Modelling the Concentration Distribution of Non-Buoyant Aerosols Released from Transient Point Sources into the Atmosphere

by

Xiaoying Cao

A thesis submitted to the Department of Chemical Engineering

In conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

(October, 2007)

Abstract

Neural network models were developed to model the short-term concentration distribution of aerosols released from point sources. Those models were based on data from a wide range field experiments (November 2002, March, May and August 2003). The study focused on relative dispersion from the puff centroid. The influence of puff/cloud meandering and large-scale gusts were not considered, the modelling was limited to studying the dispersion caused by small-scale turbulence. The data collected were based on short range/time dispersion, usually shorter than 150 s. The ANN (Artificial Neural Network) models considered explicitly a number of meteorological and turbulence parameters, as opposed to the Gaussian models that used a single fitting parameter, the dispersion coefficient. The developed ANN models were compared with predictions generated from COMBIC (Combined Obscuration Model for Battlefield Induced Contaminants), a sophisticated model based on Gaussian distributions, and a traditional Gaussian puff model using Slade's dispersion coefficients. Neural network predictions have been found to have better agreement with concentration measurements than either of the other two Gaussian puff models. All models underestimate the maximum concentration, but ANN predictions are much closer to observations. Simulations of concentration distributions under different stability conditions were also checked using the developed ANN model, and it showed that, for a short time, Gaussian distributions are a good fit for puff dispersion in the downwind, crosswind and vertical directions.

For Gaussian puff models, the key issue is to determine appropriate dispersion coefficients (standard deviations). ANN models for puff dispersion coefficients were trained and their average predictions were compared with the results of measurements. Very good agreement was observed, with a high correlation coefficient (>0.99). The ANN models for dispersion coefficients were used to analyze which input variables were more significant for puff

expansions. Dispersion time, particle position relative to the centroid, turbulent kinetic energy and insolation showed the most significant influence on puff dispersion. The Gaussian puff model with dispersion coefficients from the ANN models was compared with COMBIC and a Gaussian puff model using Slade's dispersion coefficients. Generally speaking, predictions generated by the Gaussian puff model with dispersion coefficients generated by ANN models showed better agreement with concentration measurements than the other two Gaussian puff models, by giving a much higher fraction within a factor of two, and lower normalised mean square errors.

Acknowledgements

I would like to express my gratitude to those who supported me throughout this research project, without whose help none of this would have been possible. I am particularly indebted to Dr. W. S. Andrews for providing the opportunity to undertake this project, and for his unwavering confidence in me throughout the course of this work.

Special thanks to my co-supervisor Dr. Gilles Roy at Defense Research and Development Canada, Valcartier, who took time out of his busy schedule to patiently answer my questions, conduct field trials, entertain me and make me feel perfectly at home in Valcartier. Additional thanks go to Dr. Luc Bissonnette, Dr. Luc Forand and Sylvain Cantin for their insights and assistance. Thanks to all those personnel at DRDC-Valcartier for the help of the success of field trials.

I would also like to thank the Queen's university for awarding me the DUNCAN & URLLA CARMICHAEL GRADUATE FELLOWSHIP and all the Graduate Awards.

Financial support which made this work possible is gratefully acknowledged from the Academic Research Program (ARP) and the Direct General Nuclear Safety (DGNS), both agencies of the Department of National Defense, and Natural Sciences and Engineering Research Council of Canada (NSERC).

Statement of Originality

I hereby certify that all of the work described within this thesis is the original work of the author. Any published (or unpublished) ideas and/or techniques from the work of others are fully acknowledged in accordance with the standard referencing practices.

(Xiaoying Cao)

(October, 2007)

Table of Contents

List of Tables

List of Figures

Glossary

A	Pasquill stability category: very unstable
A	receiver area (m^2)
ADAM	Air Force Dispersion Assessment Model
ANN	Artificial Neural Network
ARP	Academic Research Program
$\alpha(r)$	volume extinction coefficient (1/m)
B	Pasquill stability category: moderately unstable
BP	Backpropagation
$\beta(r)$	volume backscattering coefficient ($1.0/(ms_r)$)
C	Pasquill stability category: slightly unstable
c	speed of light
c	aerosol concentration (g/m^3)
θ	Moment of concentration
COMBIC	Combined Obscuration Model for Battlefield Induced Contaminants
C_N^2	Refractivity parameter ($1.0/m^{2/3}$)
D	Pasquill stability category: neutral
DGNS	Direct General Nuclear Safety
DRDC	Defensive Research and Development of Canada
E	Pasquill stability category: slightly stable
EC	Environment Canada
EDBD	Extended delta-bar-delta
F	Pasquill stability category: stable
$F(r)$	Lidar system constant
$f()$	linear or nonlinear transformation function
FB	fractional bias
F_2	Factor of 2
F_{10}	Factor of 10
G	Pasquill stability category: very stable

g	internal value of inputs or outputs after linear transformation
INPUFF	Integrated puff model
K	Lidar constant
k	coefficient in conversion from relative to absolute concentration
k	lidar ratio of backscattering coefficient to extinction coefficient ($1.0/s_r$)
k_x	turbulent diffusivity in the downwind direction (m^2/s)
k_y	turbulent diffusivity in the crosswind direction (m^2/s)
k_z	turbulent diffusivity in the vertical direction (m^2/s)
L	Mornin-Obukhov length (m)
LCM	Laser Cloud Mapper
LIDAR	LIght Detection and Ranging
LS	Lagrangian Stochastic
M	normalized concentration from measurement
MG	Geometric mean bias
MLFF	Multi-layer feedforward
NMSE	Normalized mean square error
NSERC	Natural Sciences and Engineering Research Council of Canada
P	Air pressure (kPa or hPa)
P	ANN prediction of normalized concentration
P_0	Initial laser pulse power (J)
$P(r)$	Lidar received signal power backscattered from distance r (W)
PBL	planetary boundary layer
PDF	probability density function
P-G	Pasquill-Gifford
PPM	Puff Particle Model
Q	aerosol release mass (g)
SEM	Scanning Electron Microscope
SPELL	SPectral Eulerian Lagrangian model
σ_x	downwind dispersion coefficient (m)
σ_{x0}	downwind dispersion coefficient corresponding to puff initial size (m)
σ_{xs}	downwind dispersion coefficient caused by wind shear

σ_{xt}	downwind dispersion coefficient caused turbulent diffusion
σ_y	crosswind dispersion coefficient (m)
σ_z	vertical dispersion coefficient (m)
T	Temperature (ºC)
t	dispersion time (s)
t_p	duration of laser pulse (s)
τ	optical depth or thickness
TG	Temperature gradient (ºC/m)
TIB	Total Integrated Backscatter
TKE	Turbulent Kinetic Energy (m^2/s^2)
ρ	aerosol density (g/m^3)
R	correlation coefficient
R	Radial distance from lidar (m)
R_b	nearest cloud boundary (m)
R_e	farthest cloud boundary (m)
RH	Relative Humidity (%)
RHG	Relative Humidity gradient (%/m)
U	mean wind speed (m/s)
$\bar{u}$	mean wind speed in downwind direction (m/s)
u	wind component in downwind direction (m/s)
V	Visibility (km)
VG	Geometric mean variance
v	wind component in crosswind direction (m/s)
w	wind component in vertical direction (m/s)
W_1	Solar irradiation (W/m^2)
W_2	Ground heat flux (W/m^2)
x	relative downwind distance (m) or x axis
x	value of inputs or outputs with physical unit
y	relative crosswind distance (m) or y axis
y	value after linear or nonlinear transformation
z	relative vertical distance (m) or z axis

z	height of cloud (m)
z_o	Surface roughness length

List of Publications

Andrews, W.S., **X. Cao,** G. Roy, L. Forand and G. Potvin (2003) Field Data Collection in Support of Dispersion Modelling. Proceedings of Battlespace Atmospheric and Cloud Impacts on Military Operations Conference 2003, Monterey CA, NRL, September 2003.

Cao. X., W.S. Andrews., G. Roy, L. Forand and G. Potvin (2005) Modelling the Dispersion of Aerosols from Point Sources. Proceedings of Battlespace Atmospheric and Cloud Impacts on Military Operations Conference (BACIMO)2005, Monterey CA, NRL, October 12-14, 2005

T. J. Devito, **X. Cao**, G. Roy, J.T. Costa and W.S. Andrews, Modelling Puff Concentration Distributions. Boundary-layer meteorology, submitted

Chapter 1
Introduction

Atmospheric flow is usually turbulent flow, and one of the most important characteristics of turbulent flow is its ability to disperse particles at a rate many orders of magnitude greater than diffusion by molecular collisions. The turbulence components (or eddies) of atmospheric flow occur over a wide range of motion scales, from the smallest scales of the order of 10^{-4} m (so-called Kolmogorov microscale) to the largest scales of the order of 10^3 to 10^4 m (Chatwin, 1990). These turbulent eddies help the dispersion of contaminants in all directions. Eddies, having a scale size larger than the instantaneous size of a cloud, hold the majority of turbulent energy and determine the direction of the flow movement. These large eddies give rise to a random motion or meandering of the entire cloud. By comparison, eddies with scales smaller than or comparable with the instantaneous size of the cloud tend to have no preferred direction on average, and their structure is of a universal form. These smaller scale eddies contribute to the expansion of the cloud. However, as the cloud expands, fluctuations characterized by a scale size which originally gave rise to a bulk displacement of the cloud may, at a later instant, be smaller than the size of the cloud, and consequently, contribute instead to the expansion. Therefore, intuitively, the cloud expands at an increasing rate because a larger and larger fraction of the spectral energy will be available for the expansion. When the cloud has grown to an extent that less and less extra eddies will be available, the growth rate will decrease (De Hann *et al.*, 1995).

Aerosol dispersion modelling is concerned with predicting the concentration distributions of a contaminant introduced into the atmosphere and its subsequent dispersion downwind. Dispersion modelling deals with two different sources, one is a continuous emission source, where the cloud formed is called a plume; the other one is an instantaneous source, with the corresponding cloud being called a puff. Most of the work to date in the field of dispersion modeling has concerned the dispersion of a pollutant from a continuous release, such as from a smokestack or evaporating pool. However, the study of dispersion from a near-instantaneous release has been sparse, both in theoretical treatment and in experimental trials, even though it is very important for environmental protection, industry hazard control and military defense.

Far more attention has been given to studying and modelling plume dispersion than on puff dispersion. Consequently, in order to enhance the understanding of non-buoyant aerosol dispersion from transient point sources at ground level into the atmosphere, a series of week-long field trials were conducted at Defense Research and Development Canada (DRDC)-Valcartier in November, 2002 and March, May and August 2003, to examine the influence of a variety of atmospheric and meteorological factors on aerosol (and, by extension, vapour) dispersion in the surface boundary layer over flat terrain. Factors considered included wind velocity, air temperature and pressure, along with their associated gradients, humidity, irradiation, ground heat flux, as well as diurnal and seasonal variations. The influence of derived turbulence statistics was also examined. These trials were aimed at removing seasonal biases and covering as broad a range of atmospheric and meteorological conditions as possible.

The Laser Cloud Mapper (LCM) of DRDC-Valcartier was used to obtain the spatial and temporal concentration distributions of aerosols released from a generator. Three different micron-sized aerosol materials were released individually every three hours, and about 500 releases were performed per trial period. The collected data were then used for further analysis and modelling.

The arrangement of this thesis is as follows: Chapter 1 is an introduction and a literature review about diffusion theories, puff modelling, lidar measurements and ANN (Artificial Neural Network) modelling applications to puff dispersions. Following this, the experimental set up and lidar inversion will be introduced in Chapter 2, and some measurements under different conditions will be checked. The influence of wind shear on dispersion in the downwind direction will be briefly discussed and illustrated. ANN modelling for puff dispersion and its relationship with key meteorological parameters are discussed in Chapter 3 and Chapter 4 respectively. Modelling of dispersion coefficients for Gaussian puff models will be discussed in Chapter 5. The growth rate of dispersion parameters with time and some other parameters in longitudinal, lateral, and vertical directions will also be investigated. ANN model predictions will be compared with two Gaussian puff models (Slade and COMBIC). The last part will be a discussion of the limitations of the model developed. Finally, the work reported will be summarized along with some possible improvements for the future.

1.1 Theory of Atmospheric Dispersion

Two basic approaches are in use to describe diffusion or dispersion in a turbulent fluid: Eulerian and Lagrangian. The Eulerian statistical description of dispersion refers to the time averaging at a fixed point in space, therefore Eulerian is better suited for

measurements and observations at fixed points in space in either a fixed or a moving frame of reference. The Lagrangian statistical description follows the movement of a volume of fluid with time, so it may be the most appropriate way to describe turbulent dispersion. This notwithstanding, it is generally impossible to perform Lagrangian measurements, and equations of Eulerian motion are much simpler than those of Lagrangian motion. A brief description of both statistical theories will be introduced below.

1.1.1 Eulerian turbulent dispersion

When the motion of the fluid is turbulent, there is usually a significant augmentation of mass transfer from regions of high to low mass fraction of a particular component. The influence of this mechanism is usually referred as turbulent dispersion, and is entirely dominant over molecular diffusion in the atmospheric boundary layer (Dobbins, 1979).

Dispersion of pollutants in the atmosphere is a random process, it is impossible to predict the instantaneous concentration due to highly irregular three-dimensional turbulent motions of the atmosphere. Ensemble averaging the instantaneous diffusion equation makes it possible to find solutions to turbulent diffusion problem. Eqn. (1.1) is the basic diffusion equation for non-reactive contaminants, ignoring molecular diffusion:

$$\frac{\partial \overline{c}}{\partial t} + \overline{u}\frac{\partial \overline{c}}{\partial x} + \overline{v}\frac{\partial \overline{c}}{\partial y} + \overline{w}\frac{\partial \overline{c}}{\partial z} = -\left(\frac{\partial(\overline{c'u'})}{\partial x} + \frac{\partial(\overline{c'v'})}{\partial y} + \frac{\partial(\overline{c'w'})}{\partial z} \right), \qquad (1.1)$$

where $\overline{c}$ is the ensemble-averaged concentration, u, v and w are wind components in x, y and z directions, overbar represents the ensemble average of the corresponding variable, and variables with primes represent their turbulent parts. The correlations $\overline{c'u'}$,

$\overline{c'v'}$ and $\overline{c'w'}$ are turbulent advection terms, and are called turbulent mass fluxes, caused by wind fluctuations in the x, y and z directions, analogous to the molecular mass fluxes. These turbulent components dominate the diffusion of the pollutant particles in the atmosphere.

The diffusion equation of Eqn. (1.1) for mean concentration is not closed. To solve Eqn. (1.1) to get mean concentration, $\bar{c}$, three other equations for the turbulent diffusion terms need to be added. The simplest and probably the most widely used closure approach in turbulent diffusion is first-order closure, based on the gradient-transport hypothesis. According to this hypothesis, turbulent mass fluxes are proportional to the gradient of the mean concentration. This is true for most situations in the surface layer, with several exceptions, such as in convective conditions (Arya, 1999). Thus

$$\overline{u'c'} = -K_x \frac{\partial \bar{c}}{\partial x}, \tag{1.2a}$$

$$\overline{v'c'} = -K_y \frac{\partial \bar{c}}{\partial y}, \tag{1.2b}$$

and

$$\overline{w'c'} = -K_z \frac{\partial \bar{c}}{\partial z}, \tag{1.2c}$$

where the "standard" orientation of axes are implied, i.e., the x axis follows the mean wind direction, for convenience, the y is the crosswind direction, and the z is the vertical direction. K_x, K_y and K_z are turbulent diffusivities in the downwind, crosswind and vertical directions respectively. Turbulent diffusivities have to be specified for the complete closure and solutions of Eqn. (1.1) can be obtained for any desired initial, boundary, and source conditions. In stationary and homogeneous turbulent atmospheric flow, eddy diffusivities are not expected to be dependent on time and space. For the

specified constants K_x, K_y, and K_z, and for an instantaneous point release, the solution of Eqn. (1.1) is expressed as

$$c(x,y,z,t) = \frac{Q}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z}\exp\left(-\frac{x^2}{2\sigma_x^{\,2}}-\frac{y^2}{2\sigma_y^{\,2}}-\frac{z^2}{2\sigma_z^{\,2}}\right), \qquad (1.3)$$

where $\sigma_x^{\,2} = 2K_x t$, $\sigma_y^{\,2} = 2K_y t$ and $\sigma_z^{\,2} = 2K_z t$. σ_x, σ_y and σ_z are called dispersion coefficients in the x, y, z directions, and are functions of dispersion time. Eqn. (1.3) shows that the concentration distributions in the x, y, z directions are Gaussian in shape.

Based on Eqn. (1.3), many Gaussian puff models or modifications were developed, the differences of those Gaussian models are the use of different dispersion coefficients, e.g., INPUFF (Petersen and Lavdas, 1986) and COMBIC (Ayres and DeSutter, 1995).

In practice, Eqn. (1.3) is only useful for more or less uniform flow and homogeneous turbulence since the first-order gradient transfer hypothesis requires that the length scale of the turbulence be small compared with that of the contaminant distributions (Smith and Hay, 1961). Under conditions with strong shear and non-homogeneous turbulence or convective condition, the application of concentration distribution represented by Eqn. (1.3) may not be satisfactory.

1.1.2 Statistical theory of turbulent diffusion (Lagrangian approach)

Eulerian measurements include fixed-point measurements of fluid motion, and always involve a continually changing sample of fluid and particles. It is relatively easy to do the measurements, but the Eulerian motion may not be the most appropriate for describing the dispersion of contaminants (gases or particles) in a turbulent flow. For this,

a Lagrangian description of motion associated with tagged fluid particles is found to be more appropriate. In this way, one follows the paths of certain tagged particles, which are subjected to random movements in the fluid environment, and deduce their dispersion characteristics, such as concentration and spatial distribution, from the statistics of particle displacements.

Two Lagrangian dispersion theories exist, one is called 1-particle displacement theory, dealing with particle displacement from its source, and is suitable for dispersion from a continuous source; the other one is Lagrangian two-particle displacement theory which considers the separation of two different particles with time, and is appropriate for puff dispersion. Lagrangian statistical theories relate dispersion directly to the Lagrangian correlation function, integral scales and turbulence (Arya, 1999). The Lagrangian correlation function is the key part of statistical dispersion theory. Different approximations of this function are used (Sawford, 1982; Mikkelsen *et al.*, 1987; Arya, 1999).

The statistical theory is based on assumptions of uniform mean winds, homogeneous turbulence and stationary flow (under these assumptions, the Gaussian distribution is usually considered). For a puff formed from an instantaneous source, if short range or near field dispersion is concerned, the initial spread of the puff is important to expansion of the puff, and the expansion is proportional to the dispersion time. When the dispersion time is long enough that the initial spread does not contribute to the final expansion, puff expansion is found to be proportional to $t^{1/2}$. Between the short range and long range is the inertial subrange, in which the growth of the puff is at its fastest rate and

proportional to $t^{3/2}$ (Arya, 1999). The detailed derivation of this relationship can be found in the work of Bachelor (1952), Gifford (1982), and Mikkelsen (1987).

When vertical inhomogeneity of turbulence and mean flow shear, which are characteristic features of the atmospheric surface layer, are encountered, use of the Gaussian distribution for statistical theory is questionable.

1.2 Dispersion Modelling

There are countless dispersion models available and in use, many of which vary considerably in modelling methodology. Common approaches include Lagrangian stochastic (LS) modelling, numerical solutions to the advection-diffusion equations, and Gaussian modelling.

Lagrangian stochastic particle models are widely used in research and generally accepted as being the most powerful tool to model the dispersion of atmospheric pollutants in the boundary layer, because LS models can address both diffusion and meandering of the puff centre at the same time. To solve the Lagrangian diffusion equation, however, one must deal with closure problems - approximations made involve the Lagrangian velocity field. However, the assumptions made in Lagrangian modeling are independent of the concentration field (Wilson and Sawford, 1996).

A LS particle model mathematically follows parcels of particles as they move in the atmosphere and simulates the motion of parcels as a random walk process. The Lagrangian model then calculates the particle dispersion by computing the statistics of the trajectories of a large number of the particles, which is usually excessively demanding of computing time. In order to obtain statistically reliable results, as many as in the order of 10^5 particles must be traced (de Haan, 1999). This large number of particle trajectories,

together with the often-required small time steps, leads to excessive CPU-time consumption.

The closure problem, as well as the enormous requirement of computing time, make LS models unattractive for many practical applications. Some modelers, however, employ simplifications to reduce computing time by using larger time steps, by assuming a vertically homogeneous skewness of the vertical velocity (Hurley and Physick, 1993), using a kernel method to determine the concentration fields instead of the most popular box-counting method (de Haan, 1999), introducing a higher order non-Gaussian method to solve the random displacement equation (Ermak and Nasstrom, 2000) and combining different strategies together (Schwere *et al.*, 2002). These simplifications make it possible for LS models to be used in the future as candidates for regulatory models, as well as powerful research tools.

All analytical dispersion models assume a flat and uniform surface that simply reflects the pollutants reaching it, and an idealized (horizontally homogeneous and quasi-stationary) atmospheric boundary layer. In all the available analytical models, transport and diffusion phenomena are greatly simplified and parameterized (for example, constant transport winds and diffusivities or power-law wind and diffusivity profile). For more realistic boundary conditions, variable-eddy diffusivities and height–dependent winds, it is difficult, if not impossible, to get an analytical solution of Eqn. (1.1), numerical solutions would be necessary to iteratively solve the advection-diffusion equation. This allows for the use of more realistic estimates for wind speed, temperature and eddy diffusivity profiles and solves the diffusion equation under simple forms (CCPS, 1996).

By far, the simplest and most commonly used modelling technique is Gaussian dispersion modelling. Among the most commonly used public and proprietary dispersion modelling software packages, the majority of these incorporate a Gaussian plume or puff model into the full dispersion code, such as AERMOD which uses Gaussian dispersion for stable atmospheric conditions, ADAM (Air Force Dispersion Assessment Model), a modified box and steady-state Gaussian dispersion model used in relatively uncomplicated terrain; INPUFF (Integrated puff model, Petersen and Lavdas, 1986).

Gaussian puff diffusion models have applicability limited to relatively flat and homogeneous surfaces, reasonably steady and moderate to strong winds, moderately stable and unstable conditions, neutrally buoyant or slightly buoyant emissions, and relatively short distances from simple source configurations. When near-calm, extremely stable and convective conditions are confronted, the suggested dispersion parameters may not be applicable, and more sophisticated numerical models must be used to simulate and predict dispersion in these complex flow conditions (Arya, 1999).

Many theories and experiments have shown that puff dispersion does not really obey the Gaussian distribution, especially in the vertical direction, therefore some non-Gaussian models have been developed by numerical modelling techniques (Nasstrom *et al.*, 2000; Tirabassi *et al.*, 2001; Walcek, 2002)

Other models have been developed that attempt to combine the advantages of both particle models and puff models, that is the accurate prediction of particle models as well as fast calculation, e.g., a particle-puff model or PARTPUFF (Hurley, 1994), puff-particle model or PPM (de Hann, 1998) and spectral Eulerian Lagrangian modeL (SPELL, Rizza et al, 2000), where the PARTPUFF model uses a Gaussian puff

formulation in terrain-following horizontal directions, and a Lagrangian particle formulation in the vertical direction. In the PPM, puff growth is described by relative diffusion to account for the influence of eddies smaller than the puff size to the puff expansion, while the effect of the meandering of puffs due to eddies larger than the actual puff size is simulated by introducing puff centre trajectories derived from a Lagrangian particle model. The modelling of puff meandering through the Lagrangian approach allows the simulation in inhomogeneous conditions, while grouping particles in a cluster or puff overcomes the computer consumption of a fully Lagrangian particle model. SPELL couples the effects of relative and absolute dispersion, and relates both dispersion effects to the atmospheric wind spectra. The model follows the approach in the PPM (de Hann, 1998), but is different in treating the relative dispersion. In PPM, formulae for the near field, intermediate field and far field are used, while in SPELL, the puff expansion is related directly to a model of the turbulence spectra.

Though the advanced puff-particle models above may be more appropriate for prediction of the dispersion of instantaneous releases than the simple puff model, it is still at the price of more time consumption compared with puff predictions and also the need for more detailed information about the turbulent statistics or wind spectrum.

1.3 Gaussian Puff Dispersion Coefficients

Gaussian puff models are the simplest in application and are widely used. The application of a Gaussian model ultimately requires a specification of the nature of the standard deviations of the Gaussian distribution. These dispersion coefficients describe the width of a cloud of contaminant and how the cloud grows as it moves downwind. The accuracy of a Gaussian model depends critically upon these coefficients.

A lot of effort has been put into the study of plume dispersion and many plume models are available for use today. But for the accidental instantaneous release of harmful gases, as from an explosion, or short-term releases on the order of seconds, the traditional plume models which only consider the diffusion in crosswind and vertical directions are not satisfactory for use to estimate the concentration distributions of puffs. The dispersion of a plume in the longitudinal direction is dominated by mean wind advection, while downwind diffusion and wind shear are usually ignored in plume models. For a puff formed by an instantaneous release, however, or where the release time is shorter than the dispersion time, the diffusion in the downwind direction is a result of longitudinal turbulent diffusion and the interaction of wind shear and vertical diffusion and can no longer be ignored. Wind shear even plays a major role in the total diffusion under stable conditions (Saffman, 1962). In addition, puff dispersion in the crosswind and vertical directions usually are less than those of a plume for a few minutes (Turner, 1994), so the estimation of puff dispersion by using plume models is not reliable.

The instantaneous growth of a puff is governed by the concept of relative dispersion instead of absolute dispersion as shown in Figure 1.1. According to relative dispersion (2-particle statistics), turbulent eddies smaller than the actual puff size contribute to its growth (increasing the ensemble 2-particle separation), while large eddies move the puff as a whole (without changing the separation of the particles), and increase absolute dispersion, whereas the relative dispersion remains unchanged. The instantaneous concentration profile is a reflection of relative dispersion to the centre of mass at any instant in time, while the hourly mean concentration profile is a reflection of absolute dispersion from the source.

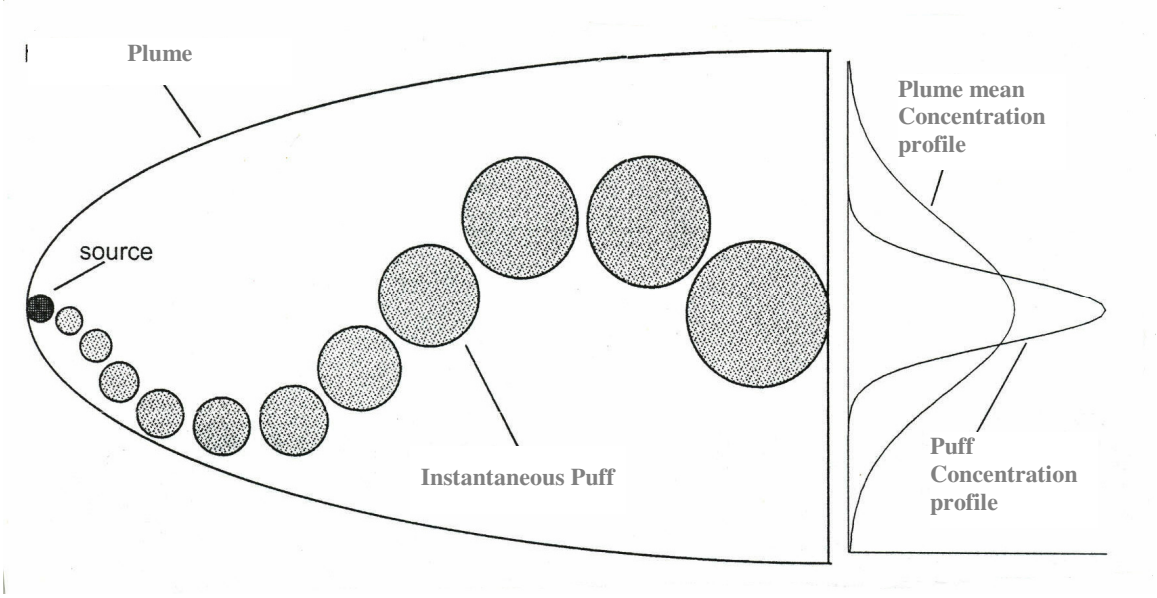


Figure 1.1 Comparison of absolute and relative dispersion (de Hann and Rotach, 1998b)

The basic expression of a Gaussian puff model is the same as expressed in Eqn. (1.3). Gaussian puff models use a moving coordinate system whose origin is at the centre of the puff, as distinct from plume models, where the coordinate system has a fixed origin (Figure 1.2). The entire coordinate system moves downwind with the puff as it is advected by the mean wind. Without loss of generality, the x axis follows the mean wind direction, i.e. downwind direction, the y and z are the crosswind and vertical directions, respectively. Thus $c(x,y,z,t)$ is pollutant concentration (g/m^3) at time t after release; Q is the mass (g) of aerosol instantaneously released at time t=0 s; three parameters, σ_x, σ_y and σ_z, are the standard deviations (m) of concentration distributions in each of the coordinate directions, or dispersion coefficients, and are used to represent the diffusion of a puff.

Directly applying the statistical theory of relative diffusion to dispersion coefficients of Gaussian puff models has been hampered by a lack of information on turbulence quantities involved in relative dispersion relations. To use the statistical

theory, standard deviations of velocity fluctuations, Lagrangian time scales or length scales in each direction and some other Lagrangian statistics have to be measured or estimated, which sometimes is impossible. The theory gives more qualitative rather than quantitative information on the dispersion coefficients.

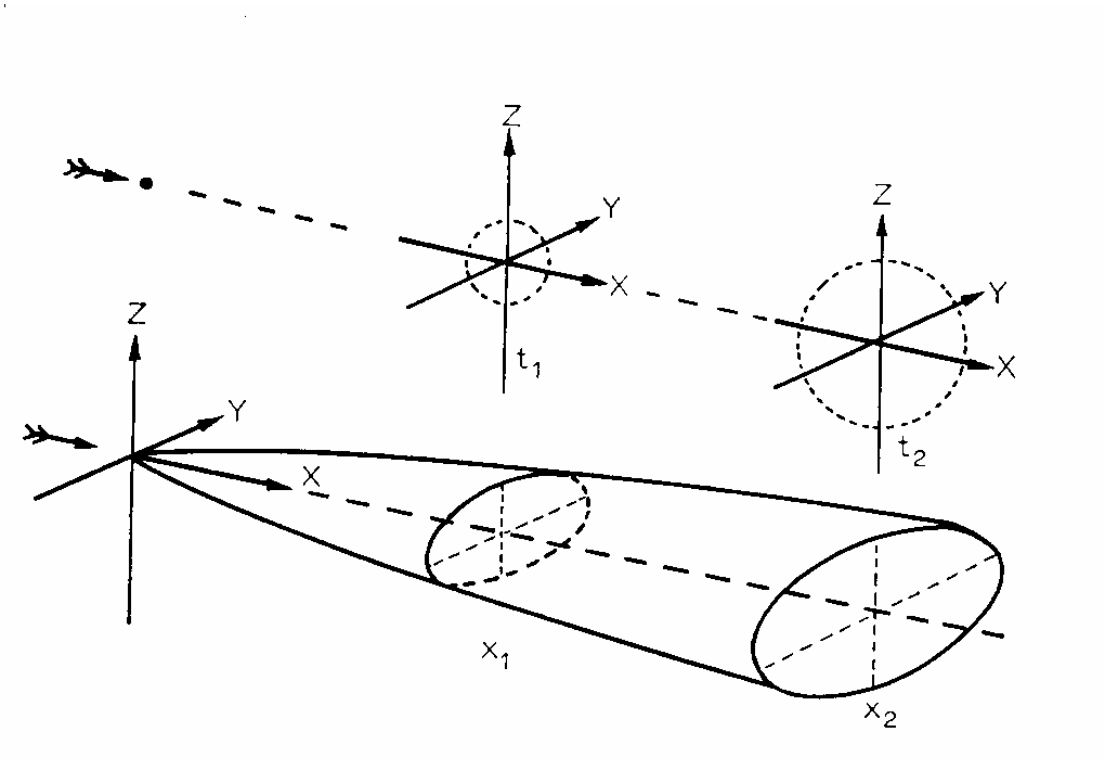


Figure 1.2 Comparison of coordinate system between Gaussian puff (top) and plume (bottom) dispersion (DeVito, 2000)

In addition, statistical theory is based on assumptions of uniform winds, homogeneous turbulence, and stationarity of the flow, which usually cannot be strictly satisfied in the atmospheric boundary layer. In a very stably stratified atmosphere, puff growth in the lateral direction is predicted as an exponential growth for the intermediate dispersion time rather than proportional to $t^{3/2}$ (Mikkelsen *et al.*, 1987).

