accurately detected. For cases where one can expect a substantial delay in the detection of a fault, it would be prudent to minimize the standard deviation of the capacities, evaluated over a period of time, and using say, $\mathbf{Z}_{t=1}$, $\mathbf{Z}_{t=10}$, and $\mathbf{Z}_{t=50}$, which would provide added robustness to the analysis.

6.2.5 Iteration procedure for Extended Implicit Kalman filter

As we observed in the previous section, the sensitivity analysis can sometimes pick a point in the vicinity of the actual source, instead of the actual point itself. The presence of turbulence in the cabin air, and general measurement noise can increase the distance of the actual solution from the first guess. In order to refine our solution, we use the property of the Implicit Kalman Filter being an optimal estimator and run the filter for the initial guess and determine the squared prediction error. The final solution is reached when the point under consideration has the least squared error when compared with its six nearest neighbors.

The algorithm for this location search is shown in Fig. 6.2 and has the following steps.

(1) Choose one of B's spatial neighboring points (grope), which has not been previously visited as the assumed location for the unknown source.

(2) For this location, and the original estimate for the capacity, run the Extended Filter and obtain a new squared prediction error, and an estimated capacity for that location.

(3) Repeat Steps 1 and 2 until all neighbors not previously considered have been covered.

(4) Compare the Squared prediction error, and for the next approximation, choose the point with the least SPE.

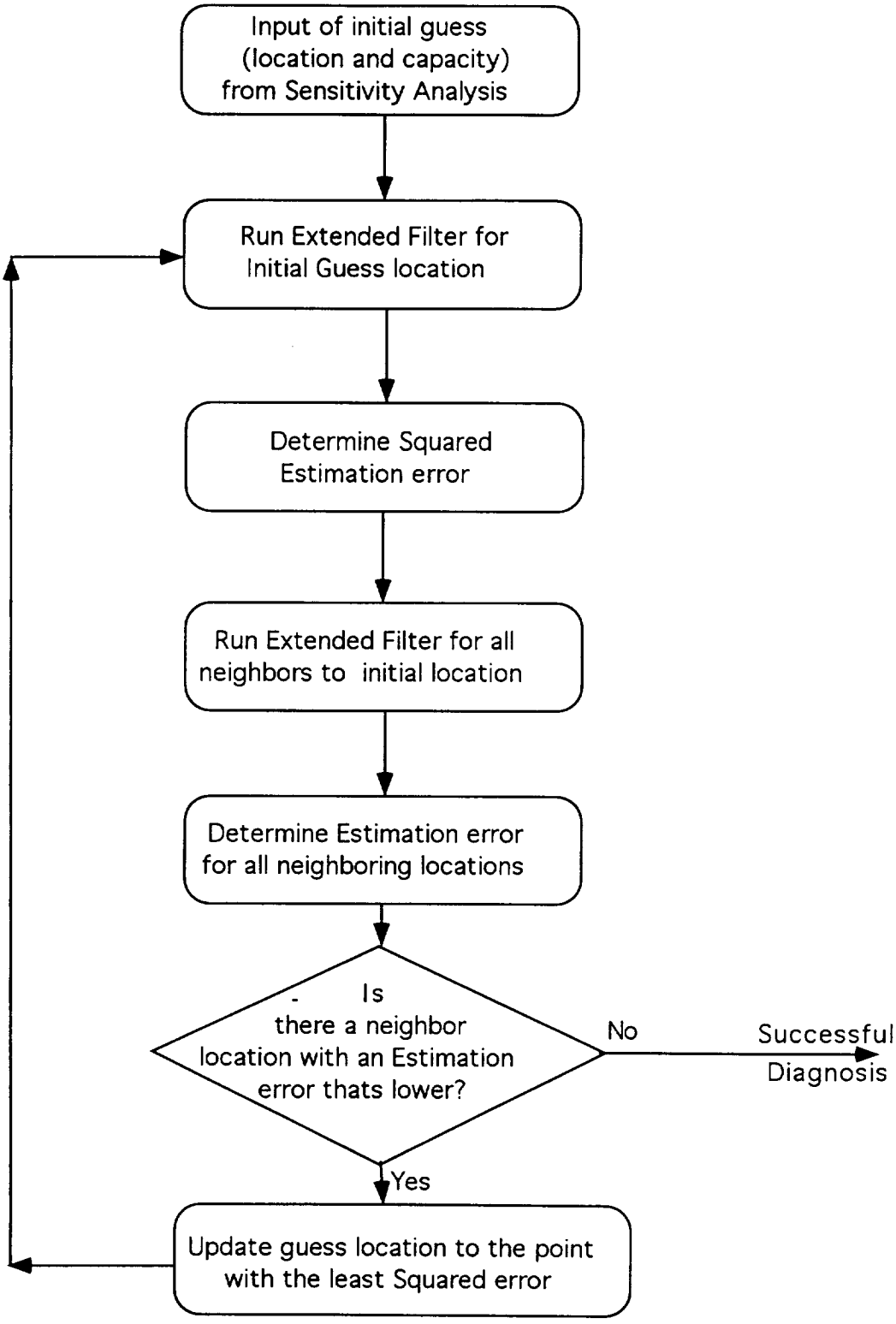

Figure 6.2: Extended Kalman Filter procedure for estimating unknown source

Table 6.14: SPE for neighbor points for a converged solution

Co-ordinates of neighbor points	SPE
7,6,9	64.557
7,5,9	293.19
7,7,9	597.72
6,6,9	9761.09
8,6,9	40655.4
7,6,8	1581.80
7,6,10	3070.07

(5) Repeat Steps 1 through 4 until a minimum is obtained.

6.3 Results

In the first test case, a source of strength 500 mg/m^3 is introduced at (7,6,9), with an initial guess of 400 mg/m^3 for the capacity at location (7,6,9) from the sensitivity analysis. The Extended Filter converges rapidly to the final solution (Fig. 6.4), and the SPE is only 64.557. While the solution appears to be acceptable on inspection, our algorithm requires that the predicted errors be examined for its neighbors. Accordingly, we run the filter for all the neighbors to (7,6,9), and confirm from Table 6.14 that the estimate is indeed the best, given the measurement data.

In the next test case, a source of strength 1500 mg/m^3 was applied at location (12,11,9), and the measurements generated were used as inputs to the sensitivity analysis algorithm. For this test case, the sensitivity analysis points to a source location at (12,10,8), and a capacity of 1150 mg/m^3. The filter does not converge for this location, and so we try the neighbors.

Based on the squared error results in Table 6.15, we shift our focus to (12,10,9) since it has least error and rerun the filter for its neighbors.

In two iterations, we are able to pinpoint the correct location and source capacity in the presence of noise. Figure 6.8 shows how the squared prediction error converges in three iterations to its final minimum value.