With these difficulties of direct application of statistical theory to puff dispersion, almost all practical Gaussian models use empirical or semi-empirical dispersion

coefficients based on field experiments, and these dispersion coefficients are usually functions of atmospheric stability, dispersion distance or time (Pasquill, 1983; Slade, 1968).

Lagrangian measurement is difficult to perform in practice, as most field experiments follow the Eulerian concept that the spread of a puff is usually obtained by measuring the time-integrated or time-averaged concentration of the air tracer (Yee *et al.*, 1998; Sato, 1995; Nickola, 1971; Slade, 1968). Slade summarized a limited number of earlier puff experiments and proposed a set of power-law relations of dispersion coefficients which are functions of atmospheric stability and downwind distance from the source (Slade, 1968).

The most commonly used dispersion coefficients are Pasquill-Gifford (P-G) expressions that are functions of P-G stability (Pasquill and Smith, 1983) and distance from source. P-G dispersion coefficients were originally developed for the Gaussian plume model, but due to the lack of knowledge of puff dispersion and the difficulty of puff experiments, some Gaussian puff models directly use dispersion coefficients for plumes or modifications, and assume the longitudinal dispersion σ_x is the same as lateral dispersion σ_y (Sato, 1995). This has been found to be not true by many analyses and field experiments (Chatwin, 1968; Draxler, 1979; Van Ulden, 1992; Sato, 1995). For dispersion from a continuous source, the dispersion concerns both contributions of diffusion from small eddies and meandering caused by larger eddies, with the corresponding coordinate system using a fixed origin. The study of puff dispersion from an instantaneous release usually uses a relative frame, the coordinate system moves with the movement of cloud. Thus, the fundamental difference between absolute and relative

dispersion is that, when the dispersion is analyzed within this relative framework, ensemble averages are computed only after the puff centres which, in each realization, are first aligned. Therefore, the dispersion of a puff should be much smaller than a plume. Puff dispersions in crosswind and vertical directions usually are less than those of a plume for a few minutes (Turner, 1994).

Nickola (1971) analyzed a series of eight puff releases using the inert radioactive gas ^{85}Kr as the tracer. He studied the puff dimensions and found that $\sigma_x > \sigma_y > \sigma_z$ regardless of stability; σ_y / σ_x and σ_z / σ_x tended to decrease as stability increased; σ_z / σ_x decreased slightly with distance or time in all stabilities, while σ_y / σ_x decreased slightly in stable conditions but increased with distance or time in unstable conditions. The observations of our field experiments also show similar trends. Nickola also noticed that σ_x is greater in stable than in the less stable atmospheres when considering puff dimensions as a function of distance alone, and that stability has a minimal effect on σ_y and σ_z. However, if considering puff dimensions as functions of dispersion time, then more reasonable results were observed that puff dimensions in a stable atmosphere are smaller than they are in neutral or unstable atmospheres. Consequently, using dispersion time in developing the relationship for puff dispersion is preferable to using distance.

The main reason of $\sigma_x > \sigma_y$ is the influence of wind shear. Wind speed increases with height, as the advection of particles by the mean wind at the top of a cloud is faster than the advection of particles at the bottom of the cloud, and the shape of the concentration distribution will show skewness toward the wind direction in the top part

(Chatwin, 1968; Draxler, 1979; Van Ulden, 1992; Sato, 1995). At the same time, the vertical diffusion by vertical turbulence is trying to destroy the skewness, the stronger the vertical diffusion, the more skewness will be destroyed. It is imaginable that the concentration shows stronger skewness under stable conditions, for which less vertical diffusion exists, than under strong convective unstable conditions.

By applying moment analysis to the diffusion equation, Saffman (1962) concluded that the wind shear exists in both the downwind and crosswind directions. However, the wind shear component in the longitudinal direction is much greater than the wind shear component in the lateral direction, especially near the ground. Therefore, in most situations, only the shear effect on the longitudinal direction needs to be considered. For medium-scale diffusion, longitudinal dispersion depends on the interaction of wind shear and vertical transport. Later studies all are in agreement with this discovery, and a general expression of downwind dispersion is (Draxler, 1979, Van Ulden, 1992):

$$\sigma_x^{\,2} = \sigma_{xs}^{\,2} + \sigma_{xt}^{\,2}\,, \tag{1.4}$$

where σ_{xt} represents the downwind turbulent component, and is approximately equal to the turbulent lateral dispersion, σ_y. The shear component, σ_{xs}, is a function of the wind shear and the vertical dispersion. Chatwin (1968) described in detail the interaction between wind shear and vertical diffusion for the neutral surface layer based on the Lagrangian similarity hypothesis, and by using the method of integral moments, he arrived at exact analytical results under neutral conditions. Chatwin also concluded that the centre of mass of the cloud, under neutral conditions, rises at a uniform velocity and is carried downstream at a velocity equal to the fluid velocity at a value of z equal to the height of the centre of mass of the cloud. Chatwin also derived a coefficient proportional

to dispersion time. That said, these conclusions are only applicable to a cloud in neutral atmospheric conditions. Wilson (1981) proposed a generalized analytical formula for σ_x valid for all stabilities, based on the work of Smith (1965) and Chatwin (1968), and a reference height used to evaluate the effective wind shear was introduced.

Van Ulden (1992) extended the work by Saffman (1962) and Chatwin (1968), by using Monin-Obukhov similarity theory for the surface layer and derived semi-empirical analytical models under all conditions for puff dispersion near the ground. The major moments of the concentration distributions given by the model agreed well with the exact solutions for neutral and stable conditions. However, no experimental validation (except the lateral diffusion which was in good agreement with experimental data) was reported for Van Ulden's model.

Yee (1998) seems the only one to have performed a detailed analysis on a series of carefully controlled repeat puff releases. Before that, very few of the puff experiments involved the controlled repeat releases of an adequate number of instantaneous clouds to permit meaningful ensemble averages of various concentration statistics. Many of them just assumed that an ensemble-mean concentration could be represented by the concentration measured in an individual trial corresponding to the single instantaneous release of a cloud, and these observations were compared directly to atmospheric diffusion models which are ensemble mean value predictors only. Yee's analysis of dispersion in the crosswind direction is in good agreement with the results of a similarity model given by Chatwin (1968), while no comparisons were made in longitudinal and vertical directions, due to the lack of data. Yee's work also verified that downwind pollutant concentration distributions were negatively skewed, while crosswind

distributions were Gaussian in nature. His study is under the assumption that the centre of mass of the cloud stayed at the same height, and wind direction remained unchanged during the measuring period.

1.4 Lidar Application to Aerosol Dispersion Measurements

LIDAR, short for LIght Detection and Ranging, is a remote sensing technique that is powered by a laser source. A lidar consists of a pulse laser source, a telescope and a fast detector. The telescope is used to collect the light of the laser backscattered by air molecules and aerosols. The time for the light to travel out to the target and back to the lidar is used to determine the range to the target. Measurement of energy at the optical or infrared wavelengths permits the determination of measurable scattering, even for very small targets. Even in a "clear" atmosphere, backscatter signals from gases and suspended particles at ranges of several kilometers may readily be detected with lidar of modest performance. It is possible to measure the position of clouds or aerosols, their motions, and most importantly, their structure (Measures, 1984).

For long-range remote sensing, pulsed lidar is usually used, since it can provide much higher power levels during the laser pulse than can be maintained with a continuous-wave lidar. It also produces higher signal-to-noise ratios for the collected radiation.

The traditional and by far most common method for measuring concentration distributions of dispersing aerosols is to use an array of detectors, situated along a line or arc downwind of the source. As the aerosol passes the array, concentration-time series or dosage counts are calculated at each detector on the array (Slade, 1968; Nickola, 1971;

Sato, 1995; Yee, 1999). The lidar presents numerous advantages over such in situ measurements.

Laser beams are coherent and can be highly collimated, and optical scanning systems guide the beams to scan a given volume with a specific pattern. In this way, large volumes of space including water can be scanned from one remote position to provide a three-dimensional map of a cloud in several seconds, and no requirement of towers for elevated measurements. Compared with setting up a number of in-situ measuring instruments, a lidar system is a less expensive, time saving, reliable and safe technique. It can readily measure hazardous or inaccessible locations, including scanning over water, and elevated measurements without towers. The most important characteristic is that a lidar probe does not disturb the process being measured. Therefore, a lidar system provides better performance than in-situ measurement instrumentation.

As a well-established technique for active remote sensing of the atmosphere, lidar systems have been used for remote sensing to track smoke plumes from smoke stacks, assess the obscuration effectiveness of military screening smokes, and monitor the cloud microphysical and optical parameters (Bennett *et al.*, 1992, 1995; Roy, 1993; Mikkelsen et. al., 1995; Jorgensen, et. al., 1997; Bissonnette *et al.*, 2002). Lidar was also used to evaluate dispersion model parameters (Middleton, 2005). Some attempts have been made to model the cloud dispersion or concentration probability density function (PDF) based on data of lidar measurements (Andrews *et al.*, 1996, 1997, 1998, 2003; Devito, 2000; Luhar *et al.*, 2001, Munro *et al.*, 2003).

However, lidar measurements need a process of lidar inversion, i.e., a conversion of the measured backscattered power signal to a measure of concentration, with several

lidar inversion methods being available (Klett, 1981, 1985; Wei, 2001; Kovalev, 2003). These inversion techniques require many assumptions that are not always valid and may cause errors (Zuev, 1983; Roy, 1993), which limit their applications in the scanning of aerosol clouds. In order to avoid errors caused by invalid assumptions in the lidar inversion and calibration of a lidar, a new method has been developed by Bissonnette and Roy (2003) and is adopted for this study This inversion technique deals with the backscattered signal by relating cloud concentrations to the backscattered signal ratio of the cloud to the atmospheric background, which assumes that the background aerosol distribution is homogeneous, and the measurements are limited to cases where particle diameter is equal or larger than the light wavelength, and light scattering is dominated by Mie scattering. Furthermore, multiple scattering effects were avoided, e.g., measurements during severe fog or rain or icy conditions were discarded. This lidar inversion technique will be described in the next chapter.

1.5 Artificial Neural Network Modelling

Artificial neural networks (ANNs) are considered a branch of artificial intelligence and found their origins in psychologists' attempts at modeling synaptic neurotransmission (Patterson, 1996). ANNs provide non-linear vector mappings, with the network itself being an interconnection of nodes or processing elements that are arranged in layers. There are usually three types of layers, input, hidden and output, as shown in Figure 1.3. The input layer contains a node for each independent variable, with each data point describing a unique input vector. Each input vector is associated with an output vector, the dependent variable(s) of interest. There may be more than one hidden layer,

depending on the complexity of the system being modelled, although one or two hidden layers are enough to address most problems (Patterson, 1996).

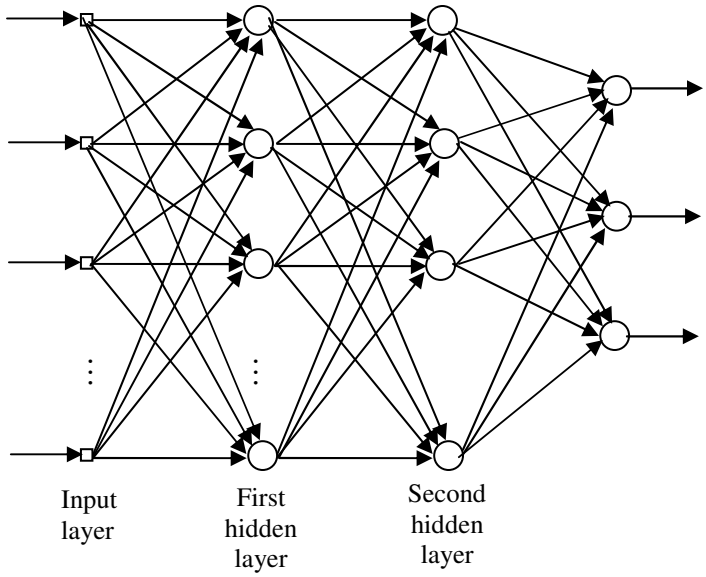


Figure 1.3 A general ANN with two hidden layers

At the start of the training process, the connecting weights throughout the network are assigned random values. The output vector generated by the model is compared to the corresponding measured output vector. The difference between predicted and measured outputs is the global error and that is propagated backward through the network. An iterative process of adjusting the connecting weights to minimize the global error is carried out. Once the global error has been minimized over the whole training set, the weights can be fixed and the network can be used to make blind predictions. A detailed description of the theory involved with ANN training can be found in Andrews *et al.* (1999).

ANN models have numerous favorable characteristics for modelling complex vector functions. They can provide multi-dimensional non-linear correlations; they are capable of reproducing synergistic effects among input variables; the input variables need

not be linearly independent; ANNs are particularly good at handling noisy and incomplete data; they provide a smoothing function; and they can estimate (through examination of the weight space or by sensitivity analysis) the relative significance of the input variables to the constructed model (Patterson, 1996; NeuralWare, 2003).

Some investigators have reported the application of neutral networks to the prediction of aerosol concentration in the atmosphere (Andrews *et al.*, 1996, 1997, 1998, 2003; Gardner and Dorling, 1999; Devito *et al.*, 2000; Cappa *et al.*, 2001; Guardani *et al.*, 2004). Gardner et al (1999) constructed several multi-layer perceptron (MLP) ANN models to predict the hourly average nitrogen oxide (NO_x=NO+NO_2) and nitrogen dioxide ($NO_{2)}$ concentrations at a specific site in the UK. The inputs of the ANN model were hourly averaged meteorological parameters, including low cloud amount, base of lowest cloud, visibility, dry bulb temperature, vapour pressure and wind speed, and took the hourly average nitrogen oxide or nitrogen as the output. The ANN models showed considerably better results than traditional multiple linear regression models. Cappa et al (2001) applied the back-propagation (BP) ANN technique to the 3-hour and 24-hour forecasting of NO_2 concentrations in Turin, Italy, providing good agreement with later observations. Similarly, Guardani and Nascimento (2004) used an ANN model for predicting ground-level ozone concentrations in large urban areas, and their model provided very good correlations between predictions and observations. They also found that ANNs were better at predicting average concentrations than absolute maximum concentrations, which were generally underestimated.

Each of these studies addressed either ambient atmospheric aerosols, or continuously released contaminants. DeVito et al. (2000) and Andrews et al. (2003) tried

to use neural network modelling techniques to predict the dispersion of clouds from instantaneous sources. DeVito et al. (2000) developed ANN models based on meteorological data corresponding to P-G stability categories A and B, i.e., from very unstable to moderately unstable conditions. The developed ANN models were compared with traditional Gaussian puff models using Slade's parameterization and COMBIC (the US Army Research Laboratory's Gaussian-based dispersion code, a Gaussian puff model using its own dispersion parameterization scheme) and showed significantly better correlations, smaller mean biases and far less variance than exhibited by the other two Gaussian puff models. Andrews et al. (2003) reported on a general ANN model which addresses puff dispersion under all possible conditions, and showed some promising results.

1.6 Objective of the Thesis

The challenge of puff dispersion modeling is (1) no sound theoretical support exists; (2) it is difficult to collect whole cloud information by traditional measuring instruments; (3) so far no one general model exists to address puff dispersion under all meteorological conditions.

Today, with the availability of lidar in tracking and measuring the whole cloud puff, providing information about time, travel distance, cloud azimuth, elevation, and signal intensity, it is possible to get the centre of mass of a whole cloud and track its movement to check and study the cloud dispersion nearly instantaneously, and give a better understanding of the whole cloud dispersion. Consequently, 4 one-week-long field trials were conducted, with a wide range of data being collected to avoid seasonal and diurnal biases.

One of the most important characteristics of the neural network modelling technique is that they can handle noisy and redundant inputs, and have the ability to build up multi-dimensional non-linear correlations. Thus, one of the goals of the research reported in this thesis is to develop a general model which can be used to predict puff dispersion under different conditions but that would have a relatively simple structure.

Most puff models use only one parameter, e.g., P-G stability category, to represent the atmospheric stability condition or turbulence intensity. The P-G stability classification scheme is simple but with several limitations, the most severe being that the P-G scheme uses only 6 or 7 discrete classes, each of which covers a rather wide range of stability and turbulence conditions. In addition, the correspondence between Pasquill's stability classes and turbulence is not unique but strongly depends on the surface roughness, the planetary boundary layer (PBL) depth and possibly other parameters. Pasquill's stability categories and the associated dispersion parameterization were proposed for a rather small roughness. The same classes, when applied to significantly rougher or smoother surfaces, could represent widely different turbulence and diffusion characteristics. Therefore, turbulence and diffusion parameters cannot be unambiguously specified in terms of these qualitative stability categories alone (Arya, 1999). The ANN model will use several atmosphere stability-related parameters including gradient of temperature and humidity and turbulence fluxes to represent the continuous varying turbulence of the atmosphere. By this way, discreteness and ambiguousness of air stability by using P-G stability only, which is adopted by most puff models, will be avoided, and more real situation will be reflected.

As discussed before, the concentration distribution of a puff in the longitudinal direction consists of not only the diffusion caused by longitudinal turbulence but also the interaction of centroid deviation caused by wind shear and vertical diffusion. A more (under stable conditions) or less (under unstable conditions) skewed concentration distribution will be expected in the sense that the top of the cloud is further downwind than the bottom. In fact, most observations show that the concentration distributions of a puff in the downwind direction are skewed non-Gaussian distributions, and for releases from or close to the ground with no deposition (total reflection at the ground), the vertical concentration distribution is non-Gaussian as well (Van Ulden, 1992).

However, due to simplicity, Gaussian puff models are still in wide use in regulatory purpose. The use of appropriate dispersion coefficients is vital to the accuracy of the model prediction. Theoretical studies make many assumptions that may not be satisfied in practice, and semi-empirical parameterizations may only be suitable for the specific conditions being studied. The second main goal of this research is to develop an ANN model to predict the dispersion coefficients for the wide range of data available. No predetermination of the stability category is needed, reflecting reality.

Aerosol puffs move in the direction corresponding to the mean wind velocity, and while moving, the puffs disperse under the influence of turbulent diffusion. When a moving puff is hit by a large-scale gust, the puff may change direction and it is also possible that the puff will be broken into two or even more parts. Some allowance will be made for puff trajectory, but it will not be a major concern in this study. The emphasis of the study is on small-scale turbulence and its influence on the expansion of puff size, with the puff following the trajectory of the centre of mass. The dispersion time involved

here is less than 150 seconds. With this short-range consideration, it is assumed that the statistical characteristics of the small-scale atmospheric turbulence are unchanged.

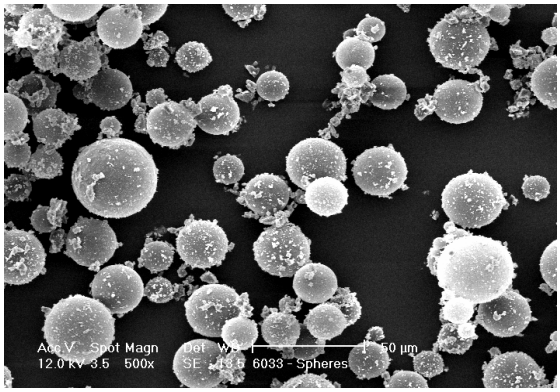
Chapter 2
Data Collection

In order to enhance understanding of non-buoyant aerosol dispersion from transient point sources into the atmosphere at ground level, field trials were conducted over flat terrain at DRDC-Valcartier intermittently from 2002 to 2003, to examine the influence of a variety of atmospheric and meteorological factors on aerosol dispersion in the surface boundary layer over flat terrain. Factors considered included wind velocity, air temperature and pressure, along with their associated gradients, humidity, ground heat flux and insolation, as well as diurnal and seasonal variations. The influence of derived turbulence statistics was also examined. The application of lidar to cloud scanning made it possible to get useful information of an entire puff cloud transient instantaneously. This four-season experiment was aimed at removing seasonal biases and covering as broad a range of atmospheric and meteorological variables and parameters as possible.

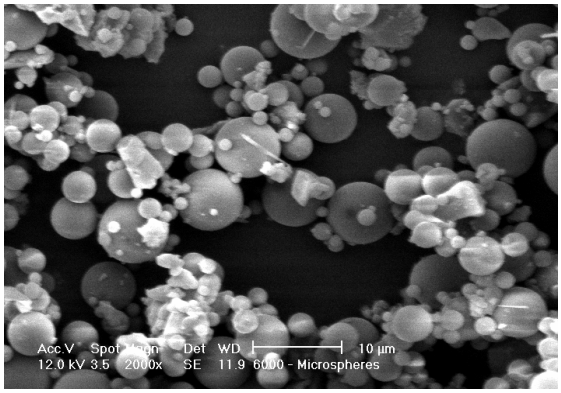
2.1 Experiment Set-up

Trials were conducted at the Canadian Forces Base Valcartier, Quebec. The trial location was a 2 km^2 level plain with high hills bordering the northern edge. The surface was either covered with snow in November and March, or covered with low brush in May and August. Artificially generated aerosol puffs were used during all experiments, with the tracers being microspheres made of PVC (~ 27 μ m, 0.2 g/cm^3), glass spheres (~ 6.5 μ m) and talc powders, where the PVC and glass powders consisted of spherical particles, while the talc particles were flakes, with their scanning electron microscope (SEM) images being shown in Figure 2.1. An aerosol generator was used to generate a

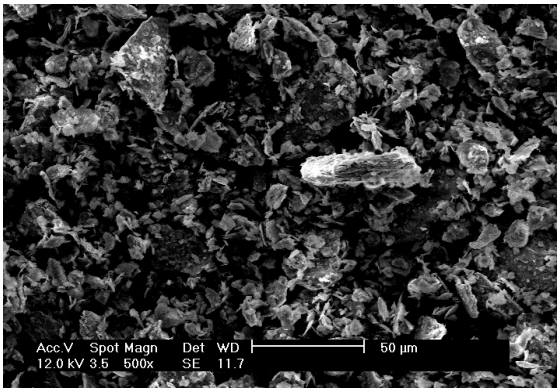
transient instantaneous release of a puff for three different aerosols individually in each trial, and it took about 2 s to finish a release. The formed cloud puff was scanned and detected by a lidar which was generally located about 200 m away from the release point. At the same time, the meteorological system, located a few meters away from the lidar, measured regular micrometeorological data such as the three components of speed, wind direction, temperature, relative humidity and their gradients, air pressure, insolation etc.



(a) microspheres of PVC



(b) glass spheres



(c) talc flakes

Figure 2.1 SEM of aerosols released

Figure 2.2 is the field layout of experimental instruments, where D_i is the possible release point depending on wind direction at the time of release. R1 is the location of the head of the scanning lidar, LCM (Lidar Scanning Mapper), and SD_1 and

SD_2 are the limits of a sand dune or berm. The digit in brackets in Figure 2.2 is the distance to the location from the LCM.

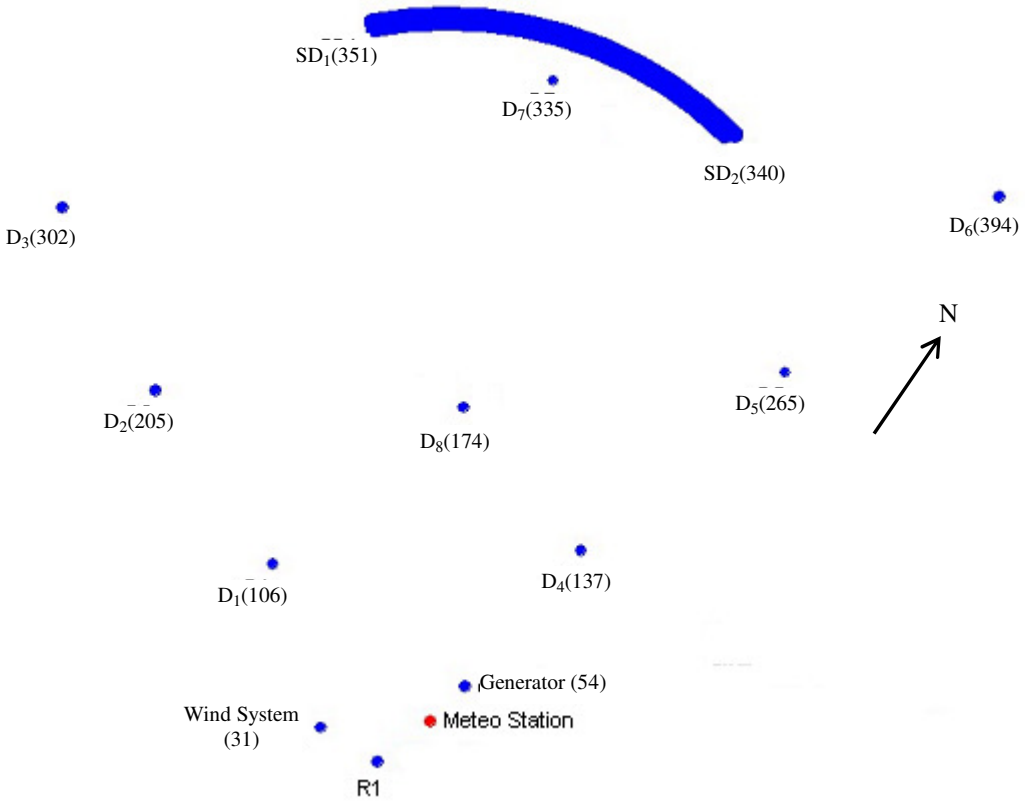


Figure 2.2 Layout of field experiment

The lidar used for the field experiments consisted of a pulsed Nd:YAG laser source with an energy of 30 mJ, operated at a wavelength of 1.064 μm and a frequency doubled Nd:YAG laser source of 20 mJ which was operated at a wavelength of 532 nm, a scanning platform, secondary optics and a receiver. The laser source emitted 12 ns pulses at a pulse repetition frequency of 100 Hz. The laser shots were directed along different lines of sight by motor-driven mirrors that scanned both in azimuth and elevation to give a raster pattern. The collecting optical system was a telescope of 200 mm diameter with a 0.762 m focal length. During lidar scanning, four channels of data could be obtained (two

channels for the 1.064 μm laser source, and two channels for the 0.532 μm laser source). However, only one channel operated by the 532 nm source was used during experiments to collect all the backscatterred energy. The signals received were amplified by logarithmic amplifiers. After the returned signals were amplified, they were digitized with 8 bits at a rate of 125 MHz and thus gave a spatial resolution of 1.2 m. The results were transferred to a computer and saved. The beam divergence was set at 2 mrad, and the field of view was 4 mrad.

The lidar was housed in a trailer, located on the southern edge of the plain, and the scanning head of the lidar was mounted on the roof of the trailer (Figure 2.3), about 4.2 m above ground level. The principle axis of the tests was in the direction of magnetic north (declination about 18 deg. toward west) and was centred on the location of the lidar.

Figure 2.3 View of scanning head of the lidar (LCM) and its support trailer

The lidar scanned the puff with a raster scanning mode. It started from a preset elevation and azimuth, and scanned the puff sweep by sweep. After the mirror completed a sweep of 82° of azimuth, it was elevated by a controlled motor and then it executed another 82° sweep in the opposite direction. Each sweep was thus a planar slice through the cloud. This process continued until the desired number of horizontal sweeps was reached. Then the mirror was returned to the starting position ready for the next scan. Each trial went through 5-8 scans, and each scan usually contained 8-10 sweeps. A limited number of scans contained 12 or 16 sweeps for convective conditions. Each sweep contained 103 shots with 400 points sampled for each shot. It took about 1 s to finish a sweep, and about 15 s for a scan containing 8 sweeps. The higher the number of scans, the more detailed was the dispersion information with time. Also a greater number of sweeps provided more vertical dispersion information. Figure 2.4 shows the raster scanning pattern of the lidar.

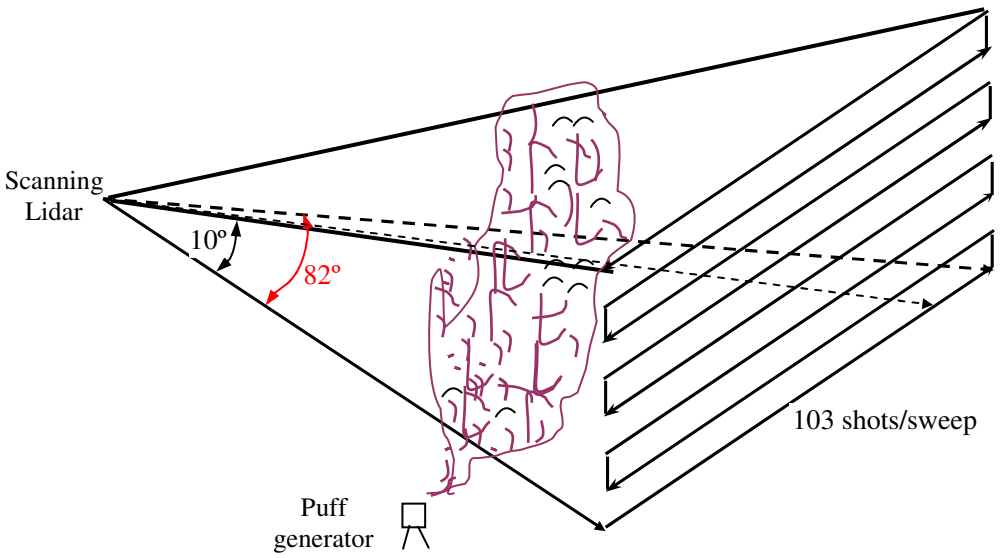


Figure 2.4 Raster scanning pattern of lidar

2.2 Meteorological Conditions

The meteorological data during each trial period were collected by a basic meteorological system. In addition, a sonic anemometer was also used to provide data on the atmospheric turbulence.

The meteorological system consisted of a number of sensors, given in Table 2.1. A data acquisition box was linked to a remote data acquisition computer, housed in the lidar support trailer. The first two pieces of equipment provided insolation and ground heat flux measurements. The third piece of equipment was a wind vane that could determine both the wind speed and direction. Air temperature and relative humidity were measured at two different heights, with one being close to the ground and the other near the maximum height of the sensor pole. Finally, a barometer was located near the data acquisition box, and a rain/snow (liquid water) gauge was located nearby on the ground. Data from all these sensors, except the rain/snow gauge, were sampled every 10 seconds. These data, along with 1-minute averages and their standard deviations, were saved to a data file. The outputs from the rain/snow gauge were logged by a HOBO^{TM} data logger and read out at the end of each trial period. The sonic anemometer was normally operated to output 4096 measurements of the sonic temperature, and three orthogonal wind speeds at a rate of 20 Hz every 15 minutes. This corresponds to a sampling period of 3-4 minutes every 15 minutes. These sonic data were subsequently analyzed by using the meteorological program METCLC which was developed at DRDC-Valcartier (Forand, 2003) to extract many meteorological (mean) and micro-meteorological (turbulence) parameters.

Table 2.1 Characteristics of the meteorological system during the trial periods

Equipment					Height (m); Trial Period			
Sensor	**Manufacture**	**Model**	**Precision**	**Units**	**1**	**2**	**3**	**4**
Insolation	Qualimetrics	3120	5	%	1.88	1.90	1.90	1.93
Ground heat flux	Qualimetrics	3120	5	%	1.73	1.70	1.75	1.72
Wind speed	Young	O5103	2	m/s	1.81	2.27	2.26	2.23
Wind direction	Young	O5103	5	deg.	1.81	2.27	2.26	2.23
(upper)Air temperature	Young	41372VC	0.2	˚C	2.12	2.08	2.13	2.07
(upper) Rel. Humidity	Young	41372VC	2	%	2.12	2.08	2.13	2.07
(lower)Air temperature	Vaisala	HMP45A	0.2	˚C	0.63	0.54	0.63	0.51
(lower) Rel. Humidity	Vaisala	HMP45A	2	%	0.63	0.54	0.63	0.51
Barometer	Qualimetrics	7105-A	1	hPa	0.75	0.75	0.75	0.50
Rain/Snow Gauge	Weather Measure Corp.	P511-E	0.5	%	0.00	0.00	0.00	0.00
Sonic Anemometer	Metek	USA-1			2.57	2.58	2.60	2.56
		Speed	0.05	m/s				
		Dir.	0.4	deg.				
		Temp.	0.01	K				
		Wind Component	0.02	m/s				

The details of statistics of the meteorological data, like wind speed, wind direction, air temperature, relative humidity, and insolation during each period can be found in a DRDC-Valcartier report by Forand (2004).

Meteorological conditions during the four trial periods varied greatly. The range of the basic meteorological conditions is summarized in Table 2.2.

Table 2.2 Range of meteorological conditions for field trial periods

Exp. Date	Max Wind Speed (m/s)	Tmax (˚C)	Tmin (˚C)	RH (%)	Max Insolation (W/m^2)
Nov. 2002	10.5	1	-20	41-100	463
Mar. 2003	8.4	-1	-30	24-98	749
May 2003	4.9	23	4	29-100	1155
Aug. 2003	7.2	28	10	25-100	1022

2.3 Lidar Inversion

With the advent of the use of lidar systems to probe the atmosphere in the past three decades, numerous techniques have been developed to extract useful optical properties from backscattered lidar returns.

The basic monostatic lidar equation assuming no multiple scattering is:

$$P(R) = P_0 \frac{ct_p}{2} \frac{A}{R^2} F(R)\beta(R)\exp(-2\int_0^R \alpha(R')\mathrm{d}R', \qquad (2.1)$$

where $P(R)$ (J) is the power of the backscattered signal from distance R, P_0 (W) is the initial laser pulse power, c is the speed of light, t_p (μs) is the duration of the laser pulse, $F(R)$ is a system specific function describing the overlap between the laser beam and the collecting optics and is usually considered to be a constant, $\beta(R)$ (1/msr) is the volume backscattering coefficient, R (m) is the distance or range through which the laser passes, as $R = ct/2$ (t is the time of flight, thus 1 μs corresponds to a distance of 150 m), A (m^2) is the telescope receiver area and $\alpha(R)$ (1/m) is the volume extinction coefficient (Roy, 1993).

Lidar inversion is used to solve Eqn. (2.1) to determine the relationship between the extinction coefficient $\alpha(R)$ and the received backscatter signal $\beta(R)$. To obtain the

solution of Eqn. (2.1), the relationship between the extinction coefficient $\alpha(R)$ and backscattering coefficient $\beta(R)$ has to be predetermined and usually a linear relationship is taken which may not be always satisfied (Roy, 1993).

A review of common inversion algorithms in use can be found in Evans (1988) and Bissonnette (1996), with the latter outlining the major problems remaining in most such inversion attempts. Bissonnette concludes that the chief problems are: the need to determine a relationship between aerosol backscatter and extinction coefficients, the common requirement to specify a boundary value at some specified range, instabilities or slow convergence of the solutions, and the need to properly account for multiple-scattering events.

Klett (1981) developed a stable inversion algorithm that assumes a power-law relation between the backscatter and extinction coefficients. This method is based on the well-known but unstable 'forward method', but requires the measurement or estimate of the extinction coefficient at some range beyond the extent of the cloud, rather than in front of it. This results in an inversion procedure that is more stable with respect to perturbations in the signal, the postulated relationship between the backscatter and extinction coefficients, and to the estimate of the boundary condition. Some suggestions concerning how to make the boundary condition estimate are given, but in practice this can be quite difficult, and these estimates become less valid as the optical depth of the cloud becomes small. In such cases, convergence is slow, and generally, only the front portion of the returned signal is of use. Klett modified this algorithm to account for deviations in the relationship between the backscatter and extinction coefficients (Klett, 1985), and showed that stable solutions could be obtained, but the basic problems of the

original method remain. Multiple scattering effects are not accounted for in Klett's formulations.

Evans (1984, 1988) developed a stable inversion algorithm that does not require the estimation of the extinction coefficient boundary condition. Each lidar return is instead divided by a clear air calibration shot, which greatly reduces the effects of system noise, enables the detection of very weak signals, and provides stable solutions for a wide range of optical depth (Evans, 1984). This method also assumes a power law relation between the backscatter and extinction coefficients, but multiple scattering effects can be accounted for to some extent directly in the algorithm. This algorithm is fairly general in the sense that it is not specific to a given aerosol. However, the numerical compensation is specific to the lidar system.

Roy *et al.* (1993) proposed a lidar inversion technique based on total integrated backscatter (TIB) calibrated curves. The TIB calibration curves obtained are general in the sense that it is not necessary to use the lidar equation to perform the inversion, and a linear relationship is not assumed between the backscattering coefficient and the extinction coefficient. However the TIB calibration curves are specific to the lidar system that has been used to obtain the calibration curves and to the aerosol material and size distribution.

Bissonnette and Roy (2003) proposed a relatively simple but effective lidar inversion technique which avoided the necessity to calibrate the lidar continuously and avoided the unreasonable assumption in generally solving the lidar equation. The basic principle of this inversion technique is described below.

The lidar received backscattered signal is composed of the signal backscattered from all the atmospheric constituents in the path: air molecules, background aerosols, as well as particles in the artificially generated aerosol puff. Taking the signal of the air molecules and background aerosols as atmospheric background signals, the backscattered signal of the released aerosol is superimposed on the backscattered atmospheric background signal. It is, thus, very easy to determine the borders of the cloud along the direction of each line of sight. A broad range is selected for the cloud and checking point by point, if a point signal intensity is twice as large as the average of the previous adjacent 3 points, then the distance from the lidar to this point is the nearest boundary (R_b) of the cloud along the line of sight. The same procedure is applied for the farthest boundary (R_e) determination (Bissonnette and Roy, 2003). The nearest and farthest distance R_b and R_e along a line of sight (shot) is shown in Figure 2.5.

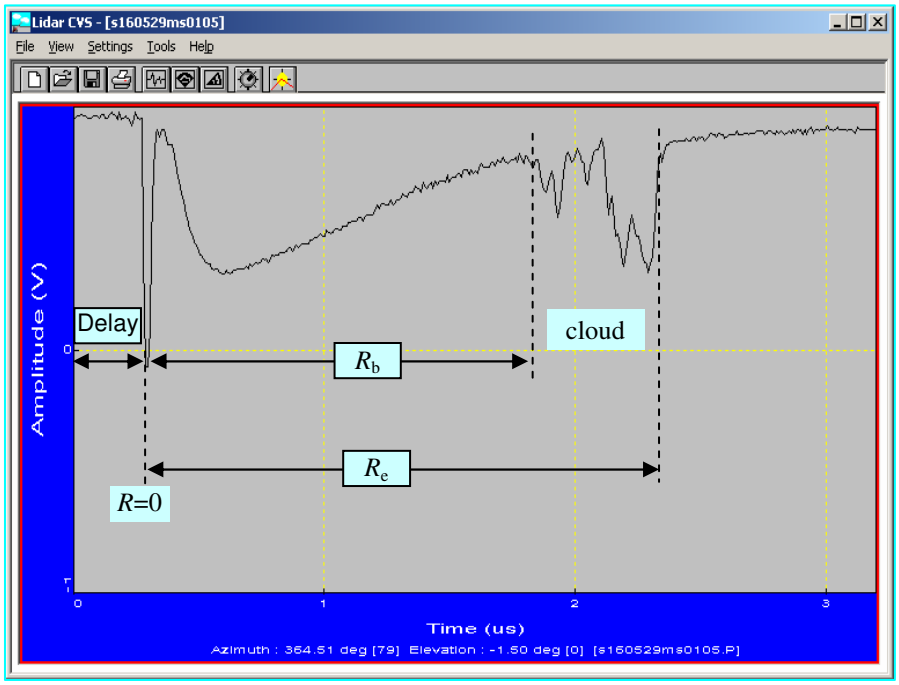


Figure 2.5 Signal view of cloud and boundary determination

When the lidar system is specified, Eqn. (2.1) can be simply expressed as:

$$P(R) = K\frac{\beta(R)}{R^2}\exp(-2\tau), \tag{2.2}$$

where $K = \frac{1}{2}P_0 c t_p A F(R)$ is a constant for a specific lidar system, and $\tau = \int_0^R \sigma(R')\mathrm{d}R'$ is the optical depth. Considering β_a as the atmospheric backscattering coefficient, β_i as the puff aerosol backscattering coefficient, P_b and P_e as the backscattering power received from points at distances R_b and R_e respectively, τ_0 as the optical thickness in the range of [0, R_b] , τ_a' and τ_i' as the optical thickness of the atmosphere and puff aerosols individually in the range of (R_b, R_e), and τ_a and τ_i as the optical thickness of the atmosphere and puff aerosols at the farthest boundary individually along the line of sight, then

$$P_b = P(R_b) = K\frac{\beta_a}{R_b^2}\exp(-2\tau_0), \qquad \text{when } R = R_b \tag{2.3}$$

$$P(R) = K\frac{[\beta_a + \beta_i(R)]}{R^2}\exp(-2\tau_0 - 2\tau_a' - 2\tau_i'), \text{when } R_b < R < R_e \tag{2.4}$$

$$P_e = P(R_e) = K\frac{\beta_a}{R_e^2}\exp(-2\tau_0 - 2\tau_a - 2\tau_i), \qquad \text{when } R = R_e. \tag{2.5}$$

Assuming the background is uniform within the cloud range, then

β_a = Constant in the interval (R_b, R_e).

By Eqn (2.3) and (2.4), it can be derived that

$$\frac{\beta_i(R)}{\beta_a} = \frac{P(R)}{P_b}\frac{R^2}{R_b^2}\exp(2\tau_a' + 2\tau_i') - 1, \text{ when } R_b \le R \le R_e, \tag{2.6}$$

and $\beta_i = 0$ when it is outside the cloud range.

It is well checked during the experimental periods that $\tau_a' + \tau_i' << 1$, thus

$$\exp(2\tau_a' + 2\tau_i') \cong 1 + 2(\tau_a' + \tau_i') + \cdots \cong 1 + a\frac{R - R_b}{R_e - R_b} + \cdots . \quad (2.7)$$

By Eqn (2.3) and (2.5), the following expression can be derived

$$\exp(2\tau_a' + 2\tau_i') = \frac{P_b R_b^{\,2}}{P_e R_e^{\,2}} . \quad (2.8)$$

Therefore

$$a \cong \frac{P_b R_b^{\,2}}{P_e R_e^{\,2}} - 1 . \quad (2.9)$$

Finally, by using Eqn.(2.6), (2.7) and (2.9), the ratio of β_i / β_a can be obtained as

$$\frac{\beta_i}{\beta_a} = \frac{P(R)}{P_b}\frac{R^2}{R_b^{\,2}}\left\{1 + \left(\frac{P_b R_b^{\,2}}{P_e R_e^{\,2}} - 1\right)\frac{(R - R_b)}{(R_e - R_b)}\right\} - 1, \text{when } R_b \le R \le R_e . \quad (2.10)$$

The relationship represented by Eqn (2.10) enables the characterization of the profile of the aerosol cloud in the direction of each line of sight, according to the relative lidar signal $P(R)/P(R_b)$. It is thus not necessary to calibrate the lidar system.

The absolute concentration of puff particle (c_i) can also be obtained by the lidar inversion provided by Eqn. (2.10) if more parameters are available. Bissonnette and Roy (2003) demonstrated that, for a non-absorbing aerosol cloud, when the size of the cloud particles is larger than the lidar wavelength, the relationship between β_i and c_i is approximately expressed as

$$c_i(R) = k\frac{\beta_i(R)}{\beta_a}, \quad (2.11)$$

and

$$k = 2.6\frac{k_a}{k_i}\frac{\rho_i r_{ie}}{V}, \quad (2.12)$$

where k_a and k_i are the lidar ratios (backscatter to extinction) of the atmospheric background and the aerosols being examined, respectively, ρ_i is the density of the aerosols r_{ie} is the mean size of the aerosols, and V is the visibility(Bissonnette and Roy, 2003). For short dispersion times for a cloud puff, these parameters are approximately unchanged and therefore k is an approximate constant during measurement. But for cloud particles under different conditions, k might be different.

For modelling purposes and for convenience to compare with other dispersion models, it is better to use the absolute concentration of puff points. Unfortunately, not all values of those necessary parameters were available. In order to make measurements under different conditions comparable, concentration normalized by the densest value of the initial cloud was used for longer term dispersion modelling.

2.4 Lidar Data Presentation

Lidar provides 4-dimensional measurements of the cloud, spatially and temporally. Data collected by the lidar can be viewed by three scan images, i.e., images of horizontal and vertical slices of the cloud, and images of the intensity along each shot. These images can help roughly estimate where the cloud location is, how it has moved and how much bottom and top have been missed.

The view of inverted lidar signals along a shot has been shown in Figure 2.5. The first peak is caused by the scattering of the laser pulse on the scanner optics. The signal digitizer and the laser are trigged at the same time and the delay is caused by the build-up time of the laser pulse in the laser cavity. This delay/distance needs to be removed in order to calculate the actual distance to target. At short ranges, the telescope does not

"see" the laser beam. As the beam travels away from the lidar, more and more of the laser beam is "seen" by the telescope until, near the peak of the signal, the entire beam is inside the telescope's field of view. At long ranges, the signal falls off as $1/r^2$, as implied in the lidar Eqn. (2.1). Signals before R_b and after R_e are called clear air or background signals, and are important in the lidar inversion.

Figures 2.6 and 2.7 are elevation and azimuth views of a cloud, respectively. The elevation view (all lines of sight are at the same elevation) provides a horizontal slice view of the cloud, while the azimuth view (all lines of sight are at same azimuth) shows a vertical cloud slice. The aerosol cloud can be seen as a small green area in the centre of the sector in Figure 2.6 and roughly 1/2 of the way along from the apex in Figure 2.7.

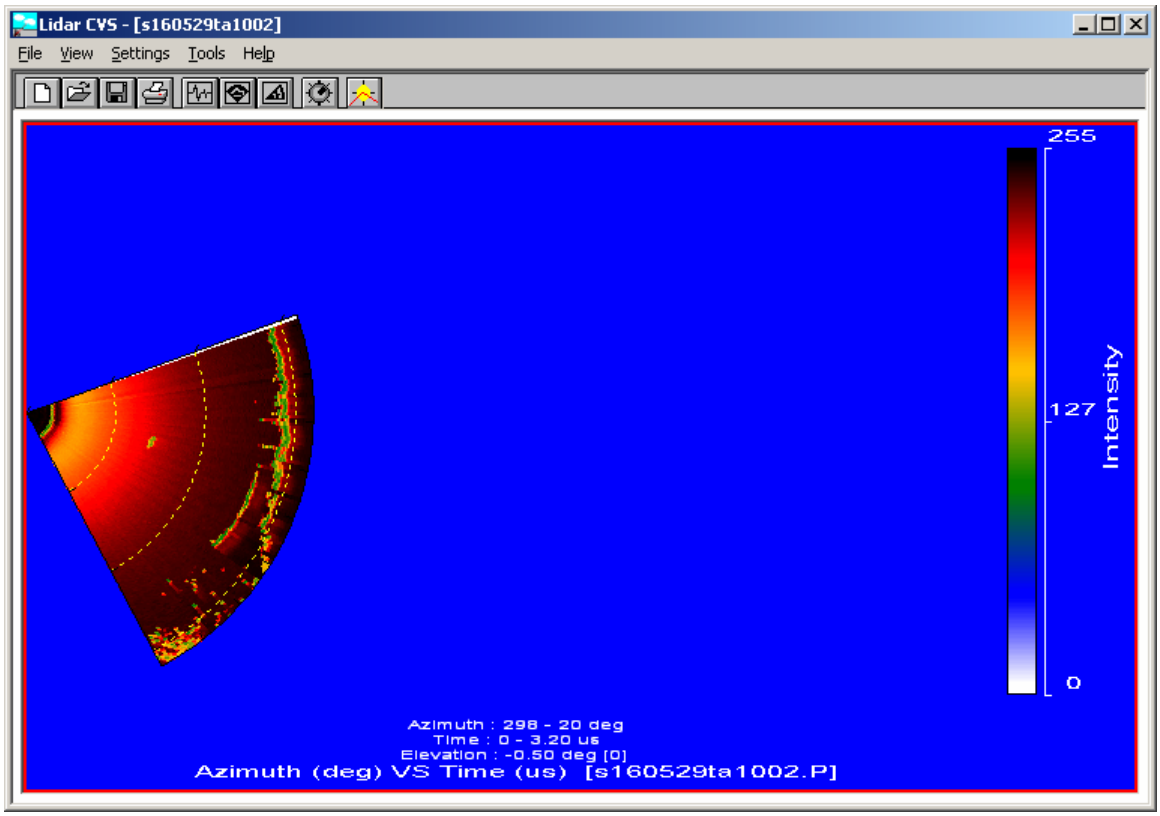


Figure 2.6 Elevation view of cloud (downward looking at a horizontal projection)

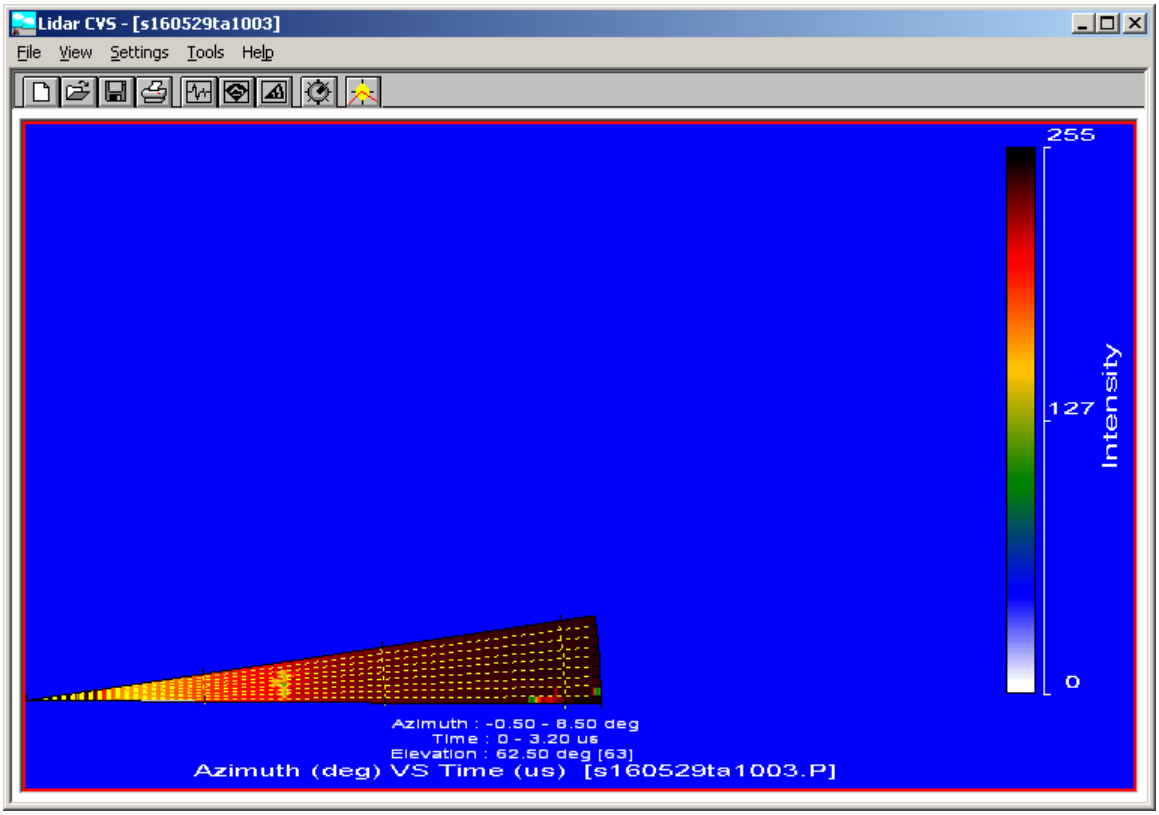


Figure 2.7 Azimuth view of cloud (downward looking at a vertical projection)

The arc in Figure 2.6 is the lidar signal of a sand dune or berm, located about 350 m north of the LCM. The sand dune in the cloud map is an important reference in correcting the lidar records of elevation and azimuth which were vital in coordinate transformation. The elevation and azimuth views can help one to figure out if the whole cloud was captured and if not, how much was missed.

It took about 15 s to finish a total scan of a cloud with 8 sweeps. The elevation interval was set at 0.5˚ in stable conditions and 1.0˚ in unstable conditions, to limit the number of sweeps without losing key information of the cloud evolution. A small number of sweeps is preferred, to make sure a full scan of the cloud can be finished in a relatively short time period so that the image can be taken as a quasi-instantaneous image of the cloud. With data from a full scan, the centre of mass of the cloud can be calculated by

using first-order moment analysis. A frame of relative diffusion, or relative distance to the centre of mass is used for later modelling under the assumption that the wind velocity was unchanged over a short time interval and only dispersion caused by small eddies is considered.

There are some limitations to the use of the lidar data at larger downwind distances. For example, sometimes the contrast between the puff and the background aerosol fell below the lidar's detection threshold, making it difficult to define the full size of the puff, especially when the cloud was getting large and had low concentrations. Also, on some occasions (when the wind was strong or under convective conditions), the lidar did not sample the complete extent of the puff. The lidar sampling window was limited and when the cloud was moving quickly, it could leave the lidar's sampling window. Consequently, these data were excluded from the data set for later modelling, leaving fewer data available for very unstable and neutral (very windy) conditions. Based on the scan images, useful clouds were selected and information of relative concentration and positions and corresponding meteorological or micrometeorological records were assembled for later modelling.

2.5 Analysis of Measurements and Results

The inverted lidar scans were analyzed in the framework of relative diffusion, where the coordinate system was centred on the centre of mass of a diffusing puff, and followed the puff as it was advected downwind. This removed the effects of the large eddies on the meandering of the puff, and only the spread of the cloud about its own centre was considered.

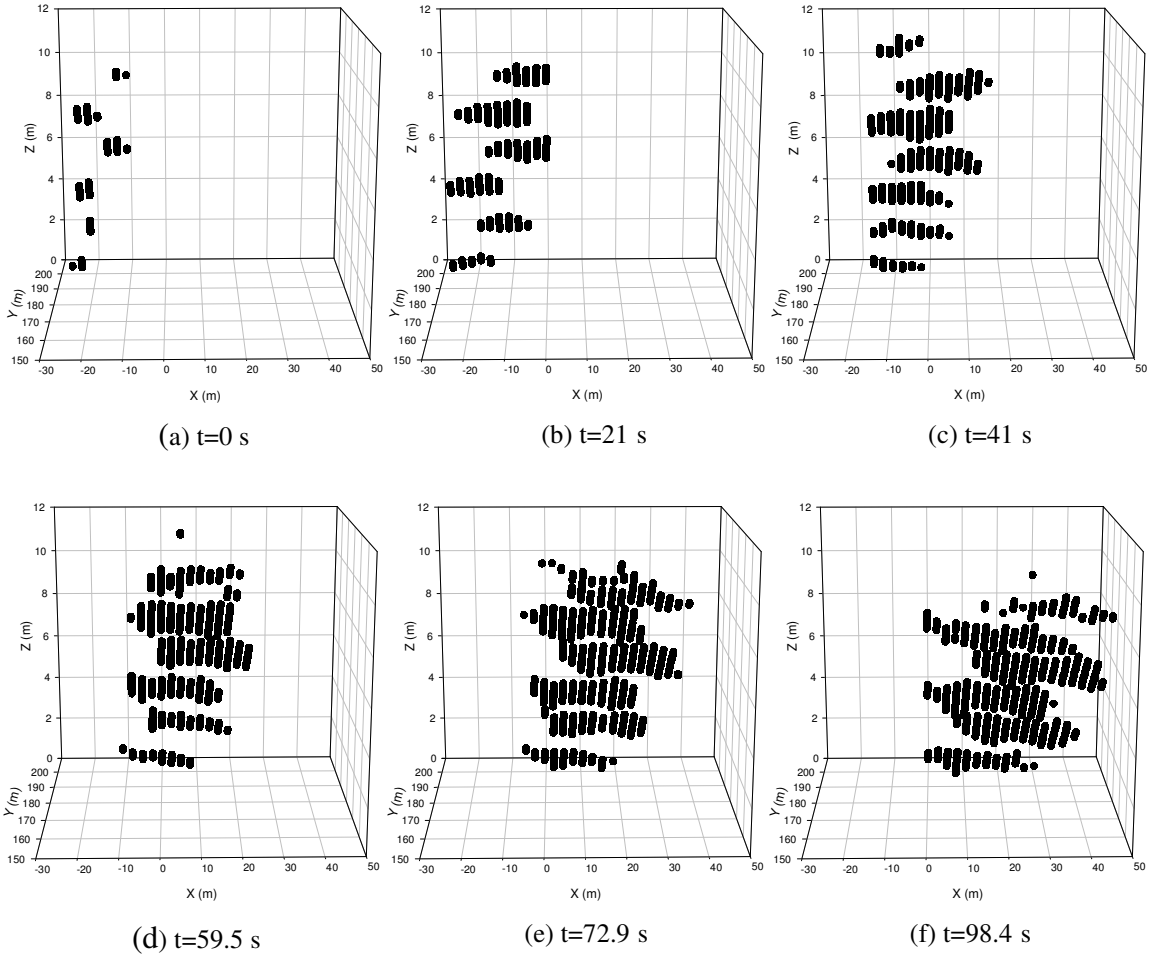
2.5.1 Cloud evolution

Puff dispersion or cloud expansion in the surface layer of the atmosphere is due to the combined effects of mechanical and convective (buoyant) atmospheric turbulence. Mechanical turbulence is produced by wind shear, in which the wind speed and direction change with height. Mechanical turbulence is dominant in the case of strong winds blowing over very rough surfaces, and in near-neutral and stable conditions. Buoyant or convective turbulence is generated by the exchange of heat between the surface and the air flow, and is dominant in the case of light or calm winds blowing over a heated surface. Puff evolution under both stable and unstable conditions was examined.

Figure 2.8 illustrates the release of talc on August 15th, 2003, with 6 scans available between 2:52 to 2:54 in the morning, with a wind speed of 0.4 m/s, a temperature of 10ºC, a relative humidity of 97.3% and an air pressure of 101.1 kPa. For this trial, the atmospheric stability was related to Pasquill stability category F, i.e., moderately stable conditions (the Pasquill stability determination followed the rules defined in COMBIC, 1995). Each scan contained 8 sweeps, but the 8th layer had no cloud points, which meant no cloud top was missed or the missing part could be ignored. Because of the systematic shift of the lidar from sweep to sweep, usually there was a 2 or 3 shot difference in the azimuth from sweep to sweep, which had to be corrected manually, trial by trial.

In Figure 2.8, x and y are the downwind and crosswind distances relative to the lidar, respectively, and z is the vertical distance above the ground. Figure 2.8(a) is a scan of the puff right after the aerosol was released. This scan was taken as the source, having the initial size. The source puff, then, contained cloud points in the bottom six sweeps.

When the dispersion time increased to 41 s, vertical dispersion had increased, so that seven sweeps caught cloud points. For the rest of the scans, the cloud depth remained almost unchanged, which means that the change of vertical dispersion with time in stable conditions was very slow, due to the fact that the temperature of the ground surface was lower than the atmosphere above, and produced a net downward flux of heat, suppressing convective turbulence. The downwind and crosswind expansion of the puff is very obvious.



(a) t=0 s (b) t=21 s (c) t=41 s

(d) t=59.5 s (e) t=72.9 s (f) t=98.4 s

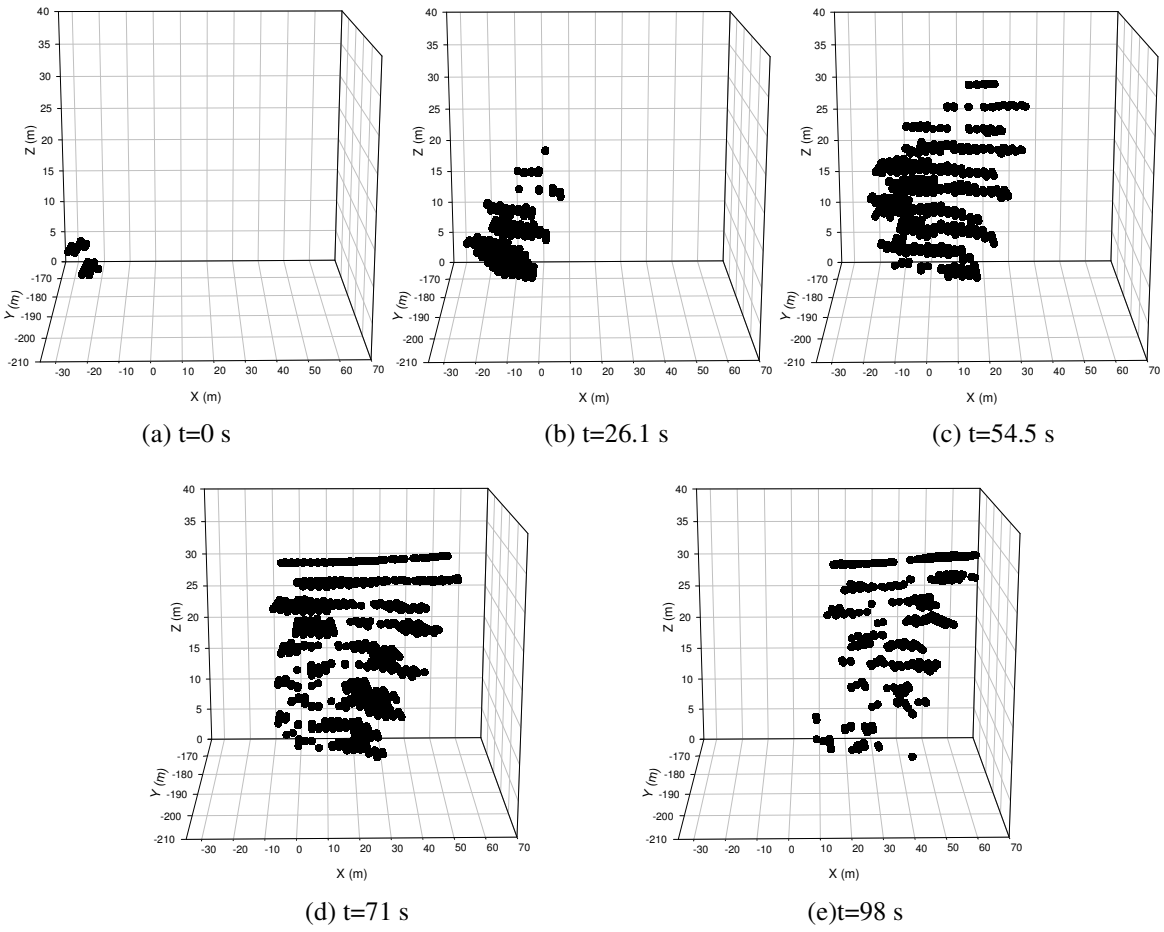
Figure 2.8 Cloud evolution in stable conditions

Another phenomenon observed in Figure 2.8 is that as the dispersion time increased, the puff showed a skewed shape, such that the top part of the cloud moved a little faster in the downwind direction than the bottom part. This is the influence of wind shear, where the wind speed increases with height in the surface layer, so that the advection wind speed of the particles at the top of the puff is faster than the advection of the particles close to the ground.