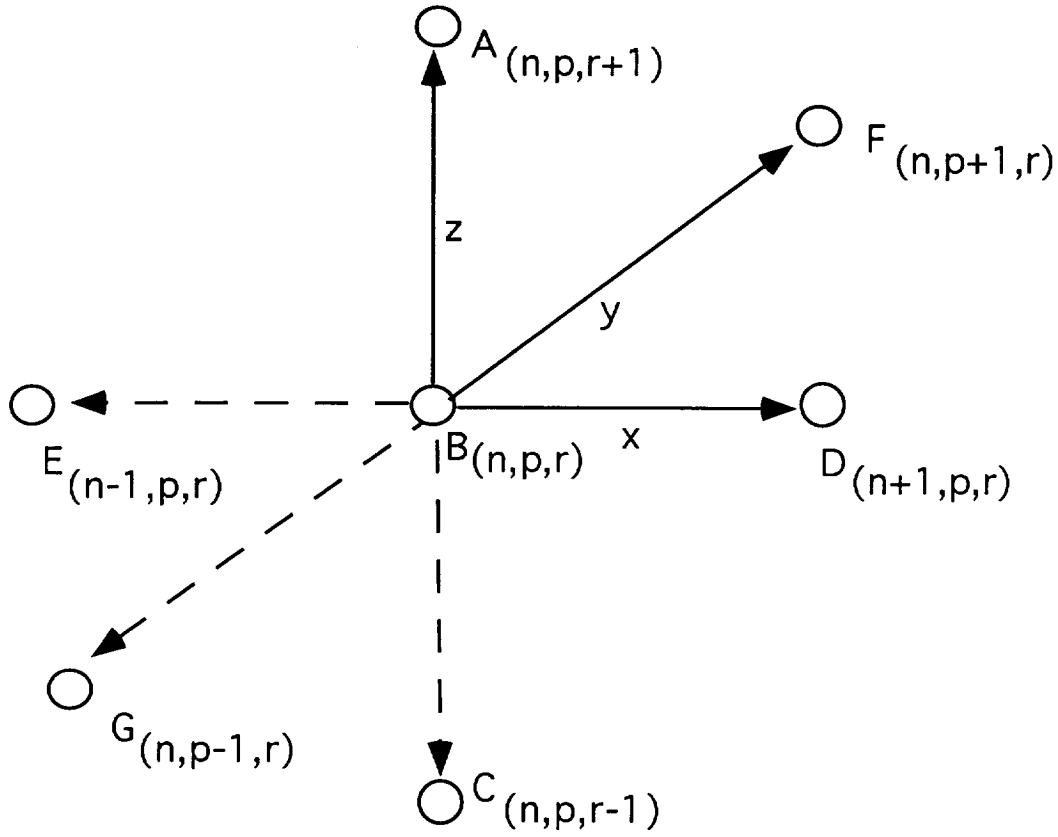

Figure 6.3: Comparing the squared estimation error across nearest-neighbors

Table 6.15: SPE for neighbor points for a converged solution-iteration 1

Co-ordinates of neighbor points	SPE
13 10 8	10545.7
11 10 8	7719.28
12 11 8	1.384 X 10^6
12 9 8	554869.0
12 10 7	552.204
12 10 9	157.859

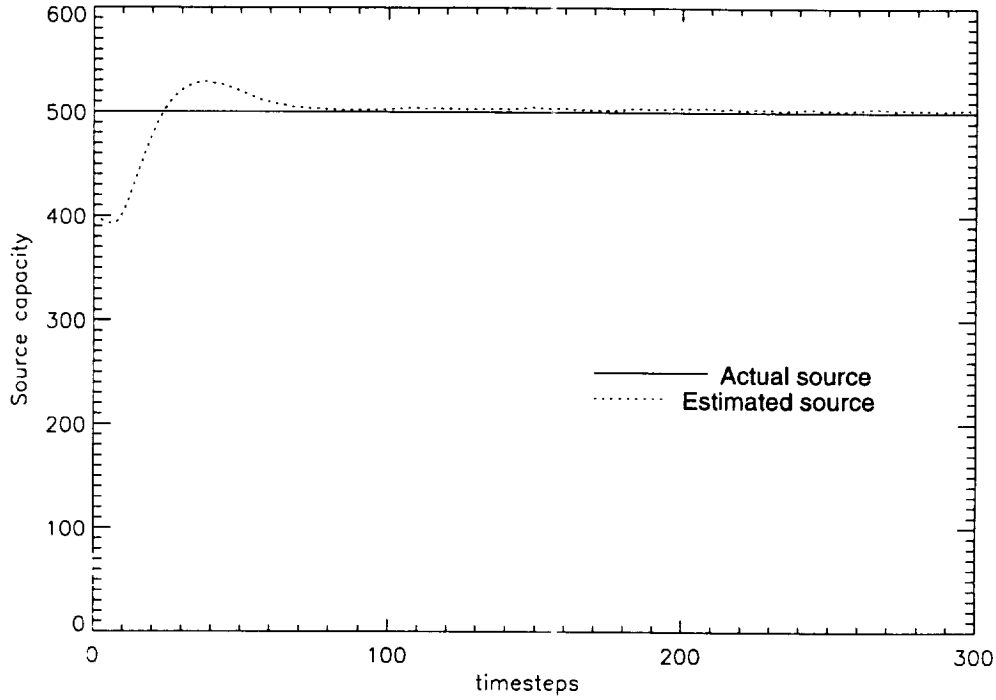

Figure 6.4: The estimation of a constant source (unknown to the model) of 500 mg/m^3 by the Extended Implicit Kalman Filte·, with an initial guess of 400 mg/m^3.

Table 6.16: SPE for neighbor points for ɛ converged solution-iteration 2

Co-ordinates of neighbor points	sum-squared error
12 11 9	72.6206
12 9 9	429.040
12 10 8	233.676
12 10 10	195.053
11 10 9	10458.9
13 10 9	10158.0

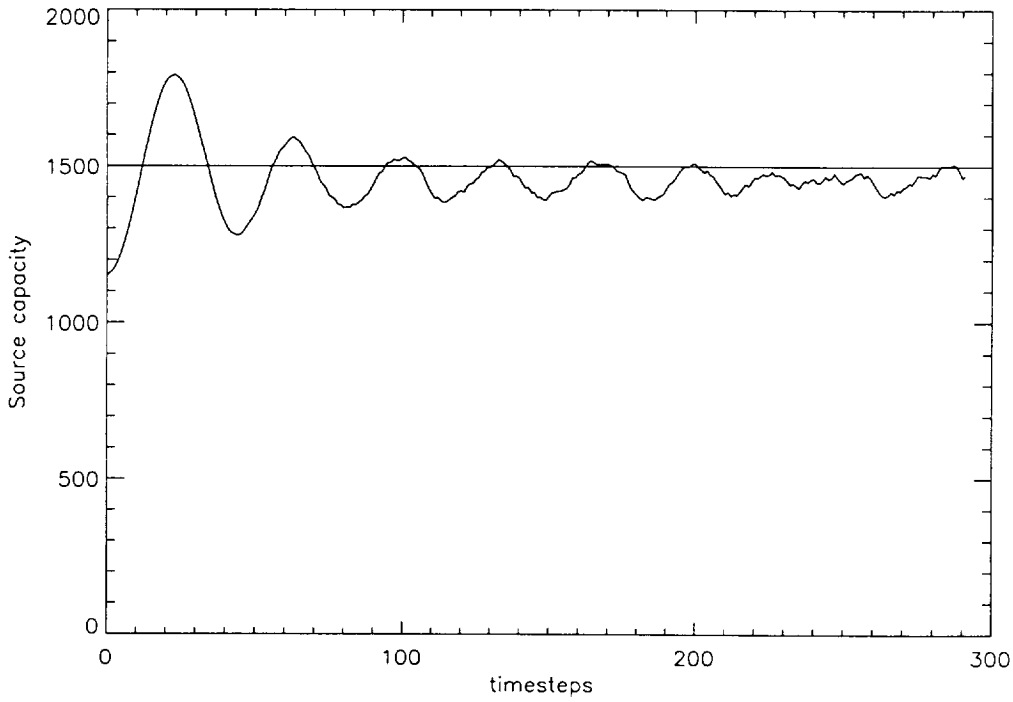

Figure 6.5: Unknown source capacity estimation – The first of a series of three iterations needed for the Extended IKF to converge to its final solution. Note that because of the wrong guess for the unknown source, the filter (wavy line) under-predicts the actual strength of the source (straight line).

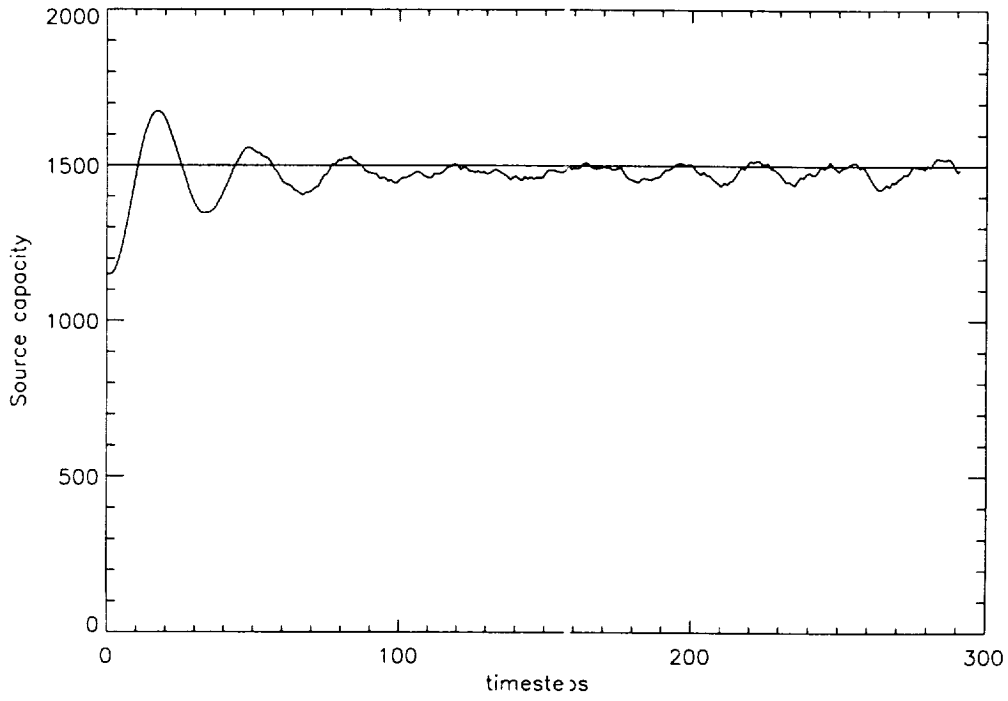

Figure 6.6: Unknown source capacity estimation– The second in a series of three iterations needed for the Extended IKF to converge to its final solution. Note that the estimation is better in this case than it is for Iteration 1, but the filter (wavy line) still slightly under-predicts the actual strength of the source (straight line).

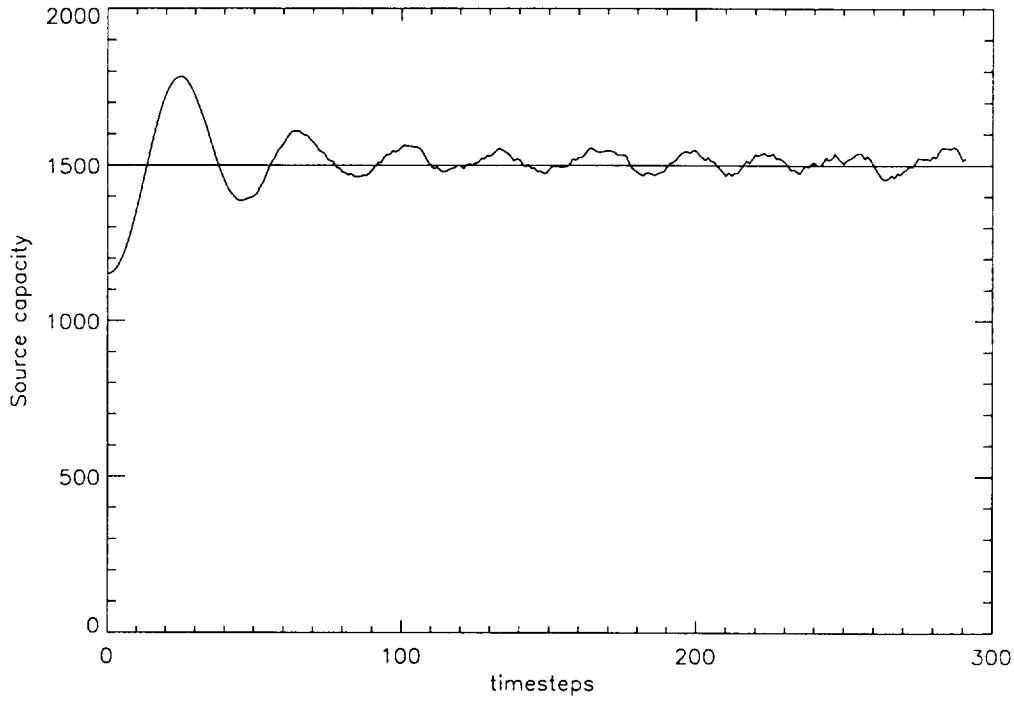

Figure 6.7: Unknown source capacity estimation– The third and final Extended IKF iteration results showing the estimated (wavy line) and correct unknown source (straight line) capacities for the correct guess for the unknown source location.

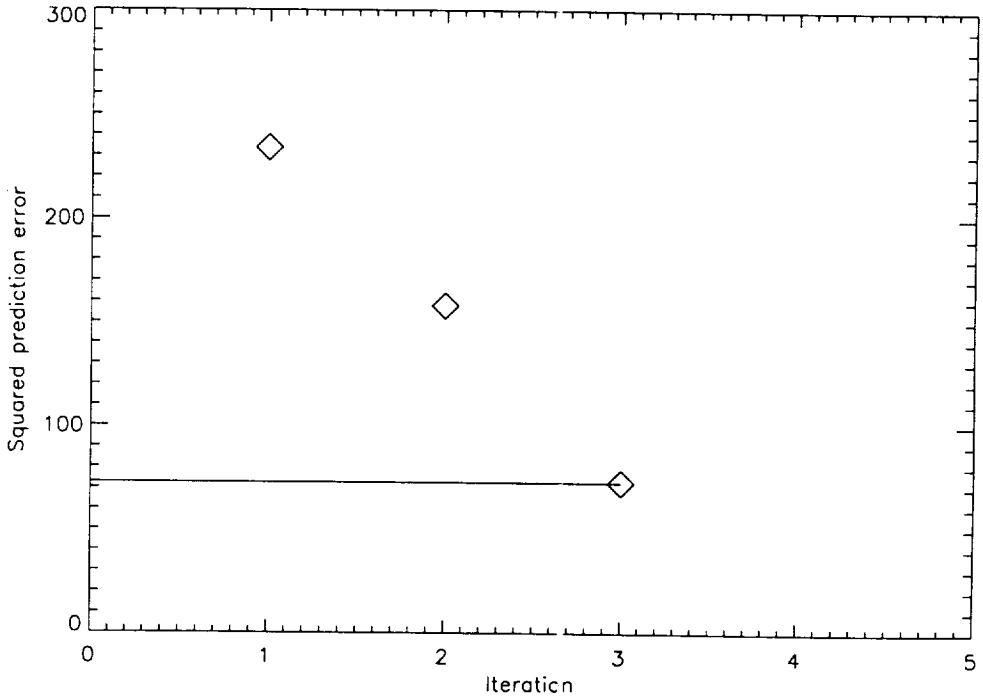

Figure 6.8: The squared prediction error varying with successive iterations until it converges to its minimum value

If for the same case, i.e., for a source of strength 1500 mg/m^3 introduced at (12,11,9), an initial guess of 1400 mg/m^3 for the capacity from the sensitivity analysis is used, the Extended Filter converges rapidly to the final solution (Fig. 6.9), and the SPE is only 22.95. The quality of the initial guess, therefore, is very crucial to the performance of the filter in terms of a quick convergence to the final solution.