For all six scans, there was only a small change in cloud depth, although horizontal dispersion increased with time, as indicated by an increasing number of cloud points in each layer. Downwind and crosswind dispersions are results of wind shear and turbulent movement. Dispersion in the horizontal plane exceeded vertical dispersion in stable conditions.

Figure 2.9 is cloud/puff evolution with time under the conditions of Pasquill stability category A, or very unstable conditions. This trial was conducted at 17:00 on May 29th, 2003, with a wind speed of 0.5 m/s, a temperature of 14ºC, a relative humidity of 94%, and a pressure of 98.1 kPa. The initial scan caught only 2 layers containing cloud points, and the initial cloud depth was about 4 m. After about 26 s, cloud points were found in up to 7 layers, and the cloud depth increased to about 19 m. When dispersion time increased to 54.5 s, up to 10 layers caught cloud points, and the cloud depth reached about 30 m. When the dispersion time increased further, the lidar vertical scanning range was not wide enough to catch the cloud top, as only 10 sweeps were preset. Interestingly, more skewness was observed in this case than the one shown above in stable conditions. The shorter the dispersion time, the more skewness was shown, which is not in agreement with common thinking that vertical diffusion destroys the

influence of wind shear in unstable conditions. Only when the dispersion time is longer, do the observations satisfy the general theory about vertical diffusion and wind shear in unstable conditions.



(a) t=0 s (b) t=26.1 s (c) t=54.5 s

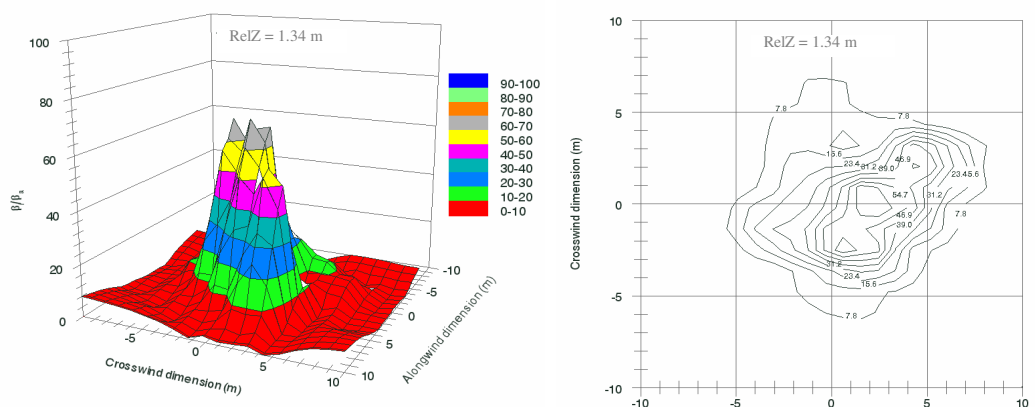
(d) t=71 s (e)t=98 s

Figure 2.9 Cloud evolution in unstable condition

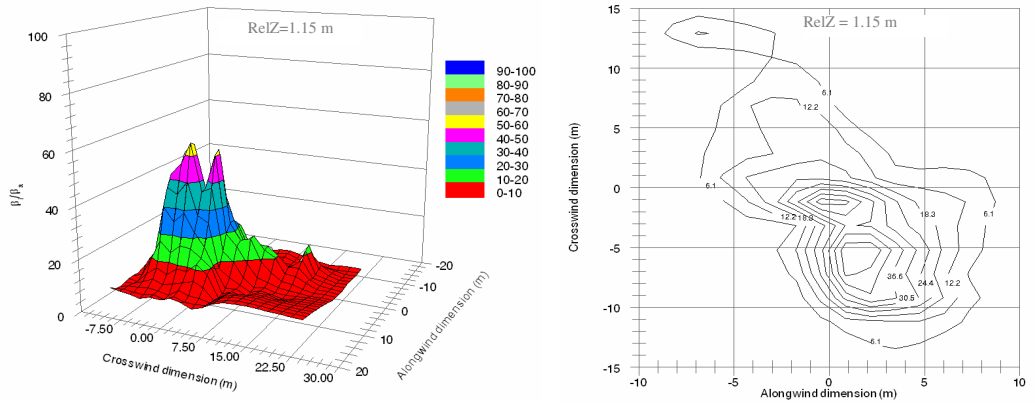
Figure 2.9 also shows that the puff, under unstable conditions, goes through a faster vertical dispersion initially than horizontal dispersion, although after a short time, horizontal dispersion will surpass vertical dispersion.

2.5.2 Concentration distributions

Figure 2.10 to 2.12 are some observations of concentration distributions from field experiments under different conditions. Figure 2.10 shows the mesh and contour plots of concentration distributions along downwind-crosswind planes at different times.



(a) t = 27.6 s



(b) t = 77.6 s

Figure 2.10 Concentration distribution over downwind-crosswind plane at times of 27.6 s (top) and 77.6 s (bottom) after release on Aug. 13, 2003, at 7:30 AM

The trial was conducted at 7:30 AM on August 13, 2003, with a transport wind speed of about 0.7 m/s, a temperature of 16.4ºC, an air pressure of 100.6 kPa, and a relative humidity of 88.9 %. The values shown in the plots are relative backscatter, β_i / β_a, after lidar inversion rather than absolute concentrations. The downwind and crosswind

coordinates are distances relative to the centre of mass of the puff. Figures 2.10 (a) and 2.10 (b) are concentration distributions at times of 27.6 s and 77.6 s after release for the cloud slices close to the centre of mass of the puff. RelZ in Figure 2.10 is relative vertical distance to the centre of mass.

It can be clearly seen that the concentration distributions show approximately Gaussian shapes locally and for the whole range, there are multiple peaks due to the turbulent nature of the atmosphere. The cloud appears to have broken into several small parts, with each part having approximately a Gaussian distribution or a Gaussian distribution with a long tail, as observed by Yee (1998) and Sato (1995). The concentration distribution shows a tail in the positive crosswind direction as dispersion time increases.

Figure 2.11 is the concentration distribution over the downwind-vertical plane with crosswind coordinates close to the centre of mass. Profiles at time 27.6 s and 77.6 s are shown. Similarly, the coordinates are relative distances to the centre of mass of the

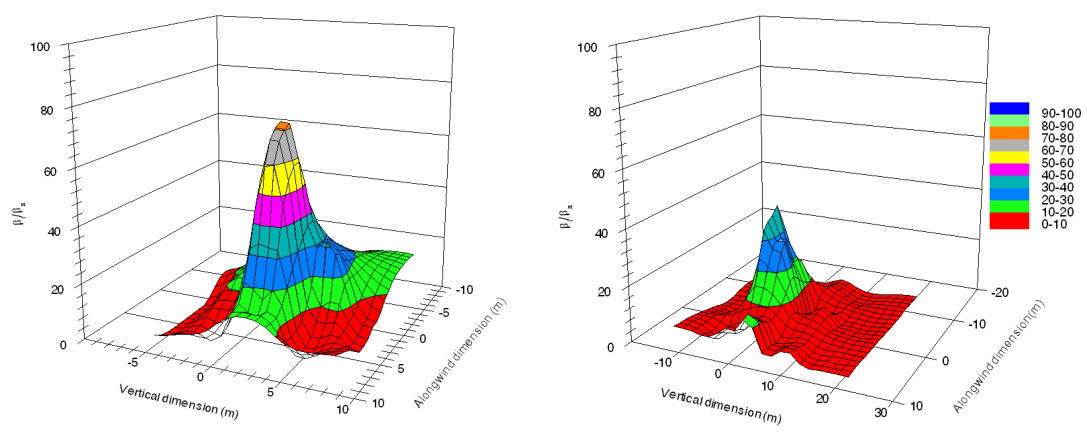
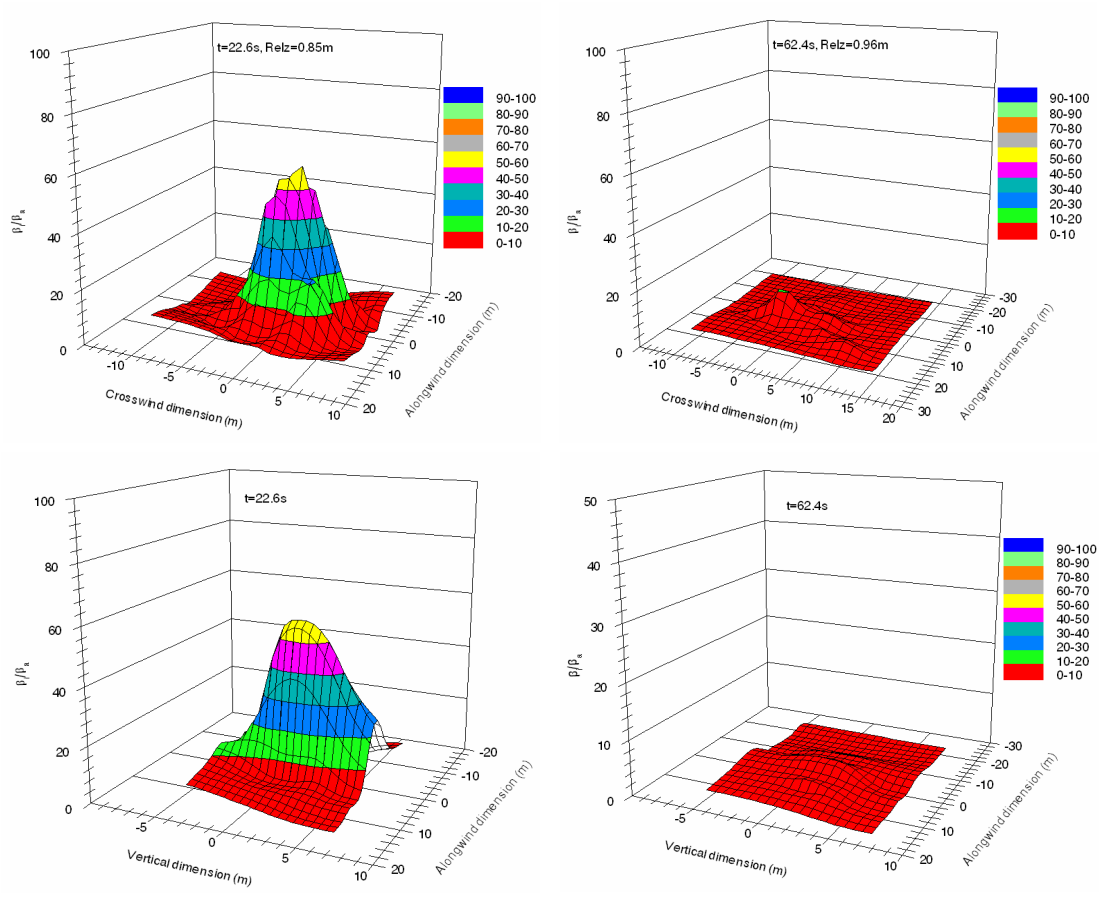


Figure 2.11 Concentration distribution over alongwind-vertical planes at release time 27.6s (left) and 77.6s (right) on Aug. 13, 2003 at 7:30 AM

cloud. Surprisingly, the vertical concentration distribution is closer to a Gaussian shape than the horizontal distribution, but it did not go down to zero for points above the centre of mass, which may be due to the fact that the last sweep with cloud points still had relatively higher concentrations, although the top of the cloud was not caught, due to the resolution of the lidar scanning, and only non zero points are plotted here. The distribution shows a long tail in the upward vertical direction with increasing time.

Figure 2.12 shows profiles of concentration distributions under unstable conditions detected at about 5 PM on August 12, 2003. The temperature was 28.1ºC, wind speed was 1.3 m/s, air pressure was 100.9 kPa, and relative humidity was 44 %. The



(a) t = 22.6 s

(b) t = 62.4 s

Figure 2.12 Concnetration distributions horizontally (top) and vertically (bottom) at times of 22.6 s and 62.4 s on August 12, 2003 at 5 PM.

concentration dropped quickly with time, as shown in Figure 2.12 both horizontally and vertically. Similarly, tails appeared with time in the crosswind and downwind directions. And since the aerosol mixed with the atmosphere quickly, it can be seen that at 62.4 s, the cloud showed a nearly uniform distribution vertically. The distribution also showed a shape close to but not exactly Gaussian. Some points were missed in the top of the cloud because of the limited lidar resolution in elevation.

These two individual cases show that puff concentration distributions are close to but not exactly Gaussian in shape. Actually, a tilted Gaussian shape can be seen with tails in both the crosswind and downwind directions. Vertically, almost all measurements miss some of the top of the cloud due to the limited elevation resolution of the lidar, which will affect the modelling result for the prediction of points far from the centre of mass vertically. This part will be discussed in the next chapter on the ANN modelling of puff dispersion.

Chapter 3
Modelling Puff Dispersion Using Neural Networks

Based on earlier experiments and observations, Gaussian distributions were established and used widely for puff dispersion. However, Gaussian models are based on a number of assumptions and simplifications that are not usually valid in a dynamic, turbulent atmosphere, and the Gaussian distribution in the vertical direction in a stably stratified boundary layer and in very convective atmospheric conditions is questionable (Arya, 1999; Hanna, 1982).

Lagrangian stochastic particle models are usually accepted to be the most powerful tools to describe dispersion from passive non-buoyant releases (Wilson and Sawford, 1996), but they typically require a large number of particles (with the order of 10^5, de Haan, 1998, 1999) to build up some statistical significance in the simulation and take excessive computing time, which greatly restrict their application in practice.

Artificial neural network (ANN) modelling is a new technique developed over the past 20 years. It tries to simulate a human brain's activities and can provide n-dimensional nonlinear correlations with the capability to reproduce synergistic effects among system variables. Theoretically, by using ANN modelling, the input variables are not required to be linearly independent, and they (ANNs) are particularly good at handling noisy and incomplete data. ANNs have been successfully used for forecasting daily maximum ozone concentrations in an industrialized urban area (Yi and Prybutok, 1996), concentration predictions of kaolin puff dispersion under unstable conditions (DeVito, 2000), short-term predictions of urban NO_2 (Cappa and Anfossi, 2001), and

ground-level ozone concentration predictions in large urban areas (Guardani and Nascimento, 2004). All these showed that neural networks are capable of more accurately predicting the concentrations, although input selection is very important. DeVito (2000) is the only one to use neural networks to model the concentration dispersion of puffs. His modelling data were based on field experiments from Pasquill stability category A (very unstable) to B (moderately unstable), and the ANN model developed was compared to COMBIC (a sophisticated Gaussian puff model) and a traditional Gaussian puff model using Slade's parameterizations for dispersion coefficients. The ANN model provided better predictions than either of those two models.

The ANN model developed by DeVito (2000) was based on data collected in unstable conditions only, thus it is not an appropriate model to be used to predict the concentration distribution of puffs under moderate and stable conditions. In addition, DeVito used discrete P-G stability categories to drive the model. A model using a complete meteorological data set (including turbulence parameters) would make a significant difference (better) compared to that using the single Pasquill stability category.

The major goal of the thesis research has been to develop a general model that can be used for all conditions. The ANN model was developed based on data collected under a wide range of meteorological conditions from November 2002 to August 2003. The data collected and processed was introduced in Chapter 2. Unlike most puff dispersion models that use only one parameter, i.e., Pasquill stability categories, each of which covers a very wide range of meteorological conditions (Pasquill and Smith, 1983), to represent the stability condition of the atmosphere, the ANN model developed used

general meteorological parameter measurements, including solar insolation, temperature gradients and some other turbulence-related parameters such as turbulent kinetic energy (TKE), Monin-Obukhov length and the structure of refractivity to represent the stability conditions of the atmosphere. Consequently, ANN predictions should be more representative of real atmospheric conditions. The developed ANN model will be compared with COMBIC and traditional Gaussian puffs with Slade's coefficients. Of interest is the concentration prediction for a whole cloud at different times as well as the maximum concentration at different points of time.

3.1 Modelling Data Selection and Processing

The data collected from field experiments from November 2002 to August 2003 were used for the ANN modelling. The details of the field experiments and data collection can be found in Andrews et al (2003) and Chapter 2. Not all experimental data were useful for modelling, so useful clouds had to be identified before the modelling.

A total of 1820 puff releases were conducted, with each release consisting of 6 scans on average, giving about 11,000 scans available. However, not all of them were suitable for modelling purposes. As stated in Chapter 2, on some occasions when the wind was strong, or under convective conditions, the lidar did not sample the complete extent of the puff. The lidar sampling window was limited and the cloud moved so quickly that significant parts of the cloud were not scanned, thus data from these releases were not used. For other cases, in very stable conditions, clouds hardly moved and thus overlapped, so it was impossible to distinguish the new cloud signals from old ones. These clouds too were removed in the selection of useful clouds for modelling. Finally, 60 releases were selected to be used for further analysis, including 54 scans of

microsphere clouds, 78 scans of glass spheres and 86 scans of talc. Table 3.1 is a summary of the meteorological ranges of all useful clouds.

Table 3.1 Range of meteorological conditions

	Wind speed (m/s)	T (°C)	RH (%)	Insolation (W/m^2)	Air Pressure (kPa)
Min	0.05	-22	29	0	98.1
max	4.6	28.4	100	856	101.1

Lidar signals were converted into relative concentrations β_i / β_a by Eqn. (2.10). If the cloud particle size distribution, density of aerosol particles, LIDAR ratios of atmospheric aerosols and cloud particles, visibility and LIDAR wavelength were available, then β_i / β_a could be converted into an absolute concentration with units of g/cm^3 using Eqn. (2.11) and Eqn. (2.12). Unfortunately, most parameters in Eqn. (2.12) were not measured at the site, some parameters such as visibility can be roughly estimated by using environmental records at a nearby airport (about 15 km away), and lidar ratios can be estimated by experience, although the estimations are only rough and might have an error of a factor of two. Under different conditions, the constant k may be different, therefore values of β_i / β_a from different clouds, after the lidar inversion, cannot be compared directly. One way to make them comparable is by using the value at the centre of mass of the initial cloud to normalize each point of the scans from the same release. For convenience, the peak value of the initial cloud was used to normalize all other points at all the time for the same release, based on the assumption that the maximum concentration is at the centroid.

Usually the coordinate system was selected with the downwind direction as the x axis, the crosswind direction as the y axis and the vertical direction as the z axis. Thus it was very important to determine the transport wind direction, to locate the x axis. The

meteorology system was set up about 200 m away from the release spot, and due to turbulence, the wind velocity at the release site might have a significant variation from the measuring site. Some observations showed that movement of some puffs followed a direction much different from that recorded at the measuring site, especially when the wind was light and variable. The wind speed and direction at the release site can be derived by measuring the distance of puff centers and time used from scan to scan, while the wind direction can be determined by tracking the movement of the centre of mass of puffs. The derived wind speed and direction were called the transport wind speed and direction, and their averages were taken as the effective wind speed and direction during puff dispersion. It was assumed that, for a short range or short time of dispersion, the transport wind velocity remained unchanged (this was actually observed for most trials).

The inverted lidar scans were analyzed in the framework of relative diffusion, where the coordinate system was centred on the centre of mass of a diffusing puff, and followed the puff as it was advected downwind. This removed the effect of the large eddies on the meandering of the puff, i.e., the meander of the puff was not considered as part of the diffusion process, only the spread of the cloud about its own centre was considered.

Since the actual puff release took about 2 seconds to accomplish, the study of puff dispersion started after the release was finished. The first available scan after complete release was taken as the source, and its size was used as the initial size of the puff. All dispersion times was measured relative to the manually selected source. Since all the useful clouds selected were transported with wind speeds less than 4.5 m/s, and with 2 s release time, the initial puff dimension was about or smaller than $10 \times 10 \times 10$ m^3.

3.2 ANN Modelling

Factors considered in developing a successful neural network model included selecting a group of representative data of conditions concerned, determining the appropriate inputs and output(s), and selecting training, test and validation sets.

The useful clouds covered a relatively wide range of conditions from stable to unstable, and these data were used for building an ANN model. The first step to be considered was to decide which variables should be used as inputs and which should be used as outputs.

3.2.1 Input/output selection for ANN

Outputs are usually selected by considering the variables or parameters to be predicted. For this case, the concentration distribution was the target, and the normalized concentration was used for the output of the ANN modelling.

The selection of inputs needs more consideration. When building models such as neural networks, it seems natural to assume that having more information is always better than having less. Instinctively, one would think that models using more input variables should do no worse, as they are less likely to not include a relevant control parameter. The reality of the situation is counter-intuitive; adding inputs gives the model more factors to consider, thus extra variables can confuse and dilute the outcomes.

Working with many variables expands the size of the input space, which consists of all the possible combinations of input values. If the input variables are not correlated, the addition of each new input multiplies rather than adds to the size of the input space. Moreover, the number of required examples for uncorrelated inputs grows geometrically

with the number of input variables (Dwinnell, 1998). But to determine which ones to keep and which ones to drop from all possible inputs candidates is a daunting task.

During field experiments, meteorological instruments and devices were set up and a number of micrometeorological data were collected, with these parameters being more or less related to the intensity of turbulence. Based on the work of other modellers, there is no doubt that the relative positions and dispersion time are directly related to the concentration distribution. Intuitively, the initial size of the puff formed by releasing the same amount of aerosols will affect the concentration distribution too, with the larger the initial size, the lower the concentration will be at the same point. Other parameters considered were the turbulence or stability parameters.

There are many basic meteorological parameters related to stability conditions, such as wind speed, insolation, temperature, pressure, relative humidity and their gradients. The meteorological variables are correlated to each other, as are their influences to the concentration distribution. The common Gaussian models use only one fitting parameter, based on, for example, the P-G stability category, to represent the turbulence or stability condition. However, each stability category covers a wide range of atmospheric conditions. One of the outstanding characteristics of ANNs is that they can use inputs that are correlated and still provide good performance. Consequently, measurements of meteorological variables were used directly, so that their direct influences on the concentration distribution predictions could be determined.

In addition to the basic meteorological measurements, the turbulence parameters such as the variance, covariance of variables, and structure parameters were derived by Forand (1999, 2003). All those variables were able to be investigated as inputs of the

ANN. But as discussed above, more inputs will make the network more complicated and may confuse the network and dilute the outcome.

Since structure parameters are functions of variance, covariance and correlated time or length, structure parameters plus TKE (turbulent kinetic energy) and MOL (Monin-Obukhov length) were selected to represent the turbulence condition, combined with other meteorological measurements. In all, 35 inputs were considered initially. The number of network inputs was gradually reduced by removing inputs contributing little to the output, until finally 21 inputs were left, with these inputs being summarized in Table 3.2.

Table 3.2 Variables considered as inputs for ANN models

Model input variables
1. Initial size of cloud in downwind direction, Rx_0 (m)
2. Initial size of cloud in crosswind direction, Ry_0 (m)
3. Initial size of cloud in vertical direction, Rz_0 (m)
4. Dispersion time, t (s)
5. Transport wind speed, U (m/s)
6. Solar elevation, α, (radians)
7. Particle shape, 1-sphere, 0-nonsphere
8. Mean size of particles, size (um)
9. Relative downwind distance, RelX (m)
10. Relative crosswind distance, RelY (m)
11. Relative vertical distance, RelZ (m)
12. Insolation, Inso (W/m^2)
13. Net ground heat flux, d_Inso (W/m^2)
14. Air temperature, T (°C)
15. Gradient of temperature , dT/dz (°C/m)
16. Relative humidity, RH (%)
17. Gradient of RH, dRH/dz (%/m)
18. Air pressure, P (hPa)
19. Monin-Obukhov length, MOL (m)
20. Turbulent kinetic energy, TKE (m^2/s^2)
21. Refractivity parameter, C_N^2 ($1/m^{2/3}$)

3.2.2 Selection of training, test and validation sets

The purpose of developing a neural model is to produce a formula that captures essential relationships among data. Once developed, this formula is used to map from a new set of inputs to corresponding outputs. The training set is the set of points that are used to fit the parameters of the model. The test set measures how well the model predicts. The test set is also used as part of the model building process, to prevent over-fitting. The validation set is used to estimate model performance in a deployed environment.

The validation set should be typical of the deployed environment. The main goal of this research is to develop a general model appropriate for broad conditions. In order to best evaluate the models developed, four releases under different conditions were set aside for validation, with these validation trials being shown in Table 3.3. The validation trials were selected from experiments in August from stable to unstable conditions. For convenience and later comparison with other models, the corresponding P-G stability categories were determined for each trial.

Table 3.3 Atmospheric conditions for validation set

Trial	Release time	Dispersion time (s)	wind speed (m/s)	Temp. (°C)	Air Pres. (kPa)	RH (%)	P-G Stability
1	02:24	22.9-104.7	0.50	9.9	100.3	97.8	G(very stable)
2	16:52	22.6-62.4	1.31	28.1	100.9	43.9	B(moderately unstable)
3	05:40	26.1-119.8	0.54	10.8	100.2	97.4	D-E(Neutral to Slightly stable)
4	23:20	17.9-65.0	0.58	11.1	100.5	94.8	F(stable)

After the validation trials were set aside, the rest of the data were used for training and test set selection.

The training set is very important for building a successful neural network model. A good training set should contain all the relationships between inputs and output(s) with a minimum of training data. Among all the available data, more than 95% were for low concentrations (normalized concentration is lower than 0.001). It was not felt necessary to use all the low concentration data, as too many low concentration data points would lead to the model greatly underestimate the high concentration values. Consequently, about 15% of the low concentration points along each shot were selected, ensuring that boundary information would not be lost, together with all other higher concentration data points to form a new data set for training and test set selection. Of this subset, 75% were randomly selected for the training set, and the remaining 25% for the test set. In selecting the training and test sets, a rule was followed that all data in the test set was to be within the range of training set. Also, since there was a very limited number of very high concentration points, some high concentration points were duplicated to help balance the output distribution in all the data range, i.e., some duplications were performed to make the output near-uniformly distributed through the output space.

3.2.3 Transformation

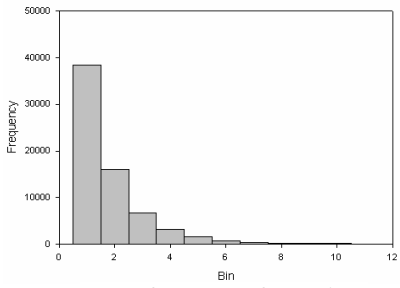
Data transformation may assist neural network development if the transformations make the training data uniformly distributed through the input and output spaces (Wasserman, 1993). The same transformations should be applied to all the training, test and validation set data. Usually there are three steps for transformations, namely mapping the original range of each input and output to a new range by linear scaling, performing nonlinear transformation for each variable, and then performing internal linear scaling to a range [-1,1] or[0, 1], depending on the transfer function used.

The first two steps can be symbolized as:

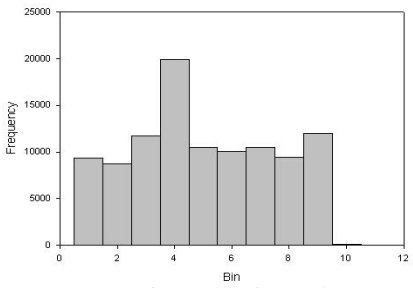
$$y = f\left[g_{\min} + (x - x_{\min}) * (g_{\max} - g_{\min}) / (x_{\max} - x_{\min})\right], \tag{3.1}$$

where $g_{\max}$ and $g_{\min}$ are the maximal and minimal values of each input and output range after linear transformation. $x_{\max}$ and $x_{\min}$ are maximal and minimal values of each input variable in physical units, and f is a nonlinear or linear function. The initial two steps may need to be repeated several times to find the best $g_{\max}$, $g_{\min}$ and the function f.

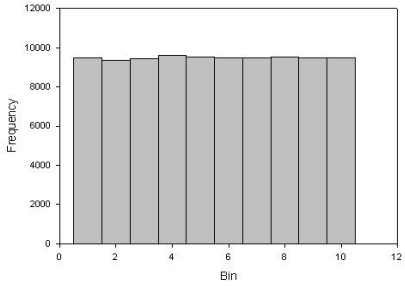
Figure 3.1 shows how transformation helped the output for the ANN be uniformly distributed in the output space. Before transformation, it can be seen that the original output was mainly distributed in the low concentration range, even after many of them had been excluded (Figure 3.1a). After the linear and non-linear transformations discussed above, the frequency distribution showed a more uniform shape in the output space (Figure 3.1b), and by duplicating some records, Figure 3.1c shows a good uniform distribution.



(a) Before Transformation



(b) After Transformation



(c) After Transformation and duplication

Figure 3.1 Relative concentration frequency distribution before and after transformation

Some of the input frequency distributions after transformations are shown in Figure 3.2.

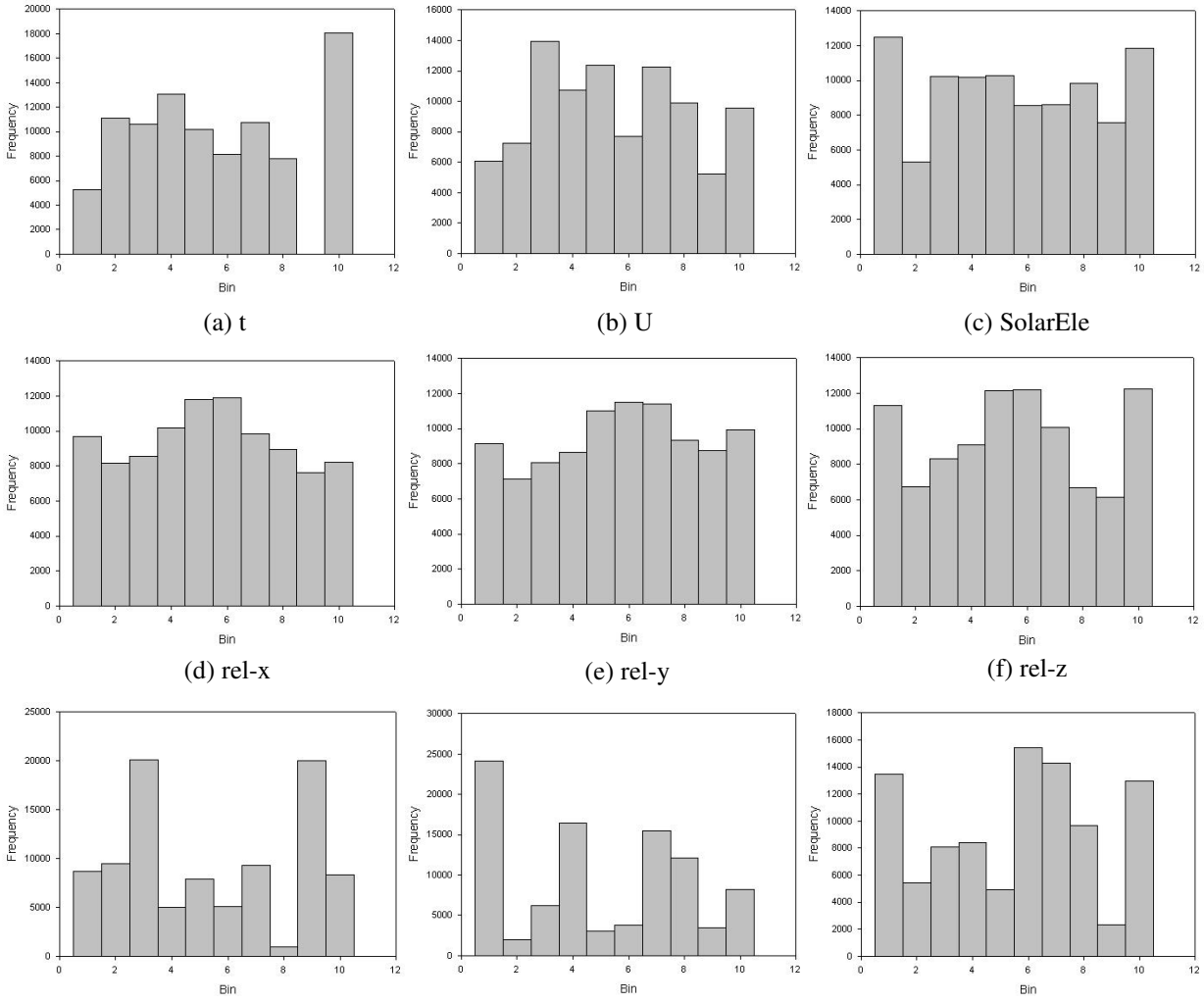


Figure 3.2 Input frequency distributions after transformations

It is ideal if all inputs and outputs are uniformly distributed in input and output space by transformations and duplications, but this is extremely difficult, if not impossible, considering the number of variables involved and that changing one variable distribution will affect other variable distributions. Compromises have to be made in order to retain the emphasis on the output (Baughman and Liu, 1995).

A good systematic description of how to better prepare data for neural network data analysis can be found from Yu et. al (2006).

The training and test sets, being subsets of the overall transformed data set, have similar distributions.

3.2.4 Neural network modelling

The commercial shell used in this study was NeuralWorks Professional II/Plus by NeuralWare (2003). The particular paradigm used was a multi-layer feedforward (MLFF) backpropagation (BP) network using the extended delta-bar-delta (EDBD) learning rule (Minai and Williams, 1990), where the prediction error is minimized using gradient descent and learning and momentum coefficients are adjusted dynamically during the training process. A detailed description of the theory involved with ANN training can be found in Andrews (1999).

Several MLFF neural networks were trained, with each having a significant number of hidden nodes. The **SaveBest** and **Prune** functions of Neural Professional II/Plus were used in the training process. One of the problems that can occur with BP networks is over training, i.e., the network remembers the specific records for training, and performs well on the training data, but poorly on independent test data. **SaveBest** allows the user to run train/test cycles and save the network with the best test result during ANN learning. **SaveBest** is one way used to avoid an ANN being over trained.

Another way to deal with the problem of over training is pruning. A network of a minimal complexity which does well on a training set will generalize better than a more complex network (Patterson, 1996). Pruning a network is the process of removing

unnecessary processing elements and connections, including removing both input and hidden nodes. The prunning of input nodes has been done by decreasing the initial 35 inputs to 21 inputs, and in the process of ANN learning, pruning means removing hidden nodes. Pruning attempts to minimize both network complexity and the error over a training set by reducing the number of hidden nodes without reducing the model's ability of predicting. Training continues until the simplest model (fewest hidden nodes) is found having no appreciable degradation in performance.

The best ANN model was employed and was compared with other puff models. The developed ANN model was evaluated by statistical performance measures which gave an idea of the discrepancy between predictions. Table 3.4 is a summary of the performance evaluation of the best ANN model. The statistical evaluation of the ANN model was based on the following model performance measures: normalized mean square error (NMSE), correlation coefficient (R), fractional bias (FB), factor of 2 (F_2) and factor of 10 (F_{10}), geometric mean bias (MG) and geometric mean variance (VG) (Mohan and Siddiqui, 1997), where

$$\text{NMSE} = \frac{1}{N}\sum_i \frac{(P_i - M_i)^2}{\overline{P}\,\overline{M}}, \tag{3.2}$$

$$\text{R} = \frac{\sum (P_i - \overline{P})(M_i - \overline{M})}{\sqrt{\sum (P_i - \overline{P})^2 \sum (M_j - \overline{M})^2}}, \tag{3.3}$$

$$\text{FB} = \frac{\overline{M} - \overline{P}}{0.5\left(\overline{M} + \overline{P}\right)}, \tag{3.4}$$

$$\text{F}_2 = \text{fraction of data where } 0.5 \le P/M \le 2.0, \tag{3.5}$$

$$\text{F}_{10} = \text{fraction of data where } 0.1 \le P/M \le 10.0, \tag{3.6}$$

$$\mathrm{MG} = \exp\left(\frac{1}{N}\sum \ln M_i - \frac{1}{N}\sum \ln P_i\right) = \exp\left(\frac{1}{N}\sum \ln\left(\frac{M_i}{P_i}\right)\right), \qquad (3.7)$$

$$\mathrm{VG} = \exp\left[\frac{1}{N}\sum_i \left(\ln M_i - \ln P_i\right)^2\right] = \exp\left[\frac{1}{N}\sum_i \left(\ln \frac{M_i}{P_i}\right)^2\right], \qquad (3.8)$$

where M is the normalized concentration from measurement, P is the ANN prediction of normalized concentration, and $\overline{P}$ and $\overline{M}$ are the corresponding averages over the data set to be used.

A perfect model would have NMSE and FB equal to 0.0, and R, F_2, F_{10}, MG and VG equal to 1.0. Geometric mean bias values of 0.5 and 2.0 can be thought of as "factor of two" over-predictions and under-predictions in the mean, respectively. A geometric variance value of about 1.6 indicates a typical factor of two scatter between the individual pairs of observed and predicted values.

A value of 1.99 in absolute fractional bias indicates no agreement between model predictions and observations, 1.0 corresponds to model predictions within a factor of 3 of the observations and for a value of 0.67 they are within a factor 2. Negative values of FB indicate a model is over-predicting and positive values mean a model is under-predicting. Fractional bias is non-linear and bounded by ± 2.

MG and VG are used here because the data set contains pairs of data with M_i/P_i equal to 10, 100, or more, so the logarithmic forms may be more appropriate measures of model performance than are the common linear forms (Mohan and Siddiqui, 1997).

Since validation trials represent typical experimental conditions, but have not been involved in the training process of the ANN, it is more reasonable to evaluate ANN

performance over validation sets than over the training and test sets. ANN evaluations over validation sets are summarized in Table 3.4.

Table 3.4 ANN model evaluations over validation sets

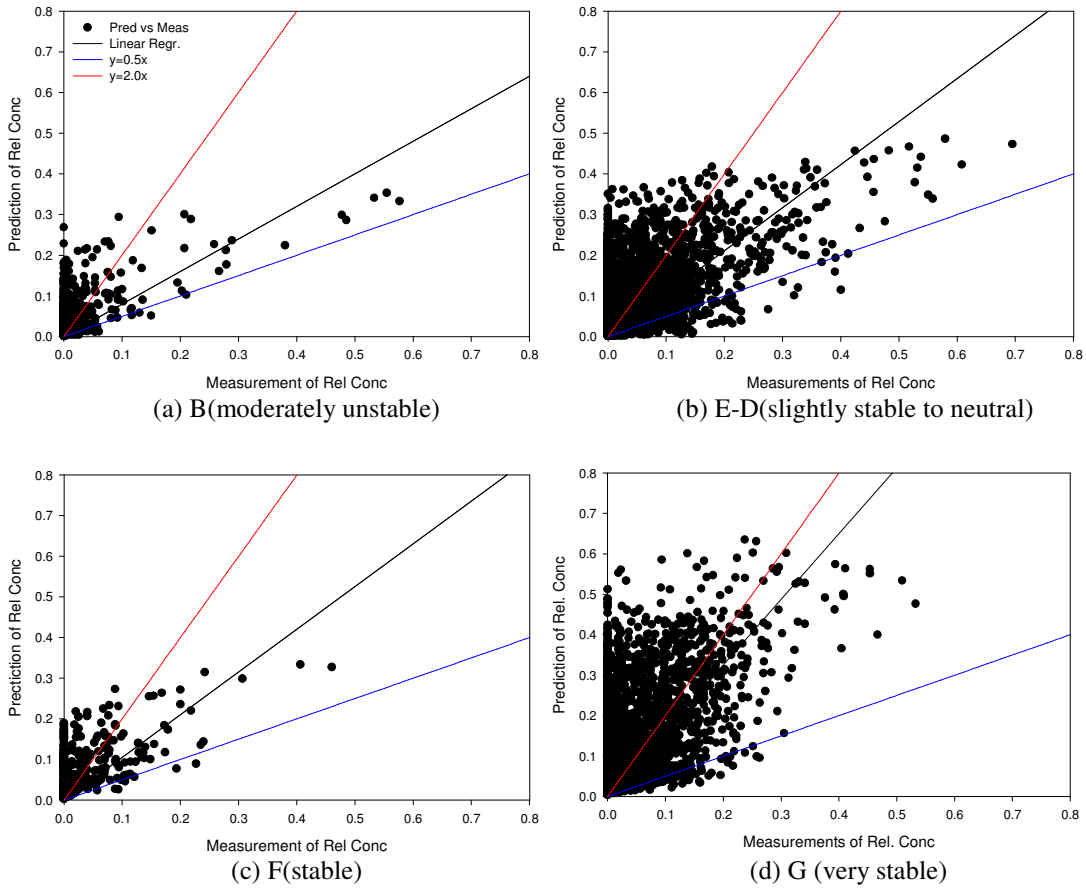
Validation	NMSE	R	FB	F_2	F_{10}	MG	VG
G(very stable)	1.76	0.58	-0.75	0.42	0.86	0.36	11.32
B(moderately unstable)	1.49	0.68	-0.45	0.43	0.85	0.42	11.01
E-D(slightly stable to neutral)	1.16	0.64	-0.46	0.44	0.89	0.45	8.41
F(stable)	0.79	0.66	-0.41	0.50	0.94	0.47	4.73

Considering the randomness of the turbulence for each trial and that the measurements are instantaneous values, the correlation coefficient of 0.6 or even higher is considered acceptable. The fractional bias for each case is negative which means the ANN overestimated the normalized concentrations, which, as will be shown later, is mainly due to the overestimations of points far from the centre of mass. Some 40% to 50% of predictions are in the range of a factor of two of measurements for all four of the validation trials.

Figure 3.3 shows the relationship between ANN predictions and measurements over each validation set. Each point represents a measurement and the corresponding predicted value. The blue and red lines are for predictions that are half and twice the value of the measurements. The black lines are linear regressions of the relationship between predictions and measurements, ideal results would be at a 45˚ angle.

The ANN under-estimated the high concentrations in all conditions, although all high concentration predictions were in the range of a factor of 2. Medium concentrations were well predicted, while low concentrations were over-predicted. As there were far more low concentration points included in the validation set than high concentration

points, it is not surprising to see the negative FB shown in Table 3.4, which means that, in general, the ANN over-estimated the concentration points.



(a) B(moderately unstable) (b) E-D(slightly stable to neutral)

(c) F(stable) (d) G (very stable)

Figure 3.3 ANN predictions vs measurements

The ANN under-estimated the high concentrations in all conditions, although all high concentration predictions were in the range of a factor of 2. Medium concentrations were well predicted, while low concentrations were over-predicted. As there were far more low concentration points included in the validation set than high concentration points, it is not surprising to see the negative FB shown in Table 3.4, which means that, in general, the ANN over-estimated the concentration points.

3.3 Model Comparison

The Combined Obscuration Model for Battlefield Induced Contaminants, or COMBIC, is a program used by the US Army to estimate spatial and temporal variations in transmittance through a battlefield obscured by smoke and dust clouds (Ayres and DeSutter, 1995). Its modification can be used to get densities directly (Mungiole and Wetmore, 2001).

For puff dispersion, COMBIC uses the standard Gaussian puff model, together with a number of enhancements to account for additional effects including models for predicting vertical profiles of wind-speed, temperature and pressure, a thermal buoyancy model, and ground interaction models. COMBIC uses two methodologies to define the puff's dispersion coefficients, σ_x, σ_y and σ_z, which are directly related to cloud dimensions. For downwind travel times less than 30 s, σ_x, σ_y and σ_z are determined by a power law, based upon the downwind distance *x*:

$$\sigma_x(x) = 0.740 A_i x^{0.9}, \tag{3.9}$$

$$\sigma_y(x) = 0.667 A_i x^{0.9}, \tag{3.10}$$

and

$$\sigma_z(x) = C_{ij} x^{D_{ij}}, \tag{3.11}$$

where, the coefficients *A*, *C* and *D* are those used by Hansen (1979), the subscript *i* denotes the Pasquil stability category, and *j* denotes the surface roughness index.

For downwind distances greater than the distance associated with a downwind travel time of 30 s, the values of σ_x, σ_y and σ_z depend on the fractional stability category, wind speed, scaling ratio and surface roughness length. A uniform wind

direction is used in COMBIC and details about dispersion coefficients can be found in manual written by Ayres and Desutter (1995).

After studying a limited number of field trials of puff releases, Slade suggested that σ_y and σ_z can be represented by power-law functions, given the downwind distance and stability (Slade, 1968; Turner, 1994). No suggestions were provided for the downwind dispersion coefficient, σ_x, and in general, one assumed that $\sigma_x = \sigma_y$. The initial dimensions of a puff are approximated by finding a virtual distance (x_v) to give the appropriate initial standard deviation for each direction, and σ_y and σ_z are determined as functions of $x + x_v$.

Table 3.5 provides a statistical evaluation of ANN, COMBIC and Slade's Guassian puff model against all validation sets, including clear air points right next to the cloud boundaries. If the linear correlation coefficient is considered, then COMBIC gave the best prediction in stable conditions. Under near neutral conditions, the three models were close to each other, and in unstable conditions, the ANN performed the best. But if one compares the predictions to measurements, ANN showed the best performance no matter what the stability condition was. The large geometric variance of COMBIC and Slade's Gaussian puff model in near neutral and stable conditions shows that there are some portions of the clouds that were significantly under-predicted, while the ANN model did not show this trait.

Table 3.5 Statistical evaluation of ANN, COMBIC and Slade's puff model. The validation sets are identified by their appropriate P-G stability categories.

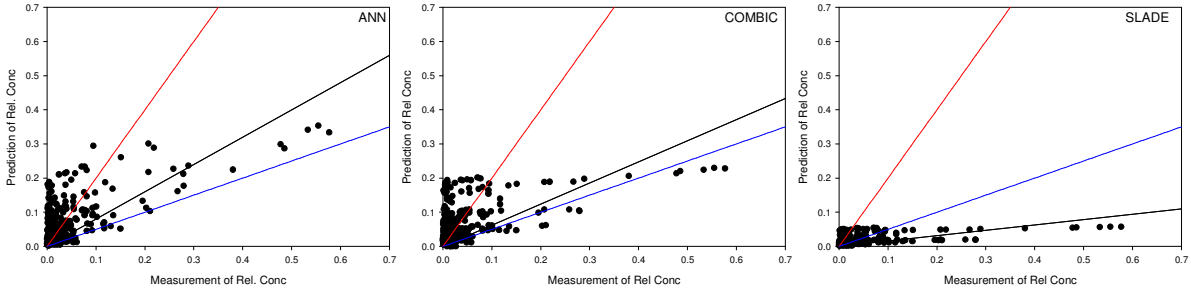
Validationn	Model	NMSE	R	FB	F_2	F_{10}	MG	VG
G very stable	**ANN**	1.76	0.58	-0.75	0.42	0.86	0.36	11.32
	COMBIC	1.02	0.62	-0.04	0.42	0.81	2.42	5446.17
	Slade's puff	2.35	0.55	-0.08	0.23	0.59	12.75	4.068E+11
B moderately unstable	**ANN**	1.49	0.68	-0.45	0.43	0.85	0.41	11.01
	COMBIC	1.90	0.53	-0.44	0.34	0.81	0.40	54.16
	Slade's puff	8.34	0.48	0.82	0.36	0.93	1.28	5.31
E-D slightly stable to neutral	**ANN**	1.16	0.64	-0.46	0.44	0.89	0.45	8.41
	COMBIC	1.14	0.66	-0.04	0.32	0.70	6.00	7.13E+07
	Slade's puff	1.44	0.65	0.04	0.26	0.55	151.86	4.57E+55
F stable	**ANN**	0.79	0.66	-0.41	0.50	0.94	0.47	4.73
	COMBIC	2.1	0.67	-0.86	0.37	0.85	0.46	72.35
	Slade's puff	3.23	0.61	-0.87	0.31	0.84	0.69	29.73

For a better illustration of model predictions, Figure 3.4 shows a comparison of the predictions from the different models with observations.

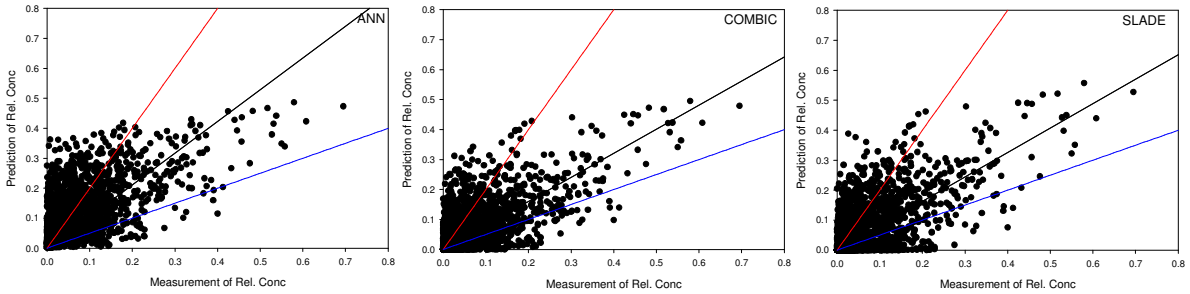
Figure 3.4 shows that, under unstable condition, both COMBIC and Slade's puff model greatly under-estimated the higher concentrations. The ANN model under-estimated the high concentration as well, but the predictions were much closer to observations than those of the other two models (Figure 3.4 a).

Under near neutral conditions (Figure 3.4 b), the performance of the three models were similar, but Slade's Gaussian puff model showed a little better prediction for high concentration points.

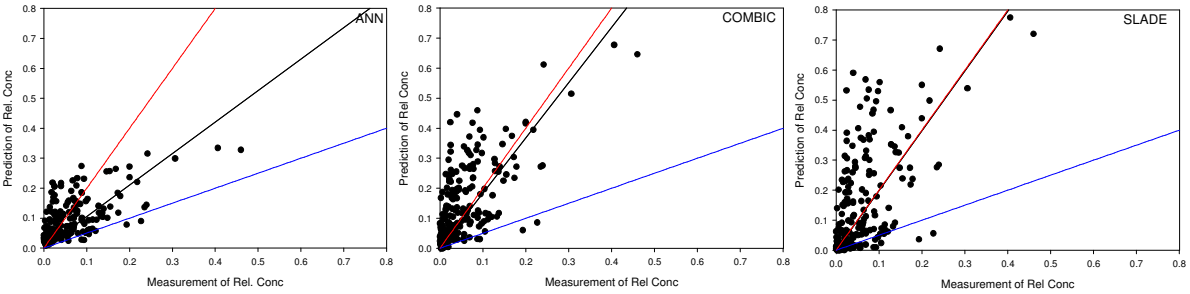
In stable conditions (Figure 3.4 c and Figure 3.4 d), all three models over-estimated low concentration points, but the ANN model was much better compared with the other two models (Figure 3.4 c). COMBIC was better when used for long dispersion times in stable condition. Figure 3.5 shows COMBIC predictions at different dispersion times for the case shown in Figure 3.4 d, at the stability condition G.



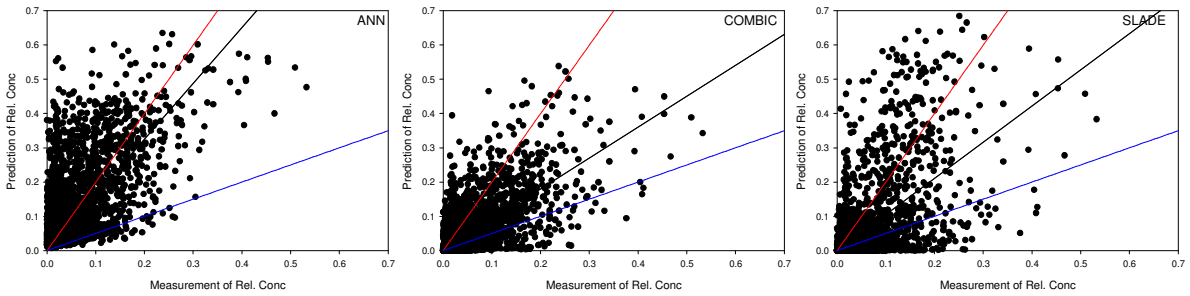
(a) B(moderately unstable)



(b) E-D(slightly stable to neutral)



(c) F(stable)



(d) G(very stable)

Figure 3.4 Prediction vs observation of different models for the validation set

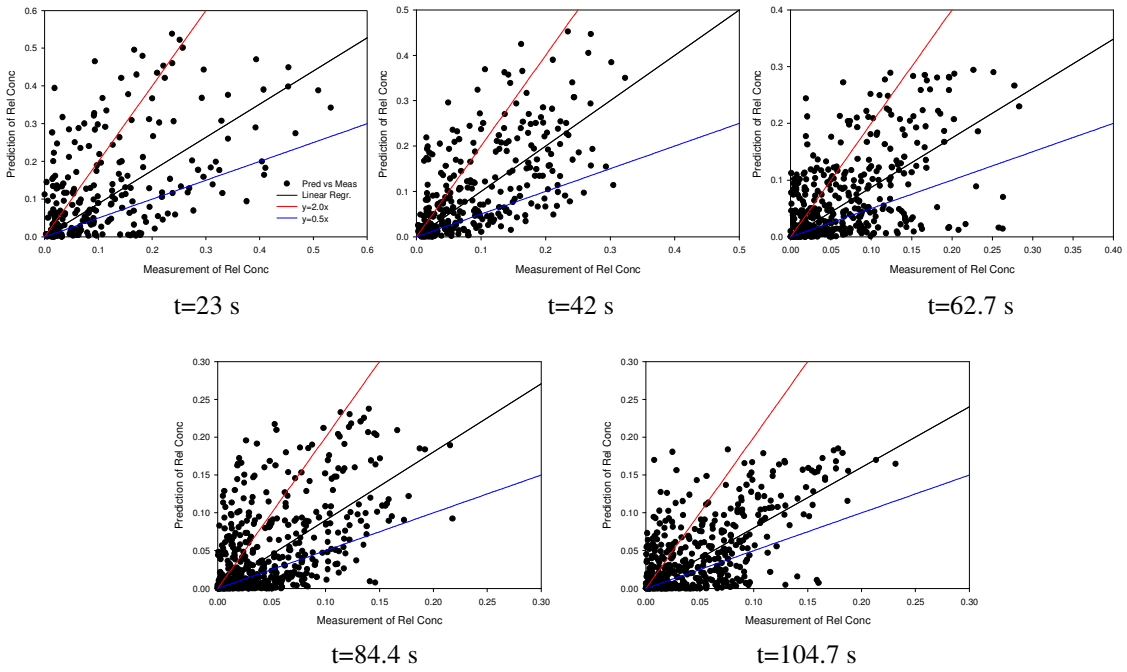


Figure 3.5 COMBIC predictions at different times in stable condition

Figure 3.5 shows that when the dispersion time is very short, COMBIC underoverestimated the concentrations, or, in other words, COMBIC underestimated the dispersion of the aerosols. As dispersion time increased, its performance of predicting higher concentrations improved, but it still overestimated the low concentrations. COMBIC did not provide good predictions when used for very short dispersion times, as can be seen from the case of stability F (Figure 3.4c) for which the dispersion time is shorter than 65 s, with greatly over-estimated concentration ditributions.

The time dependence of the predictions by ANN and Slade's Gaussian puff model were also checked, and it was found that ANN predictions for very short dispersion times were better than longer times, while Slade's puff model greatly underestimated the puff dispersion under stable conditions for both short and long times.

It is worth noting that the above predictions were concentrations normalized by the maximum concentration of the source puff. To use the ANN model, the maximum concentration of the source puff had to be estimated first. Usually, it was estimated by source size and a Guassian distribution model.

3.4 ANN Modeling Using Absolute Concentrations

Bissonnette and Roy (2003) demonstrated that signals after lidar inversion can also be converted into absolution concentrations of particles, only if additional parameters are known. The relationship between absolute concentrations and lidar inversions (ratio of backscattered coefficients of cloud points to background) is

$$c_{\mathrm{i}}(R) = 2.6\frac{k_{\mathrm{a}}}{k_{\mathrm{i}}}\frac{\rho_{\mathrm{i}} r_{\mathrm{ie}}}{V}\frac{\beta_{\mathrm{i}}(R)}{\beta_{\mathrm{a}}}, \tag{3.12}$$

where the parameters in Eqn. (3.12) are the same as those introduced in Chapter 2 for lidar inversion.

Though none of the parameters in Eqn. (3.12) were measured at the site, they can be roughly estimated by experience and the records of Environment Canada (EC) from the airport nearby (about 15 km away). Lidar values $k_{\mathrm{a}} = 0.025$ and $k_{\mathrm{i}} = 0.01$ were recommended by the lidar group of DRDC Valcartier, for which the errors might be in the range of a factor of two. The visibility data were obtained from the airport near the trial field (about 10 km away), but EC only attributed maximum visibility as 15 miles, good enough for flight, but not very accurate for the conversion to absolute concentration. In order to determine the differences between predictions of absolute concentrations and those of relative concentrations, and to assess the effectiveness of predictions from the model developed for absolute concentrations to measured values, 10

ANN models with absolute concentration were developed using the above estimates of k_i, k_a and visibility, and their average of predictions were compared with predictions from COMBIC and Slade's Gaussian puff model. Table 3.6 is a summary of the statistical evaluations.

Table 3.6 Statistical evaluation of an ANN model predicting absolute concentrations and COMBIC and Slade Gaussian models for the validation sets (indicated by their corresponding P-G stability categories)

Validation	MODEL	NMSE	R	FB	F_2	F_{10}
G very stable	**ANN**	1.17	0.46	-0.02	0.47	0.94
	COMBIC	17.59	0.22	0.56	0.06	0.16
	Slade's puff	61.90	0.15	0.47	0.04	0.11
F stable	**ANN**	0.93	0.60	-0.66	0.36	0.82
	COMBIC	4.63	0.64	1.12	0.17	0.39
	Slade's puff	6.72	0.53	1.08	0.07	0.20
C-D slightly unstable to neutral	**ANN**	3.18	0.64	-0.94	0.13	0.64
	COMBIC	53.65	0.32	1.53	0.06	0.37
	Slade's puff	102.12	0.18	1.72	0.03	0.24
B moderately unstable	**ANN**	7.40	0.42	0.93	0.28	0.86
	COMBIC	49.82	0.23	1.74	0.08	0.32
	Slade's puff	66.22	0.17	1.80	0.05	0.24

Here, ANN predictions are much better than the other two models for all conditions. The ANN has a much lower NMSE, higher correlation coefficient and the ANN's predictions are much closer to observations by providing much higher F2 and F10 values compared with the other two models. The ANN model overestimated the lower concentrations a bit for all conditions except the case in unstable conditions, and both COMBIC and Slade's Gaussian puff model underestimated the concentrations in all conditions, especially in near-neutral and unstable conditions (high FB). All models underestimated concentration distributions in unstable conditions. One possible reason for the Gaussian puff models under-estimating concentrations is that the dispersion coefficients for unstable conditions are not the proper ones and lead to larger dispersion.

Another reason might be that the parameter estimations in the conversion from relative concentration to absolute concentration may have large errors introduced and led to higher concentration measurements (for example, if visibility is actually twice the value used, the resulting absolute concentrations will be half the current estimations).

3.5 Conclusion

To summarize, the ANNs developed have the ability to model puff dispersion suitable for all conditions so long as sufficient experimental data are available and are evenly distributed for all conditions. Generally speaking, ANN models usually underestimate maximum concentrations in all conditions which have been reported by others (Yi and Prybutok, 1996; DeVito, 2000; Cappa and Anfossi, 2001). This, notwithstanding, a successful model for relative concentration has been developed that is based on a wide variety of experimental dispersion data collected under a broad range of atmospheric and aerosol conditions. This model considered explicitly a significant number of meteorological and turbulence parameters.

The developed ANN model based on relative concentration was compared with COMBIC and Slade's Gaussian puff model and demonstrated that, in general, the ANN model showed better performance than either the other two models. The ANN model generally overestimated the lower concentrations and underestimated the higher ones, especially under stable conditions. The main reason might be due to the fact that there were significantly fewer higher concentration points in the training data, compared with lower concentration points, and during the training process, the ANN had more chance to "learn" information from low concentration points than from a few points with high concentrations.

An ANN model taking absolute concentration as the output would give better performance than an ANN model using relative concentration. The maximum concentration of the source needs to be known or estimated in advance when using an ANN model for relative concentration. In practice, to get the absolute concentration of puff points, an improper estimation of the source maximum concentration will introduce errors to the final concentration predictions, so it is preferable to use absolute concentration as a direct output of an ANN model. This, however, requires a good knowledge of the values of all the required parameters. Such values were not available with a high degree of fidelity for the work reported here.

The deployment of the ANN model is limited to conditions within the range of values used for model training and development. Like all empirical correlations, the ANN model should not be used for extrapolations.

Chapter 4
ANN Simulations of Puff Dispersion

The developed ANN models discussed in Chapter 3 have demonstrated that ANNs are better at predicting puff concentrations for short dispersion times in all conditions than Gaussian models. The developed ANN can also be easily converted into a deployable code, e.g., in C or C++ (NeuralWare, 2003). So the deployment of the trained ANN is a few seconds work when necessary inputs are supplied. However, unlike other (Gaussian or Lagrangian) puff dispersion models, ANN models do not directly give an explicit explanation of the rules the puff dispersion follows. Simulation results of the developed ANN model in different conditions can provide insights into how a puff disperses.

4.1 Significance of ANN Inputs

It is very useful to determine which input variables have the most significant influence on an ANN model. This can help clarify which inputs are the best descriptive variables in modelling the process at hand, and may suggest if certain variables can be excluded from the model.

The developed ANN model for normalized concentrations was analyzed to determine the influence of each input variable on the predicted concentration. This was done by dithering, i.e., adjusting by ±5% each input value and checking how much the corresponding output changed as a percentage of an input change over the training and test sets (Neuralware, 2003), a process analogous to taking partial derivatives. Figure 4.1

shows the output change with each of input change over the training (blue bar) and test (brown bar) sets.