The Extended Filter is only as good as the model and the sensor system are. If all the sensors are crowded in one area of the cabin, faults in other regions can go undetected.

As an example, consider a fault at location (13,10,20). A constant source of strength 500 mg/m^3 is applied over a time of 300 time steps, and the sensitivity analysis was used to estimate the location and capacity for the unknown source. The guess location was a significant distance away from the correct location. The filter was then run on these measurements, assuming the correct location as the guess location with an initial guess of 400 mg/m^3. Fig. 6.10 shows the tracking for this case, and it indicates that the response of the filter has been too slow, and even though the filter is starting to approach the final solution, the solution does not converge within the desired duration of time. It is therefore crucial to be aware of 'dead-zones' which are outside the domain of observability and consequently, detectability.

6.3.1 Varying functions

The Extended Filter can be applied even to varying sources. If the sources are varying very quickly, then our assumption is violated, and therefore the solution is likely to be oscillatory and posses a high residual error. Figure 6.11 is an example of a varying source that was tracked by the filter.The filter exhibits oscillatory behaviour because the unknown source function is changing rapidly and because it operates with no knowledge about the nature of this unknown source function.

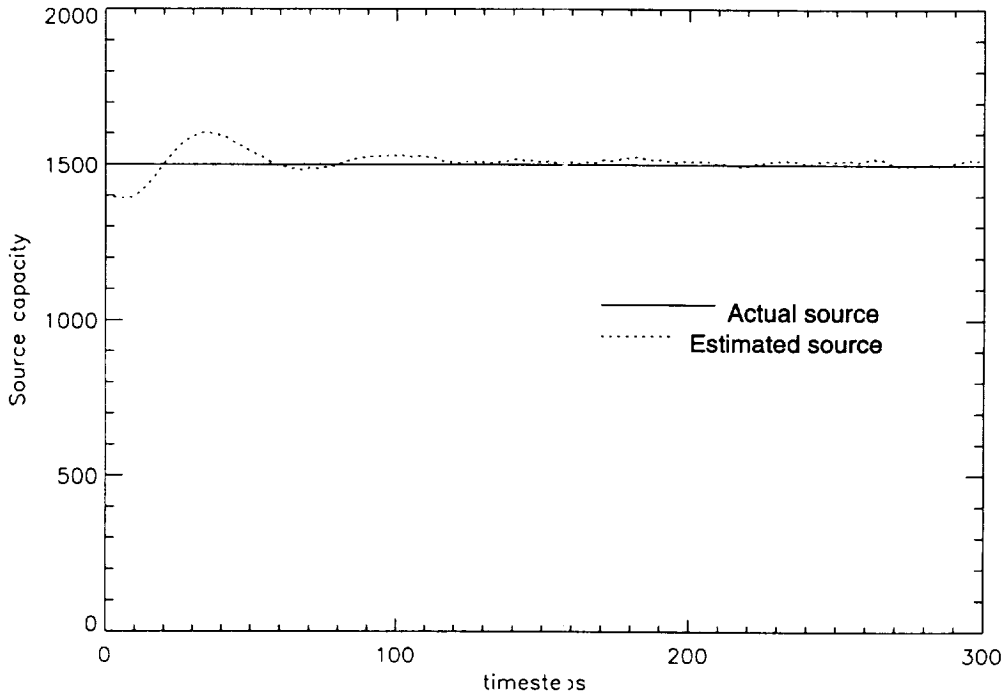

Figure 6.9: Unknown source capacity estimation– A very good first guess can lead to the Extended IKF rapidly converging to the correct solution.

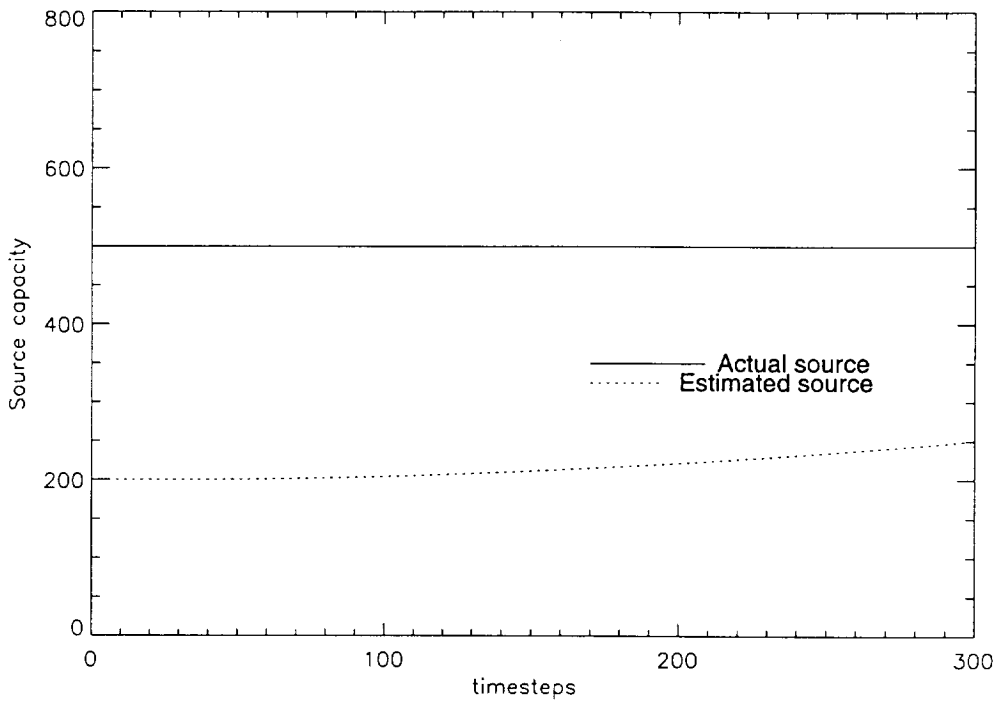

Figure 6.10: Unknown source capacity estimation when the fault is outside the active observability range

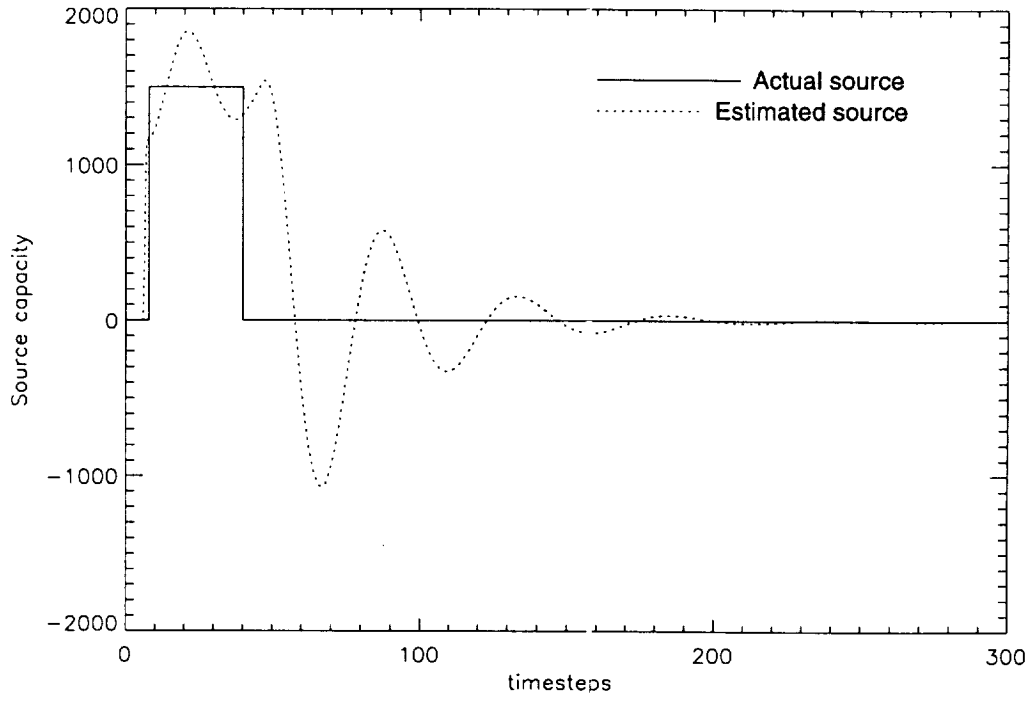

Figure 6.11: Varying source capacity estimation. The Extended Filter is able to diagnose rapidly varying functions, though wit ı some overshoot and undershoot.

In another test case, a source function that varies more slowly is used. Figure 6.12 shows the input and estimated source functions. The overshoot and undershoot is unavoidable since the filter has no *a priori* knowledge about the function.

If there were any information about the nature of the source function; for example, if a cylinder had sprung a leak, and the escape of contaminant was governed by a certain equation, Eq. 6.9 could be modified to reflect that information, and the filter would be modified accordingly. In the absence of any information, the current assumption that the source remains constant or slowly varies is found to produce satisfactory results.

6.4 Sensitivity analysis–maps

The calculation of the sensitivity coefficients, although time consuming is crucial to obtaining a good initial guess for the location and capacity of the unknown source. In addition, it also serves as a useful tool in determining the observability, and consequently the fault detectability, of the system.

The system under study presents some unique features with respect to its observability and controllability. The usual methods of determining observability and controllability (Ramirez, 1994) fail here. In the strict sense, if the state of the system is defined to be the concentrations at each mesh point, and if the controls involve varying the inlet and outlet velocities and their concentrations, this system is neither controllable nor observable. This is true for many distributed-parameter systems, especially of the reaction-diffusion parabolic type because of the way a perturbation propagates through the system, and for systems where the noise can be substantial. The sensitivity matrix contains useful information about the speed with which a perturbation travels through the system, and can be used to plot zones of observability.

Visualization of the zones would be critical in fault detection and diagnosis, especially for real-time applications, so that the operator can easily ascertain the

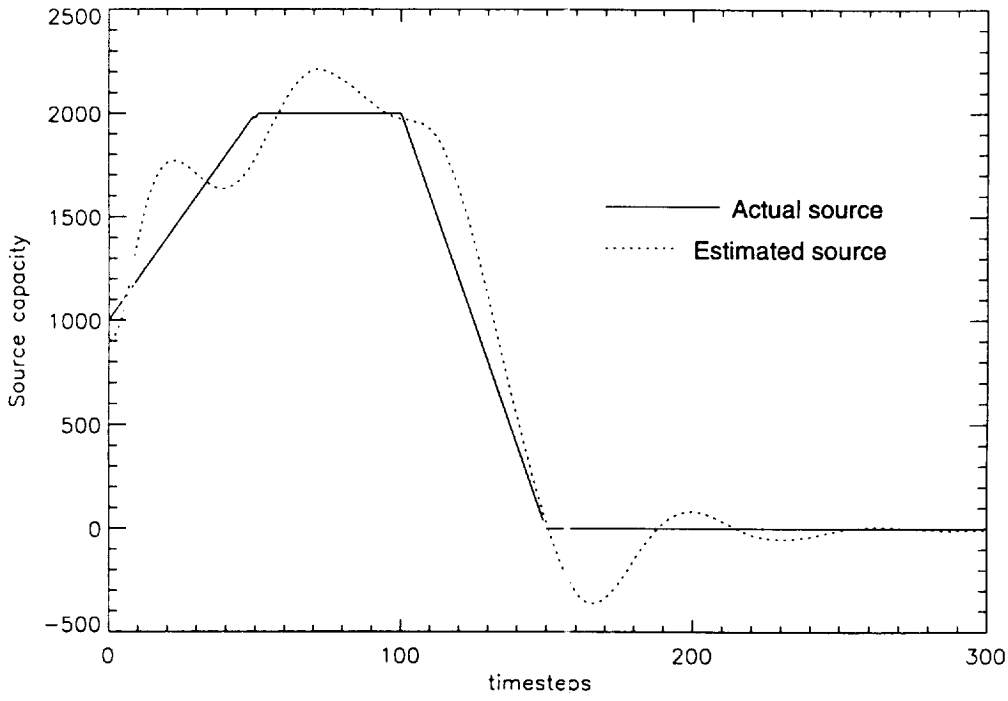

Figure 6.12: Varying source capacity estimat on. The Extended Filter can track varying unknown sources, provided that the sources lie in the region of observability.

certainty of the diagnosis. The zone is visualized through contours of equal sensitivity coefficients. Grid points that are along a constant sensitivity coefficient have equal observability and detectability. The sensor uncertainties can then be used to determine the lower limit of the sensitivity coefficient that can be detected. Grid points that have Z values below the threshold cannot be detected, and lie outside the contoured zones.

Figures (6.13-6.14) show these zones for two sample slices across the cabin. The observable zones lie within the region bounded by the contours. The overlap zone of the six sensors is clearly visible. Experiments also show that faults occurring outside the zone are difficult or impossible to diagnose accurately, though an accurate estimation of unknown sources in this zone is sometimes possible when the sensor uncertainties are very small compared to the model uncertainty (accurate sensors). The observability contours can also serve as an excellent tool for sensor selection and placement because the observability contours for different sensor configurations can be visualized in order to achieve maximum coverage of the module.

6.5 Summary and Conclusions

In this Chapter, a method for tackling a specific kind of inverse problem was developed. Inverse problems are usually ill-posed, posses no unique solutions, and small data errors can cause very large errors in the estimated solution to the problem. Most inverse problems tackled in the literature are boundary-value inverse problems, and sometimes assume a certain mathematical functional form for the function that is to be estimated. The inverse problem that arises out of the attempt to estimate the capacity and location of an unknown source provides the added complication that it is not a boundary-value problem, and that fact that there is little or no information about the nature of the source. Our solution technique consists of a combination of two commonly used techniques, sensitivity analysis, and Kalman filtering, which provides a very effective tool for estimating the location and