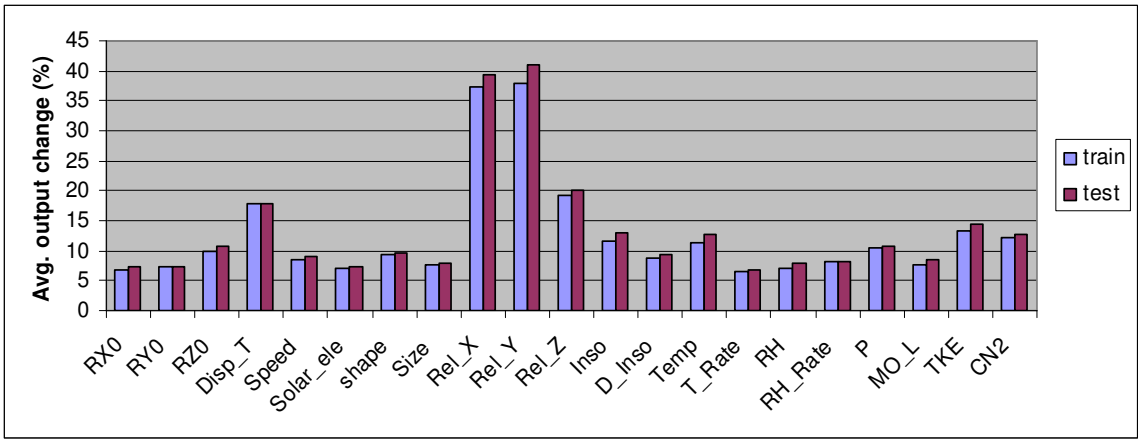


Figure 4.1 Change of ANN output over train/test sets with each input changed by 5%

Figure 4.1 shows that position within the cloud, relative to the centre of mass, and dispersion time play the most important roles in the output, e.g., a 5% change of the relative downwind and crosswind distances caused about a 40% change of the output, and a 5% change of the relative vertical distance caused approximately a 20% change of the output. A 5% change of dispersion time (or, time after release) caused about a 18% change of the output. These 4 inputs variables are major inputs in the process of ANN development. There is no doubt that the positions relative to the centre of mass and the dispersion time are key factors influencing the concentration distribution. Intuitively, the closer to the centre of mass, the higher the concentration; the longer the dispersion time, the more expansion of the puff, and the lower the concentration.

In addition to the four inputs mentioned above, TKE, the structure parameter of refractivity (C_N^2), solar insolation, air temperature, and air pressure also showed relatively large influences on the output, which may imply that these parameters acting

together are more related to the intensity of turbulence and stability. However, their contribution to the output is not significantly greater than the other stability parameters, so the remaining stability-related inputs should still be retained.

4.2 Horizontal Concentration Distributions

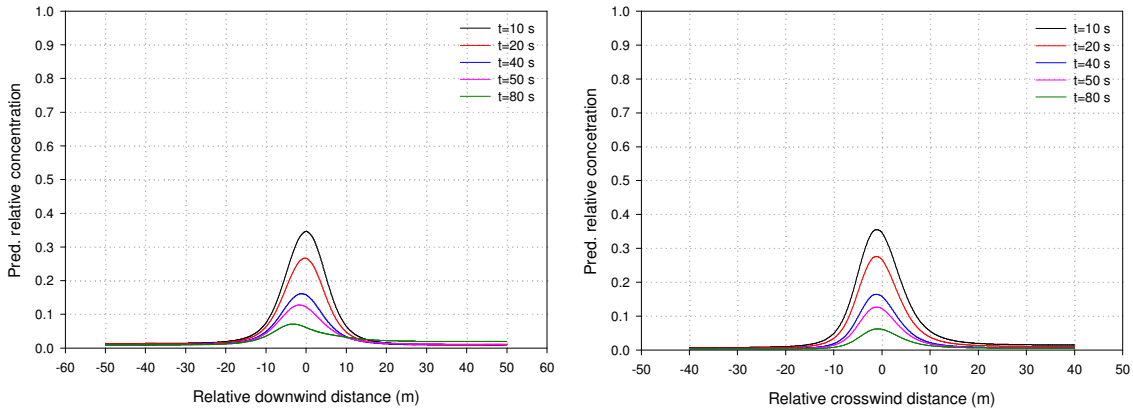
The sensitivity analysis above can help determine which inputs are more important to the output, and if the concentration distribution is known, then those more important parameters may help develop a relatively simple model by following the shape of the concentration distribution.

The concentration distributions of puff dispersions under different conditions were examined by ANN simulations. Three artificially generated cases representing P-G stability conditions unstable (B), neutral (D) and stable (F) were used for the simulations. The basic parameters related with these five conditions are listed in table 4.1.

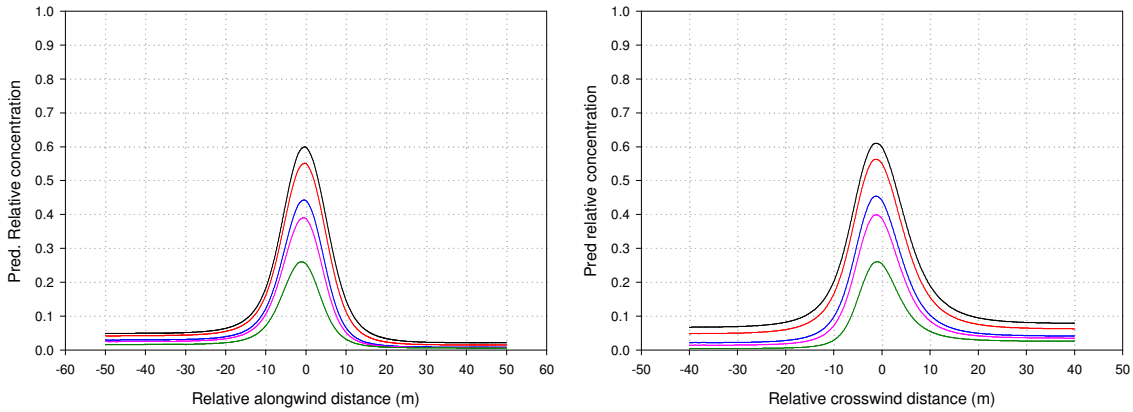
Table 4.1 Conditions of simulation cases

Stability	Rx_0 m	Ry_0 m	Rz_0 m	Speed m/s	Temp °C	TG °C/m	RH %	P kPa	TKE m^2/s^2
B	19	16	7	1.2	27.5	-0.84	44.25	100.94	204
D	8	6	2.5	3.2	-5	-0.34	74.92	98.14	232
F	6	5	2.5	0.2	10.2	0.61	98.47	100.35	11.2

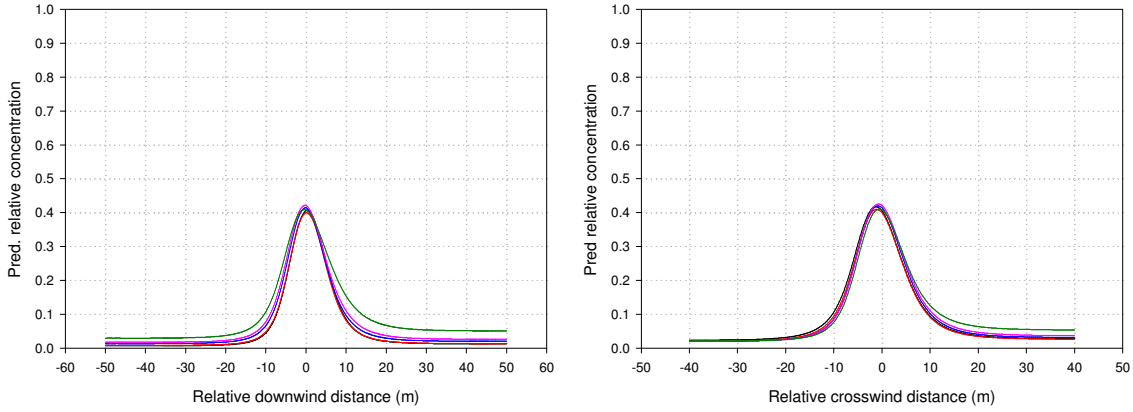
Five dispersion time snapshots from 10 s to 80 s were used in the simulations. This can help to examine the relationship between concentration and dispersion time. Figure 4.2 shows simulation results in the downwind direction, with y=0 and z=0, and in the crosswind direction, with x=0 and z=0, under different stability conditions. They are direct predictions of concentration relative to the maximum concentration in each puff.



(a) Unstable



(b) Neutral



(c) Stable

Figure 4.2 Predicted relative concentration distributions in the downwind (left) and crosswind (right) directions for (a) unstable, (b) neutral and (c) stable conditions

It is obvious from Figure 4.2 that the direct predictions of peak values of relative concentrations in stable conditions is much lower than expected, e.g., only 0.43 was predicted after 10 s. This phenomenon has been reported in Chapter 3 when comparing the predictions and measurements. However, the subsequent simulation results of each case show that the decreases in peak values under different conditions are reasonable. Table 4.2 shows the peak value normalized by the 10-second peak value prediction.

Table 4.2 Peak concentrations at increasing times under different conditions

Stability	Normalized peak concentration				
	t =10 s	t =20 s	t =40 s	t =50 s	t =80 s
Unstable	1	0.78	0.49	0.40	0.26
Neutral	1	0.92	0.74	0.65	0.43
Stable	1	0.98	0.97	0.97	0.96

From Table 4.2, we can find that the peak concentration dropped at the fastest rate in the unstable condition, followed by the neutral condition, and then the stable condition. Taking the 10 s prediction of peak concentration as the reference, after 30 s, the peak concentration dropped to about 49% in unstable conditions, 74% for neutral conditions and still remained 97% of the original value under stable conditions. After 70 s, it was 26% in unstable conditions, 43% in neutral conditions and still 96% remained in stable conditions.

This may imply that the ANN model learned the trend of how concentration changes with time, but not the absolute value. From the analysis in chapter 3, it is already known that the ANN under-estimated high concentrations of cloud points and over-estimated low concentrations.

Another phenomenon noticed from Figure 4.2 is that the predicted dispersion (width of the curve) increased with time both in the downwind and crosswind directions

when in stable conditions (Figure 4.2c). Under neutral conditions, however, the dispersion decreased with time, both in the x and y directions, i.e., the puff shrank with time. The simulation in unstable conditions showed a similar trend as in stable conditions in the downwind direction. The crosswind dispersion in unstable conditions was close to the result for the neutral conditions. The main reason may be that there were far fewer data points contained in the data set from neutral and unstable conditions, because clouds moved fast horizontally (under very windy (neutral) conditions) and vertically (in unstable conditions) and were out of the scanning window of the lidar very quickly. Consequently, the ANN did not get enough information about dispersion under neutral and unstable conditions from the training data. The other reason for the case under unstable conditions is that the cloud expanded quickly vertically, and the concentrations were diluted rapidly. The lidar inversion used in the study could only identify a cloud point when its backscattered signal was at least twice as strong as the background signal (Bissonnette and Roy, 2003). If the cloud quickly mixed with the background, or cloud backscattered signals were not strong enough, then no cloud information would be recorded. As a result, some cloud information would be lost and the puff would appear truncated. This possibility was very small for the case under stable conditions. There, cloud points usually had much higher concentrations at the boundaries, so the backscattered signals were so strong that there was no problem for the lidar to detect the cloud signals. There was a much lower chance of losing cloud information in stable conditions than in unstable conditions.

Simulations in stable and unstable conditions also showed that the peak concentration values appeared at the centre of mass initially, moved slightly to the

upwind direction when dispersion time increased, and the downwind dispersions showed tails in the downwind direction. However the skewness of the dispersion shape was very small.

In order to better compare the variation trend, the predictions were normalized by the prediction of the centroid value at 0.5 s which was treated as the centroid concentration of the source. Figure 4.3 shows 3-D concentration distributions of the ANN predictions along plane RelZ = 0 under unstable, neutral and stable conditions at t=20 s and 50 s, respectively. It can be seen that the distributions in the downwind and crosswind directions look Gaussian in shape, although all of them have long tails along each axis, which will be shown in later 2-D profiles, with nonzero concentrations at positions far from the centroid.

Figure 4.4 shows distribution regressions in the downwind and crosswind directions at different times under unstable conditions, and Figure 4.5 shows dispersions under stable conditions. Both Lorentzian and Gaussian regressions are good fits to the predictions, similar to results found by DeVito (2000). The high concentration predictions of the Lorentzian and Gaussian regressions were very close, but Lorentzian predictions of the low concentrations were closer to ANN predictions than were the Gaussian predictions, therefore the Lorentzian regressions showed higher correlation coefficients. The Gaussian and Lorentzian regressions are shown as follows:

$$\text{Gaussian regression:} \quad y = a\exp\left[-0.5\left(\frac{x-x_0}{b}\right)^2\right], \tag{4.1}$$

$$\text{Lorentzian regression:} \quad y = \frac{c}{1+\left(\frac{x-x_0}{d}\right)^2}. \tag{4.2}$$

(a) Unstable

(b) Neutral

(c) Stable

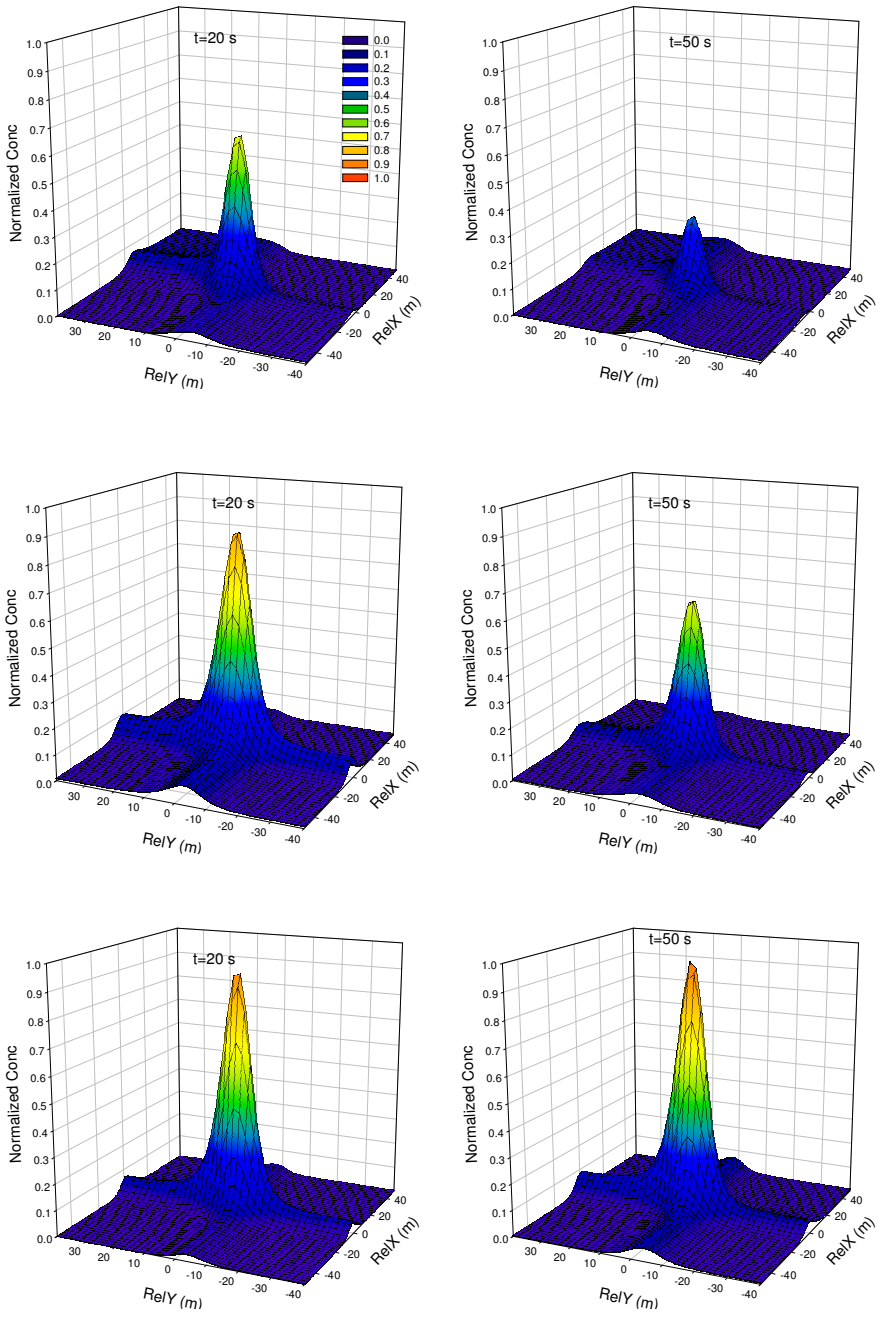
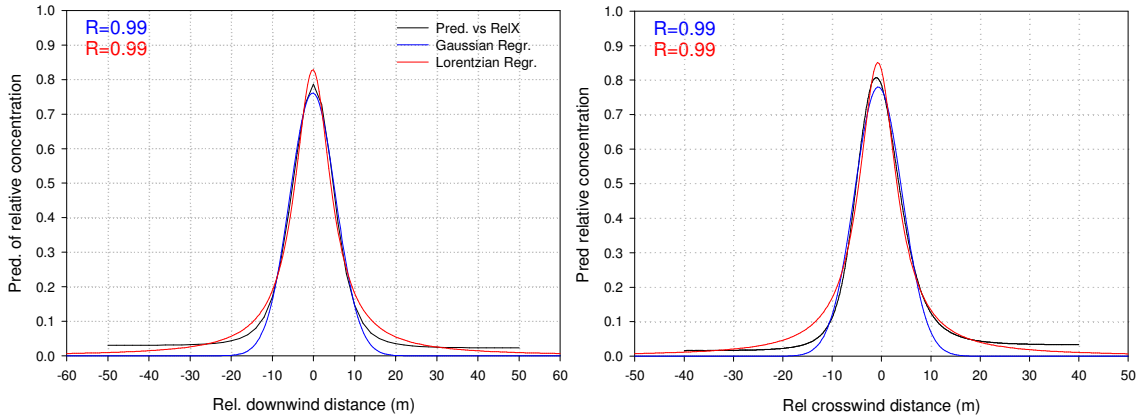
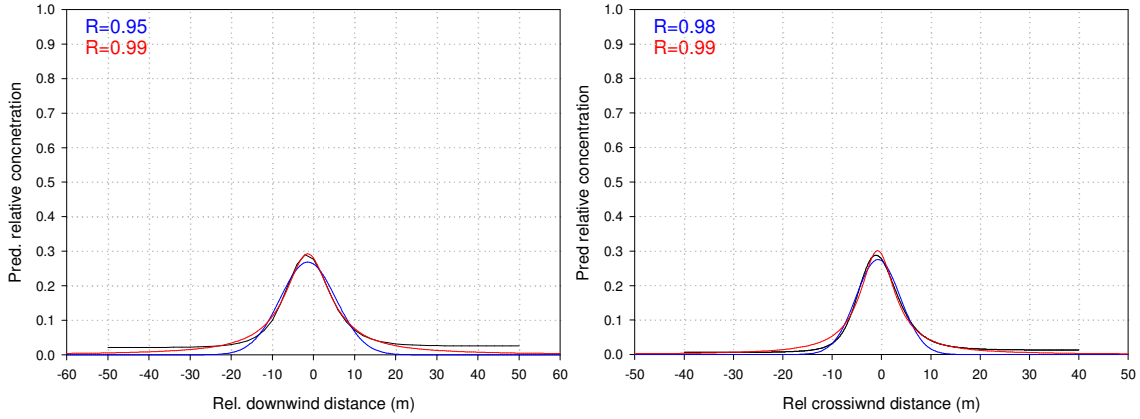


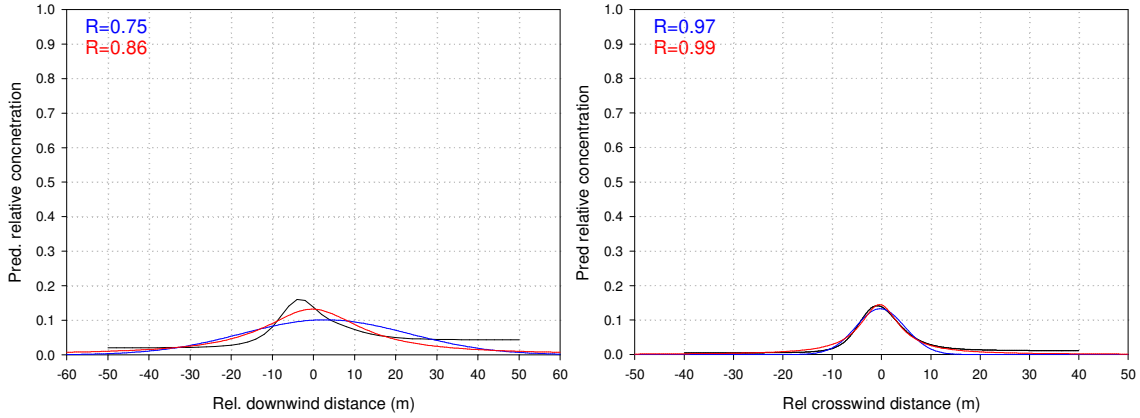
Figure 4.3 Predicted concentration distributions in the downwind and crosswind directions at 20 s (left) and 50 s (right) under (a) unstable, (b) neutral and (c) stable conditions



(a) t=10 s



(b) t=50 s



(c) t=80 s

Figure 4.4 Predicted concentration distributions in x (left) and y (right) directions in unstable conditions

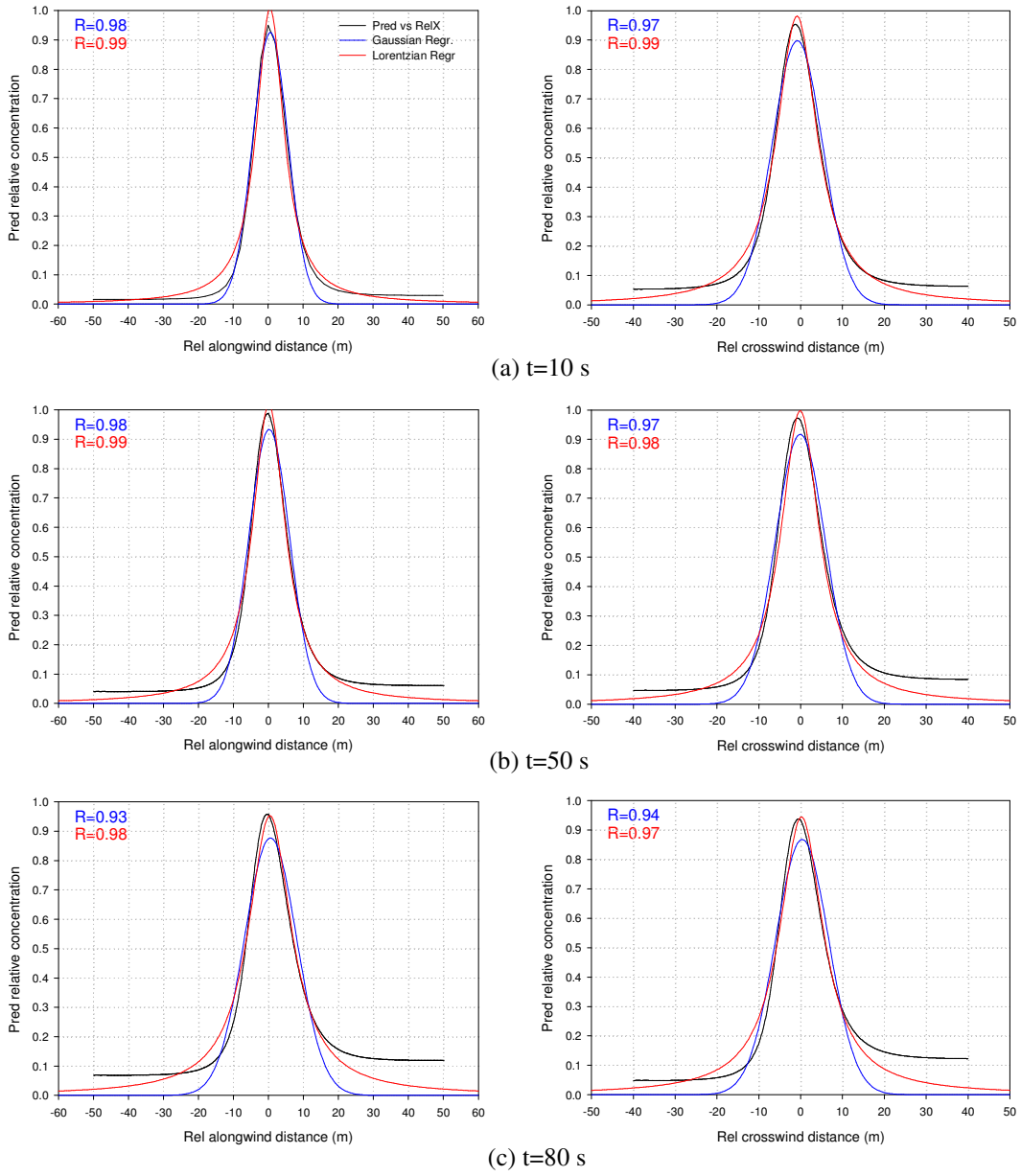


Figure 4.5 Predicted concentration distributions in x (left) and y (right) directions in stable conditions

The main difference between Lorentzian and Gaussian regressions is that the Gaussian regression rapidly approaches to zero as distance from the centre of mass increases, while the Lorentzian curves tend to zero much more gradually, which is

consistent with the mathematical expressions of the Gaussian and Lorentzian regressions. Lorentzian predictions drop at a faster rate initially than Gaussian predictions, but the Lorentzian rate of decrease decreases with increasing distance from the centroid, while the Gaussian rate of decrease increases with distance.

The long tails which are characteristic of Lorentzian curves are not physically realistic. By comparison, Gaussian curves are closer to the observations from field trials. The behavior of the tails of the ANN's predicted distributions can be largely attributed to the number of clear air data points included in the data set. Data points with zero concentration were incorporated into the data set to help the ANN learn the boundary of the puffs. DeVito (2000) has shown that if more clear-air data points near the puff boundaries were included in the data set, then the ANN models would learn more information of boundaries, and better predict distributions that more rapidly approach zero away from the puff centroid. However, maintaining the balance of the output frequency distribution places a strict constraint on the number of such points that can be included in the data set.

It can be seen in Figure 4.4 that the width of the distribution in the downwind direction (σ_x in the Gaussian puff model) increases with time, and the rate of increase also increases with time. It is the same for the case in stable conditions (Figure 4.5), but the increase rate is much slower than that in unstable conditions. It seems that there is not much change in width of the Gaussian distribution in the crosswind direction (σ_y in the Gaussian puff model) both in unstable and stable conditions. The regression results show the change of width in crosswind direction, but the change is very small compared with the change in the downwind direction. The widths of the distributions in the downwind

(*x*) and crosswind (*y*) directions under stable and unstable conditions are shown in Table 4.3.

Table 4.3 Width of Gaussian distributions of simulations in unstable and stable conditions

Disp. Time (s)	Unstable		Stable	
	σ_x	σ_y	σ_x	σ_y
10	5.59	4.84	5.06	6.06
50	6.80	4.50	5.86	5.99
80	20.30	4.98	7.51	6.37

Table 4.3 also demonstrates that in unstable conditions, $\sigma_x > \sigma_y$, and that this inequality increases with dispersion time. A similar relationship can be seen to exist in stable conditions, but the differences between σ_x and σ_y are much smaller than those in unstable conditions.

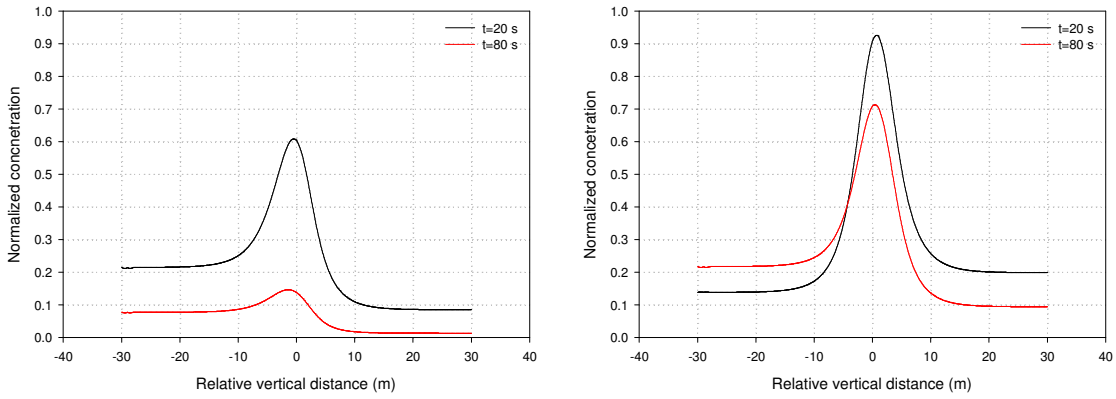
It is very interesting to notice that σ_y decreases a small amount from 10 s to 50 s, both in unstable and stable conditions, so this may be a sign that the developed ANN model may have been somewhat over-trained, in that it might have remembered the data for similar conditions. However, in the long run, the trend is reasonable that σ_y increases with dispersion time both in unstable and stable conditions..

Another conclusion from Tables 4.3 and 4.1 of the initial conditions of the simulations is that it seems that the predicted puff width has only a weak relationship with the input of initial size of the puff. The initial puff size in unstable conditions was 19 m, 16 m and 7 m in the downwind, crosswind and vertical directions, respectively, but 6 m, 5 m and 2.5 m in stable conditions. Interestingly, the predicted dispersion coefficients in the horizontal plane for dispersion time up to 50 s showed little difference between stable and unstable conditions. This may indicate that in the ANN learning process, the

initial puff size contributed little to the output. This can be further demonstrated by the sensitivity analysis shown in Figure 4.1 where change of the initial size caused little change in the output.

4.3 Vertical Concentration Distributions

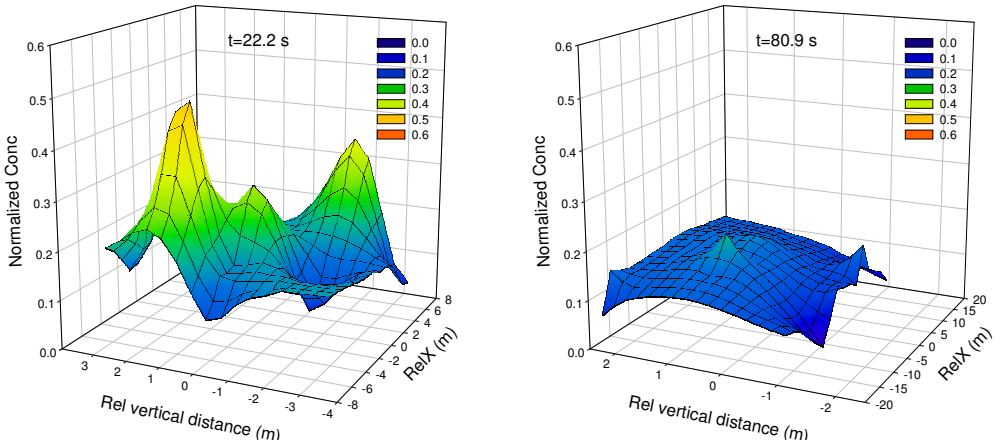
The vertical distribution is somewhat complicated. Unlike the case of horizontal distributions, where the predictions of concentrations for points far from the centre of mass are close to zero, the predictions of points vertically far from the centroid do not show the same trend, but are far higher than zero, as shown in Figure 4.6. This may be due to the fact that, in the data set for ANN learning, clear air points with zero concentrations vertically out of the cloud boundaries were not included, as shown in Figure 4.7 of observations of the vertical distribution from two cases in unstable and stable conditions, respectively. The ANN may have learned from the data set that the concentrations at the vertical boundaries are not close to zero. Therefore it is not surprising to see the prediction as shown in Figure 4.6.