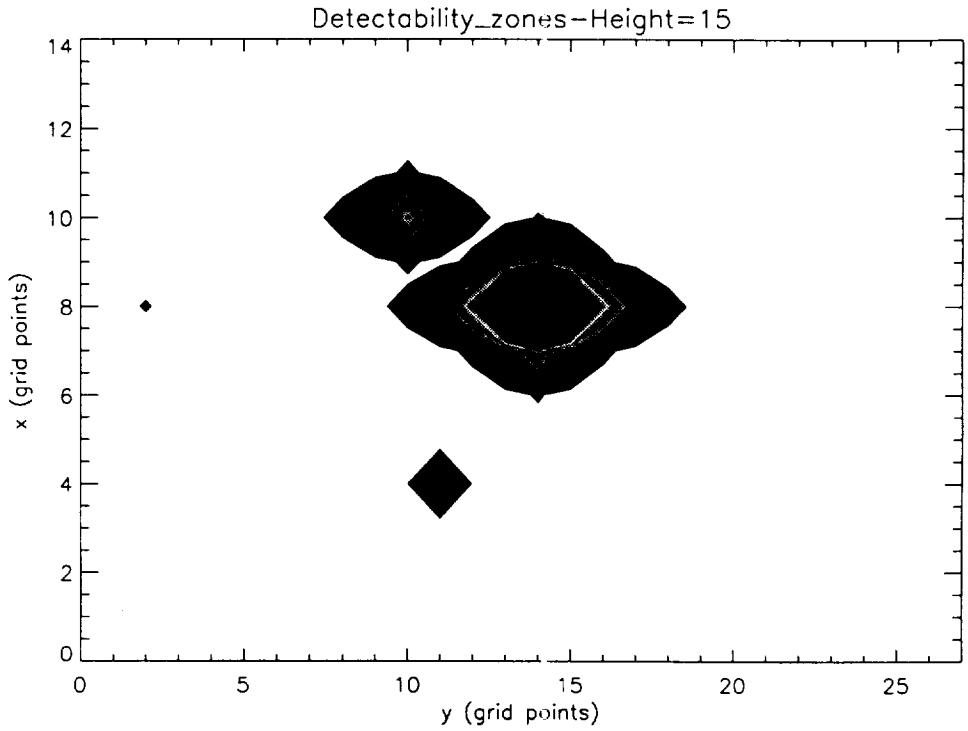

Figure 6.13: Observability and detectability co ıtours for slice at grid height 15 (1m above the floor of the cabin). The region within the contours is the observable zone.

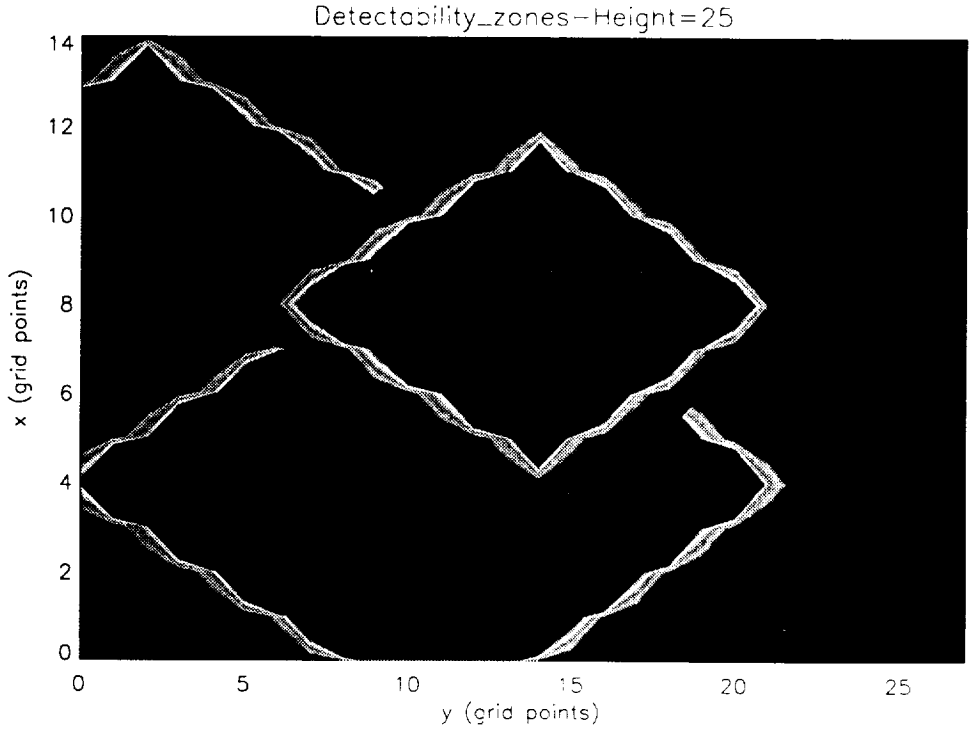

Figure 6.14: Observability and detectability contours for slice at grid height 25 (1.7m above the floor of the cabin). The observable zone lies within the contours.

capacity of the unknown sources. The two methods individually would be ineffective since sensitivity analysis does not respond very well to noisy data, and Kalman filtering is too computationally intensive to use without a good first guess for the source location. The combination of the two methods should be further investigated for its properties, and could possibly be extended to many applications that involve distributed parameter systems.

Chapter 7

Concluding remarks

7.1 Original contributions

The focus of this work was to come up with implementable algorithms for use in Space missions. This research has shown that advanced control techniques can be readily modified for use for this real-time application. Specifically, we have shown that the detailed modeling of the dispersion of air contaminants, their monitoring and detection can be accomplished with reasonable computational power. Every attempt has been made to tailor the work to NASA specifications (Technology Development Requirements, 1996).

Among the original contributions of this work are:

- Proving the inadequacy of lumped systems for air contaminant monitoring, and the limitations of two-dimensional models.
- Using CFD techniques for modeling the flow aboard the Space Station, hitherto not attempted
- Development of the Extended Implicit Filter, and using it in combination with Sensitivity analysis for joint state and parameter estimation.
- The use of sensitivity maps in identifying dead zones, and observable/detectable zones aboard the Space Station.

- Overall, providing a framework for accurate and efficient air contaminant monitoring using algorithms and methods suited for real-time implementation, and including sensor placement decisions.

7.2 Possible applications

This research can be used as a starting point for implementing future generation air contaminant monitoring systems for the Space Station and for long range Space missions. Some of the specific tasks that this research can be used for include:

- Design of habitation geometry and forced flow sources and sinks. This research can be used to identify stagnation zones in the cabin, and to design better ventilation systems to maximize crew safety and comfort.

- Sensor selection and placement issues- The sensitivity coefficients have been shown to be an effective measure of the observability and detectability of faults aboard the spacecraft. The visualization of the sensitivity coefficient provides an easy tool for facilitating sensor selection in order to ensure that all parts of the cabin are adequately covered by the measurement system.

- Analysis of proposed responses to emergencies aboard the spacecraft. The algorithms developed here can be used to simulate possible accidents, and study the effect of the remedial measures on the spacecraft atmosphere. For example, one could study how long it would take for a toxic leak in one module to spread to another module, and how long it would take for a module to be habitable after a release occurred and the removal devices are activated. In addition, proposed remediation measures for many different kinds of accidents can be analyzed to ascertain their relative merits and demerits.

- Analysis and design of fire extinguishment systems. The effect of fires, the transport of combustion products, and the air flow patterns in the likelihood

of the air supply being cut off can be studied for a variety of conditions.

- Design of automated and semi-automated control systems, which use the measurement system and model inputs to provide for operations that do not need human intervention.

7.3 Directions for future research

Computational Fluid Dynamics (CFD) has not been adequately utilized in Space station environmental control applications. CFD could prove to be a very powerful tool in the design of Space station modules for crewed missions. A detailed study of these flows is therefore in order. The next step needs to be an experimental validation of our results under both atmospheric and micro-gravity conditions. The validation needs to occur in two specific areas–one is to study room air flows, using the cabin if possible, and observing how much noise is usually present in the measurements, and in noting the performance of the filter itself. Tuning the gains via sensor and model uncertainties is another area of work that needs scrutiny. Simulations have shown that a proper choice of these parameters will affect filter gains, and the consequent performance of the filter. The Implicit Kalman Filter and Extended Filter could both be optimized to improve performance. Alternative formulations of the filter which do not require the inversion of matrices would reduce the number of operations needed per time step.

Bibliography

Alifanov, O. M. (1994). Inverse Heat Transfer Problems. Springer-Verlag.

Anderson, D. A., Tannenhill, J. C., and Pletcher, R. H. (1984). Computational Fluid Mechanics and Heat Transfer. McGraw-Hill.

Bacskay, A. S. and DaLee, R. C. (1991). Space station freedom ECLSS design configuration: A post restructure update. In Space Station ECLSS and Thermal Control, pages 195–212. Society for Automotive Engineers, Inc.

Barber, E. M., Sokgansanj, S., Lapman, W. P., and Ogilvie, J. (1982). Stability of air flow patterns in ventilated spaces. Technical Report 82-4551, American Society of Agricultural Engineers.

Barker, R. S., Russell, M. R., and Whitmer, L. R. (1991). Mathematical modeling of a four-bed molecular sieve with $C0_2$ and H_2O collection. In Space Station ECLSS and Thermal Control, pages 297–314. Society for Automotive Engineers, Inc.

Basseville, M. (1988). Detecting changes in signals and systems - a survey. Automatica, 24:309–326.

Basseville, M. (1997). Information criteria for residual generation and fault detection and isolation. Automatica, 33(5):783–803.

Basseville, M. and Nikiforov, I. V. (1993). Detection of Abrupt Changes: Theory and Application. Prentice Hall.

Beck, J. V., Blackwell, B., and Clair, C. R. S. (1985). Inverse Heat Conduction: Ill-posed Problems. John Wiley & Sons.

Bird, R. B., Stewart, and Lightfoot (1960). Transport Phenomena. John Wiley & Sons.

Blackwell, C. (1998). Elementary fault detection and location based on physical laws and M, A, and S. SAE Technical Paper Series SAE 981810, SAE International.

Carlson, N. A. (1990). Federated square root filter for decentralized parallel processes. IEEE Transactions on Aerospace and Electronic Systems, 26:517–525.

CASE/A (1990). CASE/A: Computer Aided System Engineering and Analysis User's Manual. McDonnel Douglas System Co., Huntsville, AL.

Chin, T. M., Karl, W. C., and Willsky, A. S. (1995). A distributed and iterative method for square root filtering in space-time estimation. Automatica, 31:67–82.

Chui, C. K. and Chen, G. (1991). Kalman Filtering with Real-Time Applications. Springer-Verlag.

Cole, H. E., Manuel, S., Rather, D. R., Ward, S. E., Jones, K., Perry, J., Gouzenberg, A., Savina, V., and Mikos, K. (1996). Mir space station trace contaminant assessment. SAE Technical Paper Series SAE 961472, SAE International.

Douglas, Jr., J. (1962). Alternating direction methods for three space variables. Numerische Mathematik, 4:41–63.

Douglas, Jr., J. and Rachford, Jr., H. H. (1956). On the numerical solution of heat conduction problems in two and three space variables. Transactions of the American Mathematical Society, 82:421–439.

Emami-Nacini, A., Akhter, M. M., and Rock, S. M. (1986). Robust detection, isolation and accommodation for sensor failures. Contractor Report CR-174825, NASA.

Enting, I. G. (1993). Inverse problems in atmospheric constituent studies:3. estimating errors in surface sources. Inverse Problems, 9:649–665.

Enting, I. G. and Newsam, G. N. (1990). Inverse problems in atmospheric constituent studies:2. sources in the free atmosphere. Inverse Problems, 6:349–362.

FIDAP (1993). FIDAP Version 7.6. Fluid Dynamics International, Inc., 500 Davis Street, Suite 600 Evanston IL 60201.

Frank, P. M. (1990). Fault diagnosis in dynamic systems using analytical and knowledge based redundancy- a survey and some new results. Automatica, 26:459–474.

Gan, G. (1995). Numerical investigation of local thermal discomfort in offices with displacement ventilation. Energy and Buildings, 23:73–81.

Gertler, J. (1991). Analytical redundancy methods in fault detection and isolation. In IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for Technical Processes, SAFE-Process '91, pages 9–21, Baden-Baden, Germany.

Godunov, S. K. (1959). Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics. Matematicheskii Sbornik, 47(3):271–306.

Hadamard, J. (1923). Lectures on the Cauchy problem in Linear Partial Differential Equations. Yale University Press, New Haven.

Halme, A. and Seikainaho, J. (1986). An adaptive filtering based method to detect sensor/actuator faults. In Proc. IFAC Workshop, Fault Detection and Safety in Chemical Plants, pages 158–161.

Himmelblau, D. M. (1986). Fault detection and diagnosis - today and tomorrow. In Proc. IFAC Workshop, Fault Detection and Safety in Chemical Plants, pages 95–105.

Huang, C.-H. and Wu, J.-Y. (1994). Two-dimensional inverse problem in estimating heat fluxes of an enclosure with unknown internal heat source. J. Appl. Phys., 78(1):133–141.

Jordan, T. J. (1967). Experiments on error growth associated with some linear least-squares procedures. Mathematics of Computation, 22:579–588.

Jover, J. M. and Kailath, T. (1986). A parallel architecture for kalman filter measurement update and parameter estimation. Automatica, 22:43–57.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82:35–45.

Kalman, R. E. and Bucy, R. S. (1961). New results in linear filtering and prediction theory. Journal of Basic Engineering, 83:95–100.

Kirsch, A. (1996). An Introduction to the Mathematical Theory of Inverse Problems. Springer-Verlag, New York.

Korbicz, J., Fathi, Z., and Ramirez, W. F. (1991). State estimation schemes for fault detection and diagnosis in dynamic systems. Applied Math. and Computer Science, 1(1):83.

Kuehn, T. H. (1988). Computer simulation of airflow and particle transport in cleanrooms. Journal of Environmental Sciences, 31(5):21–27.

Kurabuchi, T., Fang, J. B., and Grot, R. A. (1990). A numerical method for calculating indoor airflows using a turbulence model. Technical report, National Institute of Standards and Technology, Gaithesberg, MD.

Kurpisz, K. and Nowak, A. J. (1995). Inverse Thermal Problems. Computational Mechanics Publications, Southampton, UK and Boston, USA.

LeVeque, R. J. (1985). Intermediate boundary conditions for LOD, ADI and appropriate factorization methods. NASA Contractor Report 172591, National Aeronautics and Space Administration.

Lunar-Mars (1997). Lunar-Mars Life Support Test Project Phase II Final Report. Crew and Thermal Systems Division, JSC NASA.

Meditch, J. S. (1969). Stochastic Optimal Linear Estimation and Control. McGraw-Hill, New York.

Morf, M. and Kailath, T. (1975). Square root algorithms for least-squares estimation. IEEE Transactions on Automatic Control, 20:487–497.

Nanda, A. and Das, P. (1996). Determination of the source term in the heat conduction equation. Inverse Problems, 12:325–339.

Narayan, A. P. and Ramirez, W. F. (1998a). Intelligent system for the detection and diagnosis of spacecraft air contaminants. In Proceedings of the 28th International Conference on Environmental Systems. SAE International.

Narayan, A. P. and Ramirez, W. F. (1998b). Three-dimensional modeling, estimation, and fault diagnosis of spacecraft air contaminants. Journal of Spacecraft and Rockets, 35(4):565–574.

National Academy of Sciences (1981). Indoor Pollutants. National Academy of Sciences, Washington.

National Research Council (1984). Emergency and Continuous Exposure Limits for Selected Airborne Contaminants. National Research Council, Washington D.C.

National Research Council (1992). Guidelines for Developing Spacecraft Maximum Allowable Concentrations for Space Station Contaminants. National Research Council, Washington D.C.

National Research Council (1996). Spacecraft Maximum Allowable Concentrations for Selected Airborne Contaminants. National Research Council, Washington D.C.

Newsam, G. N. and Enting, I. G. (1988). Inverse problems in atmospheric constituent studies: 1. determination of surface sources under a diffusive transport approximation. Inverse Problems, 4:1037–1054.

Paige, C. C. and Saunders, M. A. (1977). Least squares estimation of discrete linear dynamic systems using orthogonal transformations. SIAM Journal of Numerical Analysis, 14:180–193.