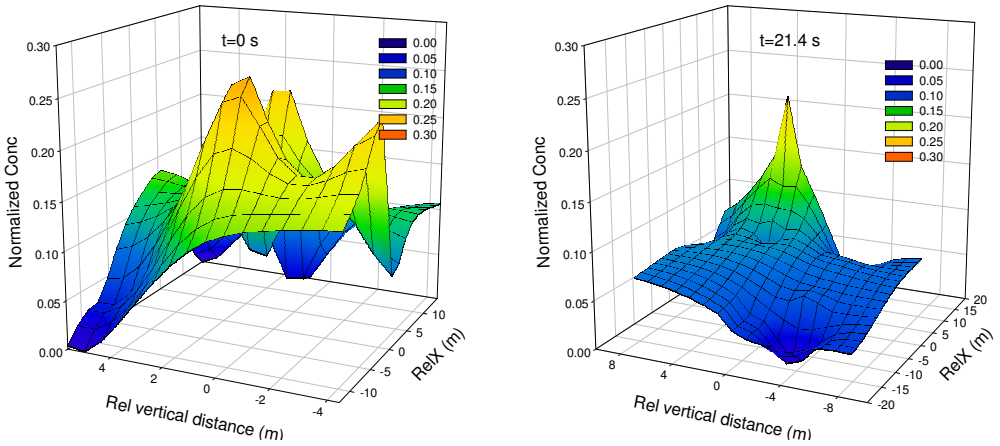
(a) Unstable (b) Stable

Figure 4.6 Prediction of vertical distribution in unstable and stable conditions

Another reason for the high value predictions at large vertical distances from the centre is that the simulations were done for vertical distances from -30 to 30 m. Under stable conditions, the vertical dispersion only extended to several meters, therefore it is an extrapolation for the ANN to provide predictions further away from the centroid. As already mentioned, an ANN's ability to extrapolate is really questionable.



(a) Stable



(b) Unstable

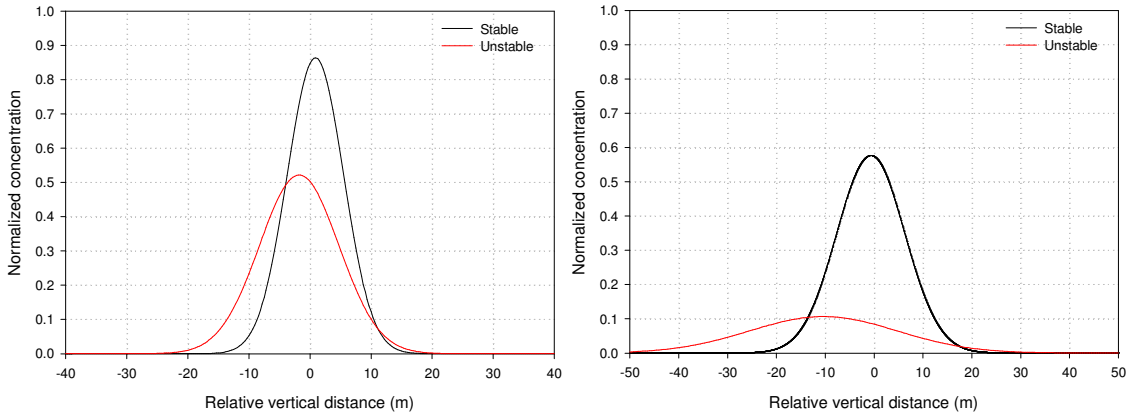
Figure 4.7 Observations of vertical distributions in unstable and stable conditions

Another phenomenon noticed in Figure 4.6 is that the predicted concentrations in positions below the centroid show higher concentrations than for locations at equal

distances above the centroid under unstable conditions at any time. For stable conditions, the predictions show the same trend, as dispersion time increases. There are two possible reasons for this. One is that the lidar scanned cloud points with a raster scanning pattern from bottom to top. Consequently, it took a few seconds to finish one sweep and then start another sweep, so the scanning of higher sweeps is several seconds later than the lower sweeps. As a result, backscattered signals from the bottom part of the cloud were received a few seconds earlier with less expansion and thus relatively higher concentrations. Also the cloud dispersed faster vertically in unstable conditions than in stable conditions, so there were more sweeps needed for cloud coverage, and so more time differences between the top and bottom part of the scanning, which caused the clear difference of concentrations at the bottom and top part of the cloud. Another reason might be due to the deposition of aerosols. Though the average size of particles was small, larger particles would still be expected to settle towards the cloud bottom more quickly than smaller particles, causing a higher concentration distribution in the bottom than in the top of the cloud.

Distributions shown in Figure 4.6 were the results of simulations in the same range both in unstable and stable conditions. However, it is well known that the size of a cloud in stable conditions is much smaller than that in unstable conditions. By observation, the radius of a cloud in stable conditions is usually less than 5 m, the simulation of up to 30 m would require the ANN to extrapolate. As already mentioned, ANNs are good at generalization rather than extrapolation (Patterson, 1996; NeuralWare, 2003). The regression of predictions in stable conditions should not include points far from the centroid. If only predictions of cloud points close to the centroid were used in

stable conditions, then the Gaussian regression gave a good fit for vertical distribution, as shown in Figure 4.8.



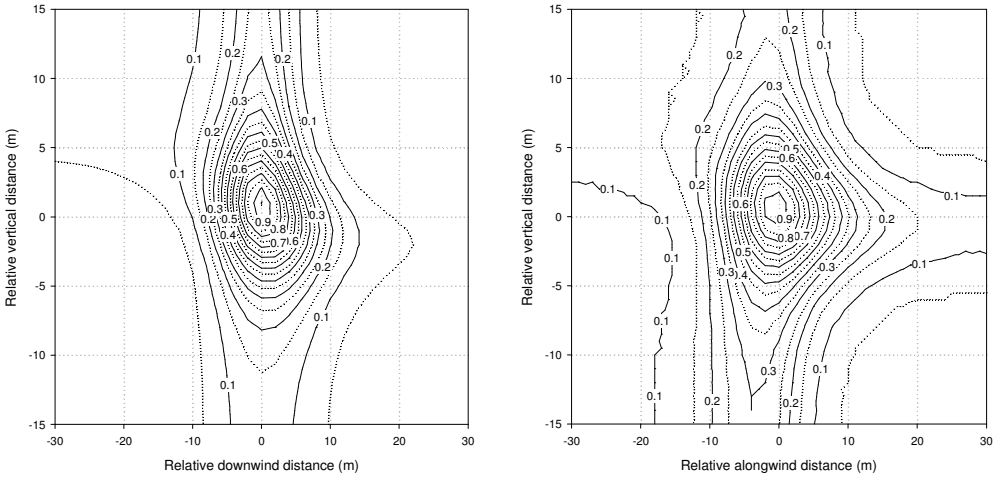
(a) t=20 s (b) t=80 s

Figure 4.8 Gaussian regression of vertical distributions

Generally speaking, cloud points above the centroid move faster than points below it, because wind speed increases with height. This should be reflected by the contours of concentration distributions over the downwind-vertical plane, where the contours should incline toward the downwind direction with height. This shear effect is mostly obvious in stable conditions. Figure 4.9 shows contour plots of a concentration distribution under stable conditions with wind speed 0.4 m/s at dispersion times of 10 s (left) and 80 s (right) over the vertical plane (y=0). Since the wind is weak, the shear effect is not very evident, but from the contour shape we can still find that somewhat skewed shape of the concentration distribution more towards the downwind direction at the top than bottom. The longer the dispersion time, the more indication of wind shear influencing the concentration distribution.

The wind shear influence can also be seen by checking the change in location of the centre of mass with height. Wind speed increases with height above ground, and the

shear influence on puff dispersion is closely related to the air stability. Under stable conditions, the vertical dispersion is weak, and wind shear shows a stronger influence on the concentration distribution, with cloud points at higher layers moving faster than at the lower layers. As a result, the local positions of the centre of mass of higher layers are farther downwind than those of lower layers. In unstable conditions, large thermal eddy movements destroy the wind shear effect, and the position of the centre of mass in different layers shows little change. The simulation results, shown in Figure 4.10, demonstrate these trends.



(a) t=10 s (b) t=80 s

Figure 4.9 Predicted concentrations along the vertical plane *y*=0 under stable conditions

Since the predictions close to the centre of mass are more reliable, only local centroids at heights from -5 m to 5 m relative to the puff centre of mass are compared and shown in Figure 4.10, which is based on predictions at 40 s under stable and unstable conditions. In the stable condition, the downwind position of the local centroid changes about 11 m when the height changes 8 m, while in the unstable condition, a 10 m change of the height only results in a 1.4 m change of the downwind position of the centroid.

This is another illustration that the trained ANN has captured or learned the effect of wind shear on puff dispersion, coupled with the effect of stability.

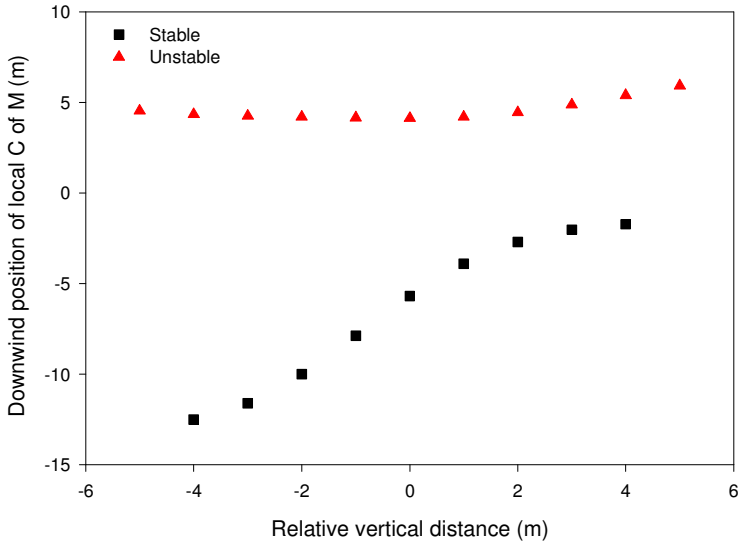


Figure 4.10 Wind shear effects under stable and unstable conditions. The position of the local or layer centre of mass (C of M) relative to the cloud C of M in the downwind direction are shown with respect to vertical distance above and below the cloud C of M.

The ANN predictions show somewhat skewed shapes of concentration distributions in the vertical plane (Figure 4.9), and the wind shear influence on downwind dispersion can be accounted for by simply displacing the puff centroid downwind a bit (Sato, 1995; Wilson, 1981). However, Figure 4.9 shows that the change of centroid is very small (only a few meters), it can be ignored when using a Gaussian distribution.

4.4 Conclusion

Three artificially generated cases under stable, neutral and unstable conditions were examined using the developed ANN model to study the puff concentration distributions. The simulations show that concentration distributions in the downwind and crosswind directions can be well represented by Gaussian and Lorentzian curves. Though

Lorentzian curves show better fits to the predictions, the long tails predicted in the downwind and crosswind directions do not reflect reality. Comparatively, Gaussian regressions appear better choices, even though they might underestimate the amount of dispersion.

The ANN developed significantly underestimated the peak concentrations in stable and unstable conditions, however when normalized by the estimated source peak value, the predictions showed reasonable trends of how peak concentrations changed with time. It showed little change in stable conditions, while the peak value dropped at a rapid rate in unstable conditions. The peak concentration in neutral conditions dropped at a medium rate.

In all conditions, Gaussian regressions provided good fits to ANN predictions. However, points close to the centroid and with high concentrations contributed significantly to the shape of the Gaussian regression in stable conditions. Moreover, under unstable conditions, the shape of the regression was again determined by the high concentration points close to the centroid for short dispersion times. As those times increased, the difference between high and low concentration points got smaller, with the result that the form of the regression broadened or the dispersion coefficients increased.

The ANN predictions in stable and unstable conditions are more reasonable than those in neutral conditions, because more data in the data set were under stable and unstable conditions. Not enough data were included for neutral conditions because clouds moved very quickly out of the scanning range of the lidar when the wind was strong. When comparing dispersion predictions in stable and unstable conditions, the ANN clearly showed the increase of cloud size in the downwind direction with time, but the

size increase was not obvious in the crosswind direction, which means the downwind dispersion was faster than crosswind dispersion. This is reasonable when considering the wind shear effect, which accelerates the downwind dispersion.

The ANN predictions of concentration distributions in stable conditions were more reliable than those in unstable conditions. Usually, cloud signals in stable conditions were strong, so all the points in stable conditions were kept in the training set and therefore full information of dispersion in stable conditions was saved. However, for cloud points in unstable conditions, a lot were of very low concentration, and in order to maintain the balance of the data in the concentration space, many of the low concentration points were removed. Consequently, it is very possible that some dispersion information was missed by removing so many points that may have been related to the boundaries of puffs. In addition, more cases in stable conditions were retained than cases in unstable conditions, so ANN generalizations in stable conditions should be better than those in unstable and neutral conditions.

The ANN also predicted the higher concentrations at locations below the puff centroid than at locations above it in both stable and unstable conditions. This could be a result of the slow deposition of large aerosol particles and the non-instantaneous scanning by the lidar at different heights.

The ANN simulations also showed the influence of the wind shear effect on puff dispersion. In simulations, the location of the centroids of individual sweeps changed with height. The higher the cloud slice, the more downwind the corresponding local centroid was located. In stable conditions, the change of the centroid in the downwind direction can be approximately expressed by a linear relationship with the relative

vertical distance, while the change of centroid location is not obvious in unstable conditions (Figure 4.10).

Though the concentration distribution in the vertical direction is complicated theoretically, the simulations still showed that a Gaussian regression gave a relatively good fit for points close to the puff centroid (Figure 4.8). The ANN predictions showed that it would take an infinite time for puff concentration to fall to zero vertically (Figure 4.6), which is definitely not the case in reality. The cause of this result is likely due to the lack of points with zero concentration outside vertical boundaries, so the ANN did not catch enough information of the vertical boundaries.

Overall, the ANN model has been able to generate simulations that are consistent with observations and that provide insights into the progress of puff dispersion over time under a variety of stability conditions.

Chapter 5
Modelling Dispersion Coefficients for Gaussian Puff Models

The simulations in Chapter 4 have shown that concentration distributions in the downwind, crosswind and vertical directions can be approximated quite well by Gaussian distributions.

In a Gaussian puff model, the concentration within a puff falls off from the centre according to a Gaussian distribution. Taking the x-axis along the direction of mean wind flow U (m/s), the y-axis as crosswind and the z-axis as vertical, the Gaussian puff model has the following form:

$$c(x,y,z,t) = \frac{Q}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z}\exp\left\{-\frac{(x-Ut)^2}{2\sigma_x^2}-\frac{y^2}{2\sigma_y^2}-\frac{z^2}{2\sigma_z^2}\right\}, \qquad (5.1)$$

which is a modification of equation (1.3), employing the assumption that the downwind centre of mass is only influenced by mean wind transport, c is the aerosol concentration (g/m^3), Q is the mass of aerosol released (g) at time t=0 (s), and the source location is taken to be the origin. The spread of the cloud is determined by the dispersion coefficients, σ_x, σ_y and σ_z (m).

Dispersion coefficients are the key fitting parameters of Gaussian models. Although different models may use different dispersion coefficients, e.g., Slade and COMBIC, dispersion coefficients in most models are usually parameterised by atmospheric stability, distance from the source and/or dispersion time (Slade, 1968; Pasquill and Smith, 1983; Ayres and Desutter, 1995).

Few parameterisations exist for puff diffusion, and many models simply employ the well-known Pasquill-Gifford parameterisation scheme, developed for plume spread (Pasquill and Smith, 1983). Slade (1968) pooled a number of earlier puff diffusion experiments to construct empirical relations to determine the dispersion coefficients based on the Pasquill stability class and downwind travel distance. The Slade parameterization takes the form of a power law with the coefficient and exponent varying with Pasquill stability class. Slade's formulae were based on far fewer observations than were the continuous plume formulae of Pasquill and Gifford (Turner, 1994), with the assumption that the longitudinal dispersion coefficient σ_x equaled the lateral dispersion coefficient $\sigma_{y.}$ This is usually not true, as Pasquill (1983) discussed that the downwind dispersion coefficient is larger than the lateral dispersion coefficient, due to the effects of wind shear. Further study of puff dispersion shows that downwind dispersion is a combined effect of vertical dispersion and wind shear (Chatwin, 1968; Draxler, 1979; Wilson, 1981; Van Ulden, 1992; Sato, 1995).

The longitudinal variance of a puff can be expressed as

$$\sigma_x^2 = \sigma_{xs}^2 + \sigma_{xt}^2 , \tag{5.2}$$

where the parameters σ_{xs} and σ_{xt} are the downwind variance caused by the vertical wind shear and turbulent longitudinal diffusivity of a puff, respectively (Draxler, 1979).

If a puff is from a point source, expression (5.2) can be used directly. Unfortunately, most field experiments or accidental releases are not point sources, but volumetric sources with initial sizes, for which expression (5.2) can be modified as

$$\sigma_x^2 = \sigma_{x0}^2 + \sigma_{xs}^2 + \sigma_{xt}^2 , \tag{5.3}$$

where σ_{x0} represents the initial puff expansion in the longitudinal or downwind direction.

This study of puff dispersion follows Eqn. (5.3), except using one parameter σ_{xts} to represent the influence of turbulent diffusivity and vertical wind shear together.

COMBIC also uses Eqn. (5.3) for downwind travel times greater than 30 s. When the downwind travel time is less than 30 s, a traditional power law is used, but with coefficients different from Slade's expressions. The Gaussian dispersion coefficients for the puff model used in COMBIC are dependent on the atmospheric stability (P-G stability category), scaling ratio (z/L) and surface roughness length (z_o) for downwind distances greater than the distance associated with a downwind travel time of 30 s. The downwind dispersion coefficient σ_x is related to the vertical dispersion coefficient. The lateral dispersion length σ_y is time dependent and responsive to changes in surface roughness length, while the vertical dispersion length σ_z is independent of the downwind travel time, but is dependent upon the surface length. All dispersion coefficients are functions of downwind distances (Ayres and DeSutter, 1995).

The dispersion coefficients discussed above were based on statistical observations under some specific conditions, and are functions of Pasquill stability categories. The use of Pasquill stability can simplify the model, although each Pasquill stability category covers a wide range of condition. Therefore, the predictions based on a simple stability parameter may not be good enough for practical use. In this study, a neural network model of dispersion coefficients was developed based on the 4 week-long field experimental data, to attempt to predict the more appropriate dispersion coefficients for Gaussian puff dispersion model for arbitrary meteorological conditions.

5.1 Dispersion Coefficients from ANN Models

The dispersion coefficients σ_x, σ_y and σ_z were calculated from experimental data by using moment analysis. The moment for concentration is defined by the following relation:

$$\theta_{nmp}(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} cx^n y^m z^p \mathrm{d}x\mathrm{d}y\mathrm{d}z\,, \quad (m,n,p \geq 0) \tag{5.4}$$

In this study, a moving coordinate frame was used. The origin of the Cartesian coordinate system was at the centre of mass of the puff. Dispersion coefficients in each direction are the variance of the concentration distribution in the same direction, and expressed as

$$\sigma_x{}^2(t) = \frac{\theta_{200}}{\theta_{000}} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 c\mathrm{d}x\mathrm{d}y\mathrm{d}z}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c\mathrm{d}x\mathrm{d}y\mathrm{d}z}. \tag{5.5}$$

Similar expressions were used for σ_y and σ_z.

In all, 60 releases, for a total of 167 scans containing almost the whole cloud information, were selected for the ANN modelling. The original cloud records were processed through lidar inversion, using the inversion technique proposed by Bissonnette and Roy (2003), as introduced in Chapter 2. The directly detected backscattered signals were inverted into concentration-related ratios of backscattering coefficients of the cloud to background. For calculating the variance of concentration distributions, no absolute concentrations were needed.

A sensitivity analysis of the ANN concentration distribution model has shown that initial cloud size and micrometeorological variables did not contribute significantly to the concentration. Therefore, only some general meteorological variables which are

easily measured were selected as inputs. Nine inputs were used in the ANN modelling of dispersion coefficients, as shown in Table 5.1. These variables were selected because they can be easily measured by general weather systems and reflect the stability of the atmosphere. For example, high temperature and high insolation usually correspond to mid-day in summer, and a negative temperature gradient indicates that the temperature at a higher position is lower than temperature at a lower position, and therefore particles will move upward due to buoyancy and extend the cloud into a larger size vertically. The larger the magnitude of the temperature gradient, the more unstable is the atmosphere, and the greater is the vertical dispersion. Pressure is usually related to cloud cover. A clear sky is usually associated with high pressure, and a lower pressure indicates a cloudy or partly cloudy sky.

Table 5.1 Inputs for ANN model and data range

Inputs	Description	Range	
t	Dispersion time, s	16.8	137.2
U	Mean transport wind speed, m/s	0.12	4.25
W_1	solar irradiation, W/m^2	0.0	856.2
W_2	Ground heat flux, W/m^2	0.0	442.7
T	Temperature, °C	-19.0	27.5
TG	Temperature gradient, °C/m	-0.98	1.70
RH	Relative Humidity, %	27.5	98.23
RHG	Relative Humidity gradient, %/m	-8.46	4.88
p	Air pressure, kPa	98.07	101.05

Since the change of dispersion coefficients with dispersion time and stability variables is significant, concentration variances in each direction were selected as the outputs of ANNs. As with the ANN development for aerosol concentration, where the first available scan containing the whole cloud was taken as source, the size of the source was different under different conditions. The influence of source size was removed by subtracting it from the final variance, i.e.

$$\sigma_{id}^2 = \sigma_i^2 - \sigma_{i0}^2,\ i = x, y, z, \tag{5.6}$$

where σ_i^2 is calculated by Eqn.(5.5), and σ_{id}^2 is variance without influence of source size . For convenience, σ_{id} will be replaced by σ_i in later description.

As with ANN modelling for concentration distributions, all inputs and outputs have to be transformed linearly or nonlinearly to make the training space more evenly distributed. Figure 5.1 shows the frequency distributions of all nine inputs and outputs σ_x, σ_y and σ_z after transformation.

Overall, 15% of all the data were selected for the test set, with the remainder being used for the training set.

ANN training started with randomly-assigned connecting weights, with the weights being adjusted during the training process until optimal performance was achieved. Due to the randomness of the learning process, training of each model started with a different random seed, i.e., a different random weight distribution to identify different traits of the training data. Thus it is believed that an average of predictions from different trained ANN models would be more effective than from only one ANN model.

Five models were trained for each σ_i, i = x, y and z, and all five models were deployed to simulate the cloud size variation with each input. The average of predictions from 5 models was used as final predictions of simulated conditions. The statistical evaluations of the ANN models are summarized in Table 5.2.

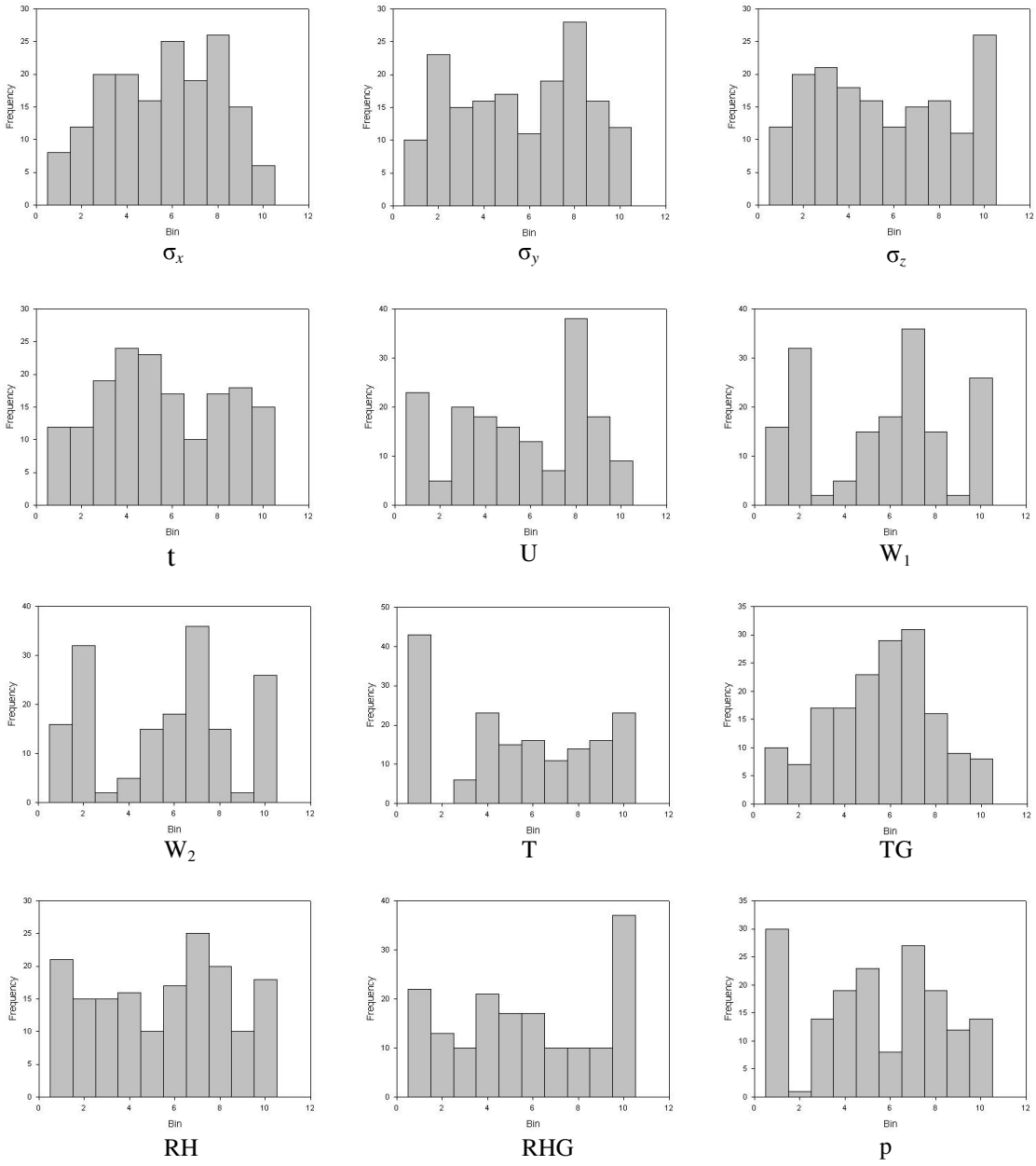


Figure 5.1 Frequency distributions of all inputs and outputs after transformation

Table 5.2 Evaluation of ANN models for dispersion coefficients

MODEL	set	NMSE	R	FB	F2	MG	VG
σ_x	trainset	0.056	0.928	-0.069	0.979	0.906	1.065
	testset	0.042	0.912	0.014	1.000	1.036	1.044
σ_y	trainset	0.030	0.962	0.001	1.000	0.976	1.018
	testset	0.016	0.972	0.027	1.000	1.005	1.020
σ_z	trainset	0.021	0.979	-0.007	1.000	0.982	1.021
	testset	0.033	0.975	-0.017	1.000	1.003	1.049

NMSE is normalized mean square error; R is linear correlation coefficient, FB is fraction bias, F_2 is proportion of prediction within a factor of 2, and MG and VG are the geometric mean and variance, respectively. The definitions of above parameters are as used in Chapter 3.

The best values for all those performance measures are:

$$NMSE = FB = 0,$$

$$R = F2 = MG = VG = 1.$$

Table 5.2 shows that ANN predictions are very close to the measurements, with a very high correlation over 0.9 for all three dispersion coefficients, and almost 100% of predictions are in the range of a factor of two. This means that the developed ANNs are good for predicting puff expansion when the atmospheric conditions and dispersion times are in the range of the ANN training set.

5.2 Model Comparison

The trained ANN models were used for estimating the puff dispersion coefficients in a Gaussian model and the predictions of relative concentrations are compared with those of both Slade's Gaussian puff model and COMBIC. Three validation sets are used (under unstable, neutral and stable conditions). Table 5.3 shows the atmospheric conditions of each validation set.

Table 5.3 Validation sets conditions

Trial	Release time	U (m/s)	W_1 (W/m^2)	W_2 (W/m^2)	T (°C)	TG (°C/m)	P (kPa)	RH (%)
A-B Unstable	16:52	1.31	449.5	76.3	27.4	-0.84	100.9	43.9
E-D Slightly stable to Neutral	5:40	0.54	9.3	5.8	10.8	-0.1	100.2	97.4
F Stable	23:20	0.58	9.3	5.1	11.1	0.85	100.5	94.85

As ANN models only consider expansion from the initial size, the final expansion coefficients in the Gaussian puff model will be

$$\sigma_i^2(t) = \sigma_i^2(0) + \sigma_{ip}^2(t), \tag{5.7}$$

where σ_{ip} is the average of σ_i predictions from the ANN models designed to generate dispersion coefficient predictions.

Table 5.4 lists the performance of an ANN model directly predicting concentration distribution (as discussed in Chapter 3 and 4), a Gaussian puff model with dispersion coefficients predicted by ANN models (Puff-ANN), a Gaussian puff model with Slade's coefficients and COMBIC for all points including clear air points outside the boundaries. It can be seen that the ANN model for direct concentration prediction gives the best prediction for most conditions, while COMBIC is the worst because it greatly underestimates the higher concentrations (high positive FB). If F_2 is considered, the ANN is still the best, and in most conditions, the performance of the Gaussian puff model with ANN parameterizations is very close to the performance of Slade's Gaussian puff model. For all 4 validation sets, COMBIC gives extremely high NMSE values, which means that most COMBIC predictions are far from the measurements.

Table 5.4 Model comparison for all points (including clear air points outside boundaries)

Validation P-G category	MODEL	NMSE	R	FB	F_2	F_{10}
A-B Unstable	ANN	3.85	0.63	-0.92	0.18	0.35
	Puff_ANN	4.13	0.55	-0.73	0.16	0.36
	COMBIC	159.60	0.47	1.74	0.05	0.16
	Slade's puff	11.87	0.40	0.27	0.15	0.38
E-D Slightly stable to Neutral	ANN	2.51	0.61	-0.79	0.24	0.49
	Puff_ANN	2.25	0.62	-0.11	0.21	0.44
	COMBIC	57.91	0.64	1.71	0.19	0.32
	Slade's puff	2.32	0.67	-0.19	0.25	0.43
F Stable	ANN	2.82	0.60	-0.92	0.20	0.37
	Puff_ANN	2.52	0.64	-0.70	0.18	0.37
	COMBIC	16.68	0.67	1.15	0.17	0.40
	Slade's puff	6.13	0.60	-0.58	0.16	0.38

ANN models are not good at extrapolating. To better compare the performance of puff dispersion models, evaluations of models based on nonzero concentration points should be more reasonable. Table 5.5 is the results of the evaluation of models for nonzero points.