Perry, J. L. (1993). Computerized atmospheric trace contaminant control simulation for manned spacecraft. Technical Report TM-108409, George C. Marshall Space Flight Center.

Ramirez, W. F. (1994). Process Control and Identification. Academic Press.

Reuter, J. L. (1998). International space station environmental control and life support system status:1997-1998. In Proceedings of the 28th International Conference on Environmental Systems. SAE International.

Richards, R. F., Munk, B. N., and Plumb, O. A. (1997a). Fire detection, location and heat release rate through inverse problem solution. part 1: Theory. Fire Safety Journal, 28:323–350.

Richards, R. F., Ribail, R. T., Bakkom, A. W., and Plumb, O. A. (1997b). Fire detection, location and heat release rate through inverse problem solution. part 2: Experiment. Fire Safety Journal, 28:351–378.

Roache, P. J. (1972). Computational Fluid Dynamics. Hermosa publishers, P. O. Box 8172, Albuquerque, NM 87108.

Roy, S., Hashemi, R. H., and Laub, A. J. (1991). Square root parallel kalman filtering using reduced order lower filters. IEEE Transactions on Aerospace and Electronic Systems, 27:276–289.

Said, N. A., Zhang, J. S., Shaw, C. Y., and Christianson, L. L. (1995). Computation of room air distribution. ASHRAE Transactions, 101:1065–1076.

Skliar, M. (1996). State estimation and fault diagnosis for distributed parameter transport processes with application to air contamination control. PhD thesis, University of Colorado.

Skliar, M. and Ramirez, W. F. (1997a). Air-quality monitoring and detection of air contamination in an enclosed environment. Journal of Spacecraft and Rockets, 34(4):522–532.

Skliar, M. and Ramirez, W. F. (1997b). Implicit kalman filtering. International Journal of Control, 66(3):393–412.

Smith, G. J. (1996). Sensor Selection Methodology for detecting and monitoring toxic concentrations of airborne contaminants in a space habitat. PhD thesis, University of Colorado.

Son, C. and Barker, R. S. (1997). U.s. lab-a module cabin air distribution in space station. SAE, 66(3):393–412.

Talenti, G. and Tonani, F. (1995). Sounding for underground gas sources by surface physico-chemical methods: a mathematical basis. Inverse Problems, 11:1265–1297.

Tam, L. T. (1998). Application of CFD in the software design of environmental control and life support systems for the international space station. SAE Technical Paper Series SAE 981579, SAE International.

Technology Development Requirements (1996). Technology Development Requirements. JPL D-13832.

Todd, P., Skliar, M., Smith, G., Morgenthaler, G., Mckinon, J., Oberdoerster, G., and Shulz, J. (1994). Inhalation risk in low-gravity spacecraft. Acta Astronautica, 33:304–315.

Whitaker, S. (1968). Introduction to Fluid Mechanics. Prentice Hall, Inc.

White, F. M. (1974). Viscous Fluid Flow. McGraw-Hill.

Whittle, G. E. and Clancy, E. M. (1991). Evaluation of cases b,d,e- presentation of results from measurements and simulations. Research Report 1.22, International Energy Agency Annex 20.

Yamamoto, T., Donovan, R. P., and Ensor, D. S. (1988). Model study for optimization of cleanroom airflows. Journal of Environmental Sciences, 31(6):24–29.

Zhang, J. S., Wu, G., and Christianson, L. (1992). Full-scale experimental results on the mean and turbulent behavior of room ventilation flows. ASHRAE Transactions, 98(2):307–318.

Nomenclature

$\mathbf{A}_1$	= State Transition Matrix (Left Hand Side)
$\mathbf{A}_2$	= State Transition Matrix (Right Hand Side)
$\mathbf{A}_x$,$\mathbf{A}_y$,$\mathbf{A}_z$	= Discrete representation of the spatial operator for the State Transition Matrices in each of the co-ordinate directions
α, β	= Zero-mean white Gaussian processes
c_μ, c_1, c_2	= Dimensionless empirical constants used in the $k-\epsilon$ turbulence model
$\mathbf{C}$	= Stochastic model disturbance transition matrix
d	= Discrete analog of the mass or eddy diffusivity (m^2s^{-1})
D_M	= Mass Diffusivity (m^2s^{-1})
$\overline{D_M}$	= Eddy Diffusivity (m^2s^{-1})
$\mathbf{e}$	Error m-tuple used in sensitivity analysis
$\mathbf{f}$	= Contaminant source capacity ($kgm^{-3}s^{-1}$)
$\mathbf{f_u}$	= Unknown contaminant source capacity ($kgm^{-3}s^{-1}$)
f_x, f_y, f_z	= Body force per unit mass acting on fluid N/m
h	= Elements in the measurement matrix, $\mathbf{H}$
$\mathbf{H}$	= Measurement matrix
$\mathbf{I}$	= Identity matrix
k	= Kinetic energy of turbulence (used in the $k-\epsilon$ model)
k_m	= Constant of integration for the unknown source capacity
l	= Loss function
$\mathbf{L}_m$	= Implicit Kalman Gain
p	= Pressure of fluid Nm^{-2}
$\mathbf{P}^y_{m+1\|m}$	= Predicted error covariance matrix
$\mathbf{P}^y_{m+1\|m+1}$	= Estimation error covariance matrix
$\mathbf{P}^q_{m+1\|m+1}$	= Covariance matrix of the state

q = Concentration of Contaminant (kgm^{-3}) (volume %)

$\mathbf{q}$ = Discrete analog of contaminant concentration (kgm^{-3}) or (volume %)

$\dot{q}$ = Volumetric heat flux addition/removal to the fluid $Jm^{-3}s^{-1}$

$\mathbf{Q}$ = Covariance matrix (diagonal matrix) of the model noise

$\mathbf{Q}_m$ = Concentration vector

$\mathbf{r}$ = Row unity matrix

$\mathbf{R}$ = Covariance matrix (diagonal matrix) of the measurement noise

$\mathbf{S}$ = Covariance matrix for the unknown source capacity noise

t = time (s)

T = Fluid temperature (K)

$\mathbf{u}$ = Velocity vector

$u,\ v,\ w$ = Velocity components in the co-ordinate directions (ms^{-1})

w_1 = Stochastic model noise

$\mathbf{W}$ = Location(s) of the unknown source(s) used in fault diagnosis

$\mathbf{x}(k)$ = State vector

$\hat{\mathbf{x}}$ = Estimate of State

$\tilde{\mathbf{x}}$ = Estimation Error

$x,\ y,\ z$ = Co-ordinate directions or positions (m)

$\mathbf{z}(i)$ = Measurement signal

$\mathbf{Z}$ = Sensitivity matrix

Subscripts

$n,\ p,\ r$ = Mesh indices in the coordinate directions

Superscripts

a = Augmented quantities, augmented to include unknown source(s) in the model

E = East

W = West

N	= North
S	= South
U	= Up
D	= Down

Greek Symbols

$\alpha_x, \alpha_y, \alpha_z$	= State transition matrices used in the double filter
β, γ	= discretization steps in the spatial directions (ADI method)
ϵ	= Dissipation rate of the turbulence (used in the $k-\epsilon$ model)
λ	= Second viscosity coefficient ($m^2 s^{-1}$)
μ	= Molecular Viscosity coefficient ($m^2 s^{-1}$)
μ_t	= Eddy Viscosity coefficient ($m^2 s^{-1}$)
Φ_k	= Error Functional
ρ	= Density of fluid (kgm^{-3})
$\sigma_k, \sigma_\epsilon$	= dimensionless empirical constants used in the turbulence model