Table 5.5 Models comparison for all nonzero concentration points

Validation P-G category	MODEL	NMSE	R	FB	F2	F10
A-B Unstable	ANN	1.49	0.68	-0.45	0.43	0.85
	Puff_ANN	1.93	0.62	-0.17	0.39	0.86
	COMBIC	1.90	0.53	-0.44	0.34	0.81
	Slade's puff	8.34	0.48	0.82	0.36	0.93
E-D Slightly stable to Neutral	ANN	1.16	0.64	-0.46	0.44	0.89
	Puff_ANN	1.31	0.65	0.21	0.33	0.73
	COMBIC	1.14	0.66	-0.04	0.32	0.70
	Slade's puff	1.44	0.65	0.04	0.26	0.55
F Stable	ANN	0.79	0.66	-0.41	0.50	0.94
	Puff_ANN	0.87	0.63	-0.28	0.46	0.93
	COMBIC	2.10	0.67	-0.86	0.37	0.85
	Slade's puff	3.23	0.61	-0.87	0.31	0.84

The same conclusions can be reached from Table 5.5 as from Table 5.4, i.e., that COMBIC is the least effective, based on F_2, F_{10}, FB and NMSE, which indicate how close predictions are to measurements. The Gaussian puff model with dispersion coefficients from ANN models appears less effective than predictions directly from an ANN model, but still better than COMBIC for all cases and better or at least as good as Slade's Gaussian puff model.

It can be concluded, based on the analyses above, that a Gaussian puff model can give reasonable estimations for puff dispersion, as long as the dispersion coefficients are given properly. Predictions of the Gaussian puff model with ANN-based dispersion coefficients is closer to measurements than the Gaussian puff model with Slade's parameterizations, and both Gaussian puff models outperform COMBIC, even though COMBIC is a more complex version of a general Gaussian puff model.

5.3 Puff Expansion

There is no doubt that the size of the cloud increases with dispersion time, and the puff size changes at different rates under different atmospheric conditions. To find out how puff dispersion changes with dispersion time and meteorological variables, ANN simulations based on three artificially generated cases were analyzed. Except for dispersion time, the other eight inputs of the ANN models were selected to represent Pasquill stability categories from unstable to stable. As ANNs should not be used to extrapolate, all vectors used for simulation should be within the space of the training data used in model development. Since the training space is formed by discrete vectors, it is hard to make sure the 9-dimension vector to be simulated is in the trained space. What one can do is make sure each of the inputs is within the training range of the

corresponding variable. The prediction will be forced to have the value of the boundary if it exceeds it.

Table 5.6 summarizes the meteorological conditions for the simulations. Dispersion time is changed from 20 to 120 s, every 20 s, for a total of 6 time instants, and 9 wind speeds are checked, from 0.15 m/s to 4.15 m/s, every 0.5 m/s.

Table 5.6 Simulation conditions for puff dispersions

Case	W_1 (W/m^2)	W_2 (W/m^2)	T (ºC)	TG (ºC/m)	RH (%)	RHG (%/m)	p (kPa)
1	5.4	5.8	10	0.65	98	2.2	100.3
2	280	248	-10	-0.16	55	-1.3	99.5
3	830	110	27	-0.92	28	0.72	101.0

Case 1 is the atmospheric condition at about 1:30 in the morning in August with little insolation. Case 2 is at 17:30 in March with moderate insolation and low temperature, where the ground was covered by snow. Case 3 is at 14:00 in August with high temperature and strong insolation.

Case 1 has a temperature gradient greater than zero, which means it is under stable conditions. Temperature gradients of cases 2, 3 are negative, so they may range from neutral to unstable conditions. Table 5.7 gives the corresponding P-G stability categories for the cases listed in Table 5.6, considering different wind speeds. The determination of P-G stability categories follows the stability definitions used in COMBIC (Ayres and DeSutter, 1995). The combination of different wind speeds and all three conditions covers all the possibilities of P-G stability categories.

Table 5.7 P-G stability under different wind speeds of cases in Table 5.6

Case	Wind speed (m/s)								
	0.15	**0.65**	**1.15**	**1.65**	**2.15**	**2.65**	**3.15**	**3.65**	**4.15**
1	G-F	F	F-E	F-E	E	E	E-D	E-D	D
2	C	C	C-D	C-D	D	D	D	D	D
3	A-B	A-B	B	B	B	B	B	B	B-C

5.3.1 Puff dispersion coefficients and dispersion times

The following analysis involves fixing the wind speed and all other conditions, while changing the dispersion time, to check how dispersion coefficients change with diffusion time.

Figures 5.2-1(a) to 5.2-1(e) are simulations of atmospheric condition or case 1 with different wind speeds. According to Table 5.7, the P-G stability category changes from very stable G to neutral condition D.

The black round symbols in Figure 5.2 represent σ_x, while the black line is its corresponding regression. Red up-triangles represent σ_y and the red line is the regression. The blue squares represent σ_z.

Figure 5.2-1 shows the relationship between puff dispersion coefficients and dispersion time under stable to neutral conditions. They can be approximated by a linear regression of the form:

$$\sigma_i = \sigma_{i0} + kt \text{ , } i=x, y \text{ or } z, \tag{5.8}$$

where σ_{i0} is the variance at t=0, and usually has a nonzero value. But as stated above, σ_i is the variance in the *i* direction, without the initial size of puff, so it should satisfy

$$\sigma_{i0} = 0 . \tag{5.9}$$

To satisfy Eqn. (5.9), it is found that σ_x can be well represented by the power law regression $\sigma_x = at^b$ for stabilities from very stable to neutral. Also, σ_y can be represented by a power law regression under very stable conditions. For other conditions, the relationship between σ_y and dispersion time follows the shape of exponential rise to the maximum, so $\sigma_y = a(1-\mathrm{e}^{-bt})$ from stable to neutral conditions with a very high

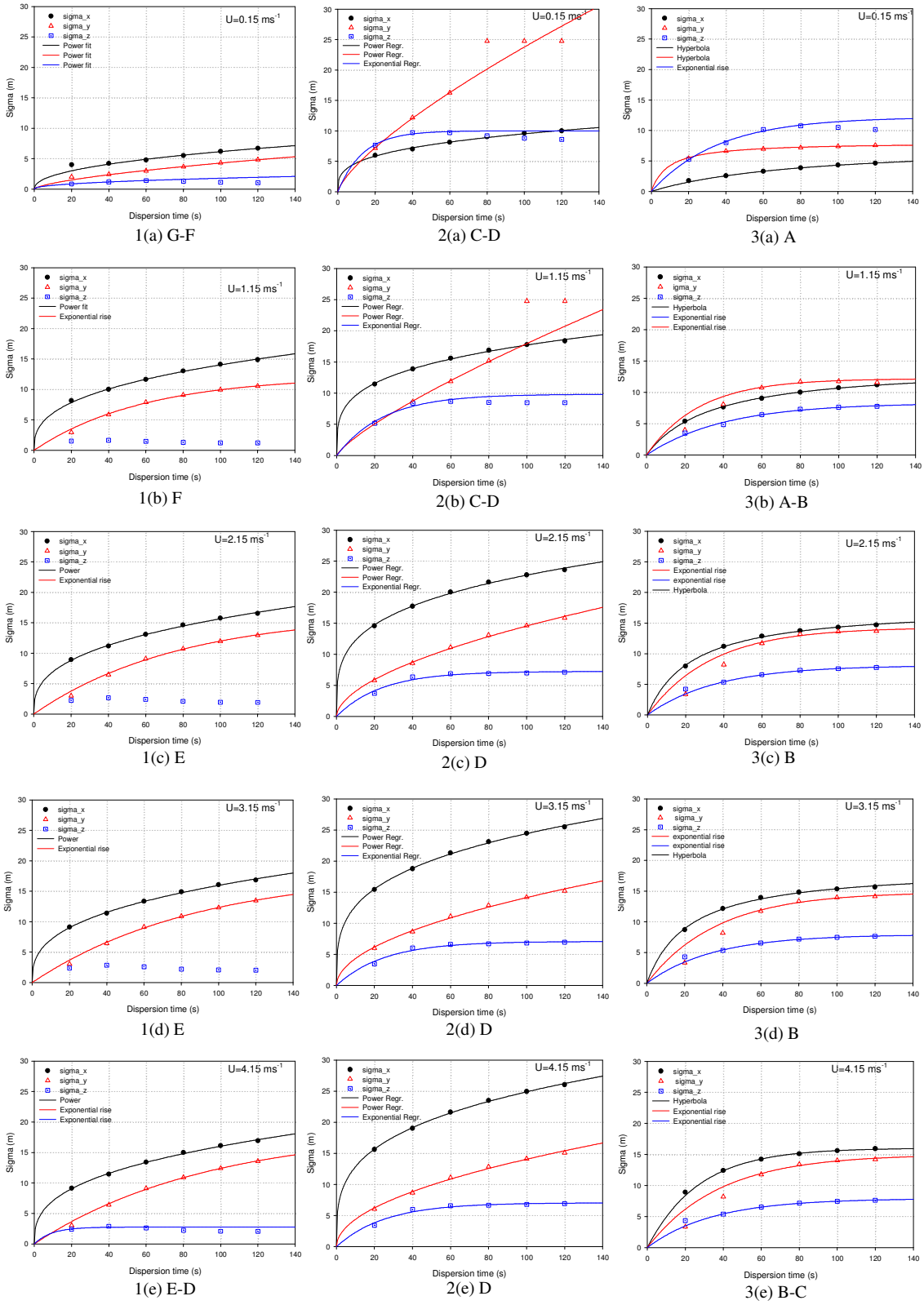


Figure 5.2 Influence of dispersion time on puff expansion. Dispersion coefficients are predicted by an ANN model.

correlation coefficient ($R > 0.99$).

The simulation result of σ_z is more complex, as it increases with time from 20 s to 40 s, but then decreases under stable and near neutral conditions. Intuitively, it is impossible that the puff gets smaller with time, unless part of it has already mixed into the atmospheric background. This may be due to the fact that the lidar scanned the puff cloud slice by slice. Consequently, it seemed unavoidable that a few cloud bottoms and tops were missed. Since the puff dispersed little vertically under stable conditions, even a few missing scans of cloud (extremities) would affect the determination of puff size, and thus lead to a smaller puff (vertically). The decreasing size of the puff with time can be found in training data from some stable conditions, therefore it is not surprising that ANN models predict a slightly decreased size of puff with increasing dispersion time. If the relationship of puff size and dispersion time is known, these data can be modified properly. For the simulation results, the change of puff size in the vertical is very small, so it is reasonable to use an expression of exponential rise to the maximum to represent the relationship of σ_z and dispersion time, i.e., $\sigma_z = a(1 - e^{-bt})$.

Figure 5.2-1 also shows that $\sigma_x > \sigma_y > \sigma_z$ under stable conditions at any simulation time. All σ_x, σ_y and σ_z increase with wind speed, but the rate of increase decreases with increases of wind speed. For case 1, when wind speed increases, the P-G stability category changes from very stable G to neutral D. Therefore the simulation is in agreement with the thought that vertical dispersion of a puff increases when atmospheric stability decreases, as the wind speeds up. When the wind is strong enough, puff size

stays almost unchanged, even as the wind further increases, which is related to neutral conditions with a strong wind.

The influence of dispersion time on puff expansion under the neutral conditions of case 2 with different wind speeds is shown from Figure 5.2-2(a) to 5.2-2(e). P-G stability changes from slightly unstable with very weak wind to neutral when the wind gets strong.

This is also the case under neutral conditions, where σ_x can be well represented by a power law, $\sigma_x = at^b$ When the wind is weak ($\leq$1.15 m/s), the simulation results of σ_y after 60 s exceed the upper limit of the training data, which may due to the fact that vectors are outside of the training space. Therefore, the regression curve of σ_y is based on results from 20 s to 60 s for weak wind (lower than 1.15 m/s). When the wind gets stronger, all predictions are within the range of the training data, and are used for the regression. A power law of $\sigma_y = at^b$ (red lines) gives the best fit (R > 0.99), as shown in Figure 5.2-2. It shows that $\sigma_y > \sigma_x$ when wind is weak no matter what the relationship is initially. But when wind gets a little stronger (greater than 2.0 m/s), it will meet the general relationship between longitudinal, lateral and vertical dispersions that $\sigma_x > \sigma_y > \sigma_z$.

Figure 5.2-2 also shows that under neutral conditions, σ_x increases with wind speed, while σ_y and σ_z correspondingly decrease. For strong wind speeds at the surface, there is a strong wind shear effect, and as a result, σ_x increases with wind speed. The crosswind dispersion is mainly a result of turbulent diffusion. If wind shear is caused by a change of wind speed with height, then wind shear has little influence on σ_y. If all other conditions remain unchanged, but wind speed increases, the stability of the atmosphere

will increase, which results in the decreased vertical dispersion. For case 2, when the wind speed is weak, it is in a transition from slightly unstable to neutral, so the wind shear effect is not obvious, and turbulent diffusion dominates the puff dispersion. Due to the randomness of turbulence, it is possible that σ_y exceeds σ_x. Figures 5.2-2(a) and 5.2-2(b) show that the chance for $\sigma_y > \sigma_x$ depends on how weak the wind is. When wind is strong enough, it shows the general relationship that $\sigma_x > \sigma_y > \sigma_z$. Similarly, when wind is strong enough, the puff shape is little changed.

Atmospheric condition 3 under different wind speeds shows the variable intensity of instability. Surprisingly, only Figure 5.2-3(a) under a very weak wind (0.15 m/s) shows a relatively larger vertical dispersion, or σ_z. When the wind speed increases, the P-G stability category changes from very unstable to moderately unstable, and vertical dispersion quickly drops off to values smaller than horizontal dispersions (σ_x and σ_y). Also, σ_z slightly decreases with wind speed. When wind increases to 1.65 m/s, the general relationship of $\sigma_x > \sigma_y > \sigma_z$ is re-established, in agreement with DeVito's ANN-based simulation developed in unstable conditions (DeVito, 2000). Figure 5.2-3 also provides useful information that the downwind dispersion, σ_x, is close to crosswind dispersion, σ_y, under unstable conditions, which may be due to the fact that vertical dispersion destroys the wind shear influence and therefore turbulent diffusion dominates the horizontal dispersions. Comparing predictions under unstable conditions with those under neutral conditions, the horizontal dispersion coefficients, σ_x and σ_y, are similar in value. However, there is a larger vertical dispersion in unstable conditions than in neutral conditions, especially when diffusion time is longer.

However, the vertical dispersion coefficient σ_z in unstable conditions used in the training set did not provide enough information. In unstable conditions, it was noticed that clouds expanded very quickly vertically. Due to limited lidar elevation, usually after 2 scans, all subsequent scans missed the cloud top, and the lidar was only able to catch signals 3 times or more stronger than that of the background. If part of the cloud has been mixed with the background, then in the records of lidar scanning, no information of these points would be included and as a result the detected puff size would remain the same or even smaller. The degree of aerosol mixing with the atmosphere, however, is directly related to the aerosol mass released. Since only small amounts of aerosols were released in the field trials to form artificial puffs, it is very possible that portions of the vertical dispersion information were lost over the relatively short times of the cloud scanning. Consequently, predictions of vertical dispersions in unstable conditions may be under-estimated by the ANNs.

Similar obesrvations were made that the predictions of vertical dispersions after 60 s decrease with time in case 2 and case 3 when the wind is weak. It is a coincidence that most of the training data under neutral and unstable conditions are under 60 s, so predictions over 60 s need extrapolation by the ANNs and may not reflect the true situation. But when the wind is a little stronger, σ_z increases with time slowly and smoothly.

To characterize the change in puff shape, the ratios σ_y / σ_x and σ_z / σ_x, can be examined. These are shown in Table 5.8. Under stable conditions, stability decreases with wind speed, while under neutral and unstable conditions, stability increases with speed.

Table 5.8 Change of puff shape with stability

Speed (m/s)	Stable		Neutral		Unstable	
	σ_y/σ_x	σ_z/σ_x	σ_y/σ_x	σ_z/σ_x	σ_y/σ_x	σ_z/σ_x
0.15	0.58	0.28	1.74	1.39	2.58	3.11
0.65	0.44	0.10	0.98	1.22	1.67	0.79
1.15	0.59	0.16	0.62	0.61	1.06	0.64
1.65	0.59	0.22	0.52	0.42	0.83	0.53
2.15	0.58	0.24	0.48	0.36	0.73	0.48
2.65	0.57	0.25	0.47	0.33	0.69	0.45
3.15	0.56	0.25	0.46	0.32	0.67	0.44
3.65	0.56	0.25	0.46	0.32	0.66	0.44
4.15	0.56	0.25	0.46	0.31	0.66	0.43

Table 5.8 indicates that σ_z / σ_x tends to decrease with stability in all conditions, and σ_y / σ_x tends to decrease with stability when it is under neutral and unstable conditions, which is in agreement with the study of Nicola (1971). However, σ_y / σ_x tends to increase with stability in stable conditions, which is the opposite of Nicola's study.

As for the relationship between $\sigma_y / \sigma_x, \sigma_z / \sigma_x$ and dispersion time, it is summarized in Table 5.9, where predictions are made under a wind speed of 1.15 m/s for stable and unstable conditions, and of 3.15 m/s for neutral conditions.

Table 5.9 Shape of puff change with time

Time (s)	Stable		Neutral		Unstable	
	σ_y/σ_x	σ_z/σ_x	σ_y/σ_x	σ_z/σ_x	σ_y/σ_x	σ_z/σ_x
20	0.36	0.18	0.39	0.23	0.74	0.65
40	0.59	0.16	0.46	0.32	1.06	0.64
60	0.68	0.13	0.52	0.31	1.19	0.71
80	0.70	0.10	0.56	0.29	1.17	0.73
100	0.70	0.09	0.58	0.28	1.10	0.71
120	0.71	0.08	0.60	0.27	1.04	0.70

Nickola (1971) found that σ_z / σ_x decreases slightly with time in all stabilities, while the simulation of the ANN shows that this is the case under stable and neutral conditions, while under unstable conditions, σ_z / σ_x slightly increases with time. The

simulation results show that σ_y / σ_x increases with time under stable and neutral conditions, and the same for unstable conditions if not considering the simulation after 60 s, which is not in agreement with Nickola's study that σ_y / σ_x decreases slightly with time in stable conditions.

To summarize, the five relationship between puff dispersion and dispersion time can be expressed as:

$$\sigma_x(t) = a_1 t^{b_1}, \quad \text{for all except unstable conditions}$$

$$\sigma_x(t) = a_2 t / (b_2 + t), \quad \text{for unstable conditions}$$

$$\sigma_y(t) = a_3 t^{b_3}, \quad \text{for stable and neutral conditions}$$

$$\sigma_y(t) = a_4 (1 - \mathrm{e}^{b_4 t}), \quad \text{for unstable conditions}$$

$$\sigma_z(t) = a_5 (1 - \mathrm{e}^{-b_5 t}), \quad \text{for all conditions,}$$

where coefficient a_i and b_i, $i = 1, 2, 3, 4, 5$ might be functions of other meteorological parameters representing atmospheric stability. Figure 5.2 shows that a_i and b_i are at least functions of wind speed. When using the derived equations for dispersion coefficients for a Gaussian puff model, one must be aware that the expression of σ_z probably underestimates the vertical dispersion in unstable conditions.

The relationship between puff dispersion and dispersion time has been discussed and expressed above, but the influence of other atmospheric parameters on puff dispersion has yet to be addressed. One approach for determining this influence is to adjust the values of these parameters, and while at the same time, ensuring the appropriate P-G stability category is maintained.

5.3.2 Influence of wind speed on puff dispersion

The influence of wind speed on puff dispersion can also be seen in Figure 5.2, where σ_x and σ_y increase with wind speed under stable and unstable conditions. In unstable and (near) neutral conditions, σ_z decreases with wind speed, but increases in table conditions. It is reasonable that, for case 1, when wind speed increases, the stability of the atmosphere decreases, and as a result, vertical dispersion σ_z increases with wind speed. For cases 2 and 3, when speed increases, the stability increases, therefore resulting in decreasing vertical dispersions. Figure 5.3 shows the detailed relationship between wind speed and puff dispersions.

Figure 5.3-1 illustrates the effect of changing the wind speed while retraining the rest of the conditions of case 1. The P-G stability gradually changes from very stable, through stable to slightly stable and neutral conditions. Figure 5.3-2 shows case 2 under different wind conditions, with stability gradually changing from slightly unstable to neutral, while Figure 5.3-3 is a combination of case 3 with varying wind speeds, where stability gradually changes from very unstable, through moderately unstable to slightly unstable. It is found that for neutral and unstable conditions, if simulation results from 0.65 m/s and 1.15 m/s are removed, then reasonable regressions will be obtained. In unstable conditions, as shown in Figure 5.3-3, when wind speed increases from 0.15 m/s to 4.15 m/s, the stability gradually changes from very unstable to slightly unstable, and vertical dispersion decreases as instability decreases, but increases with time. Similarly for case 2, when the wind speed increases, the stability increases from slightly unstable to neutral, which leads to a decrease in the vertical dispersion.

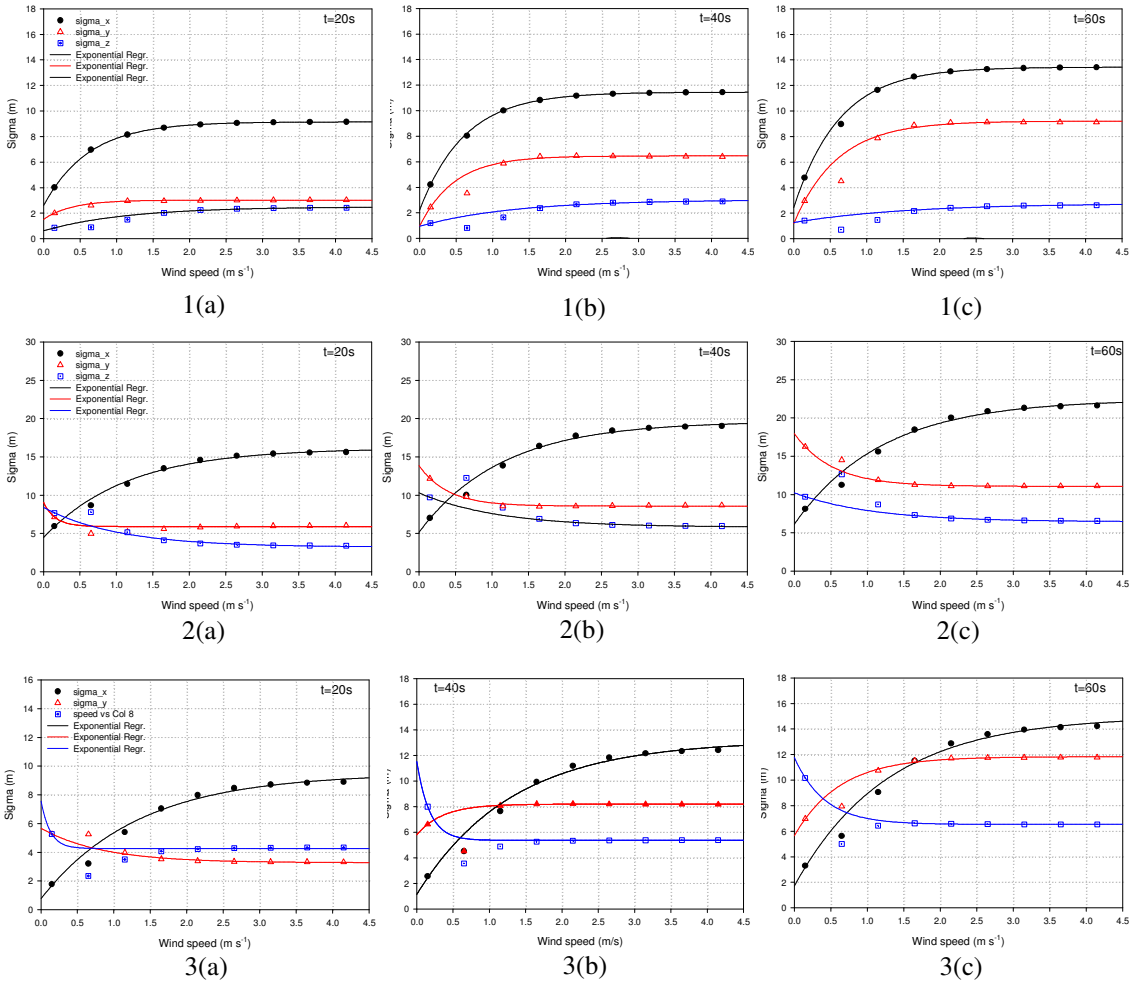


Figure 5.3 Wind speed influence on puff dispersion

Puff dispersion under stable conditions is shown in Figure 5.3-1. Puff dispersion in all directions increases with wind speed, which is reasonable, as when wind speed increases, the stability gradually decreases from moderately stable to slightly stable and finally to near neutral. This progression enhances the vertical dispersion, but since it is still in the range of stable conditions, the increase is minor. In horizontal directions, however, increased wind speed may also provide stronger turbulence and wind shear, and result in a greater increase of σ_x and σ_y. When wind is further increased to near neutral

conditions, the influence of wind speed on puff dispersion is so small that it can be ignored.

An interesting result from Figure 5.3-2 is that longitudinal dispersion, σ_x, increases with wind speed, while lateral dispersion, σ_y, decreases with wind speed under neutral conditions. The longer the dispersion time, the more obvious the trend. Also, the rate of change drops quickly with wind speed. The reason might be that with the wind getting stronger, there is more wind shear influence on σ_x, which helps increase the longitudinal dispersion. When it has reached neutral conditions, cloud size in the crosswind and vertical directions change little with speed.

Figure 5.3 also shows that the great influence of wind speed on puff dispersion only happens during the period of stability transition. When stability stays in the same category, wind influence will be minor and ignorable. Also, predictions in neutral and unstable conditions after 60 s are not very reliable, as this is outside the training space.

To summarize, the simulations shown in Figure 5.3 indicate that, in all conditions, the effect of wind speed on puff dispersions can be well represented by

$$\sigma(U) = \sigma(U)_0 + c_1(1 - e^{(-d_1 * U)}), \text{ for } \sigma_x \text{ in all conditions}$$

σ_y in stable and unstable conditions

and σ_z in stable conditions

or $\sigma(U) = \sigma(U)_0 + c_2 e^{(-d_2 * U)}$, for σ_y in neutral conditions

and σ_z in unstable and neutral conditions

The values of the coefficients σ_0, c_i and d_i, i = 1, 2, differ with stability conditions, and are functions of dispersion time.

5.3.3 Influence of temperature on puff dispersion

The influence of temperature on puff dispersion was studied by changing it through a reasonable range while ensuring the appropriate stability category was identified, then observing the effect on dispersion.

For case 1, the temperature was varied from -18°C to 12°C, every 3°C, while keeping the other parameters as shown in Table 5.6. The temperature range for case 2 was from -20°C to -4°C under neutral conditions and case 3, in unstable conditions, the temperature was changed from 19°C to 26°C to ensure that the stability remained in the same category.

As shown in Figure 5.4, for the neutral and unstable conditions of cases 2 and 3, at any wind speed, the influence of temperature on puff dispersion is so minor that it can be ignored. This also applies for dispersion in stable conditions when the temperature is below zero. However, temperature change starts to show noticeable influence on puff dispersion coefficients in stable conditions when the temperature is above zero, and this influence is dramatic when the temperature is above 6°C, where small temperature changes lead to large drops in σ_x and σ_y. This influence is intensified when wind speed increases. At the same time, σ_z increases slightly with temperature.

The reason for this phenomenon might be that when temperature is above 6°C, the air stability may have changed from stable to slightly stable or even neutral conditions, and increase of wind speed also decreases the intensity of stability. Decreasing stability enhances the vertical dispersion.

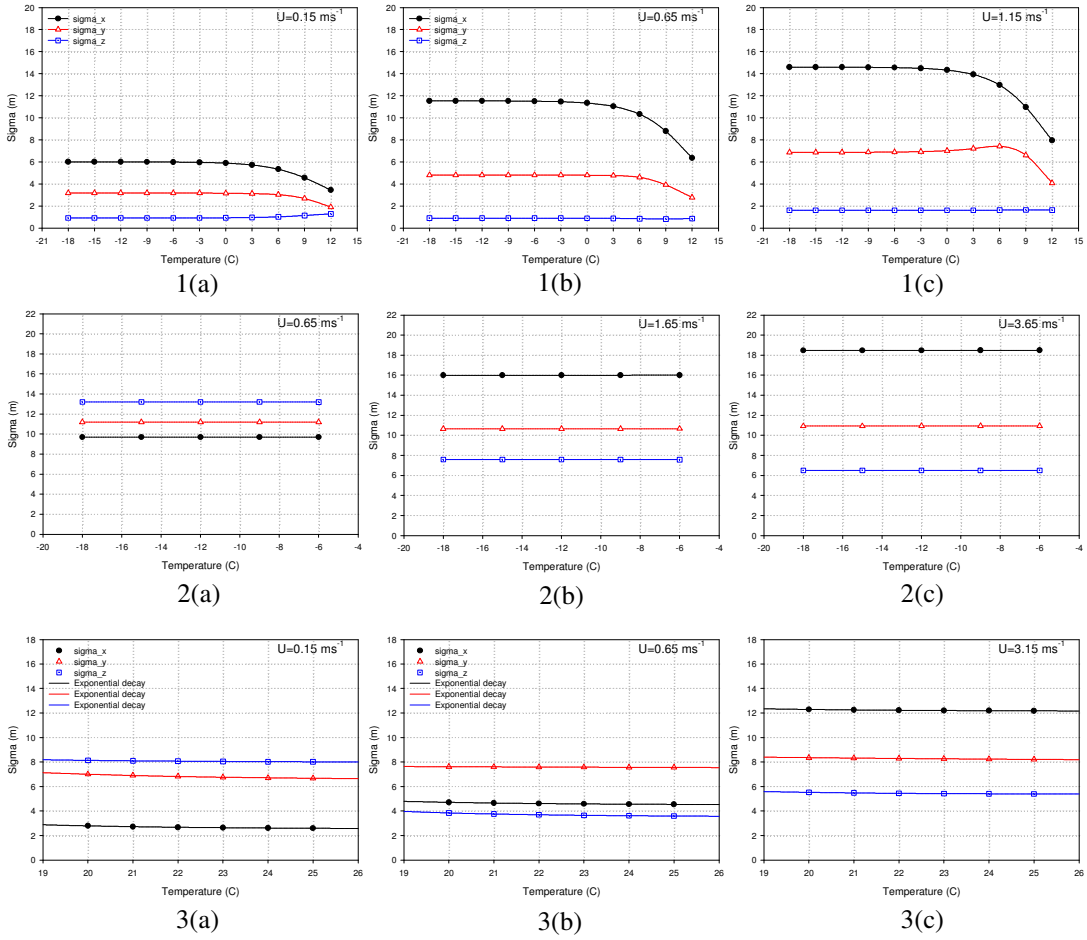


Figure 5.4 Influence of temperature to puff dispersion

5.3.4 Influence of relative humidity (RH)

Relative humidity is usually high at midnight and in the early morning, drops rapidly after the sun rises, until it is lowest just after midday. It then increases rapidly in the late afternoon and early evening and levels off around midnight. As the temperature increases, the relative humidity usually decreases, or vice versa. So when RH changes, the air temperature would also likely change. However, to study the influence of RH on puff dispersion, the other meteorological parameters were fixed, and only the RH was changed. RH changes were kept within a small range, to ensure there was no change of P-G stability category.

Figure 5.5 shows the simulation results under stable (Figure 5.5-1), neutral (Figure 5.5-2) and unstable (Figure 5.5-3) conditions.

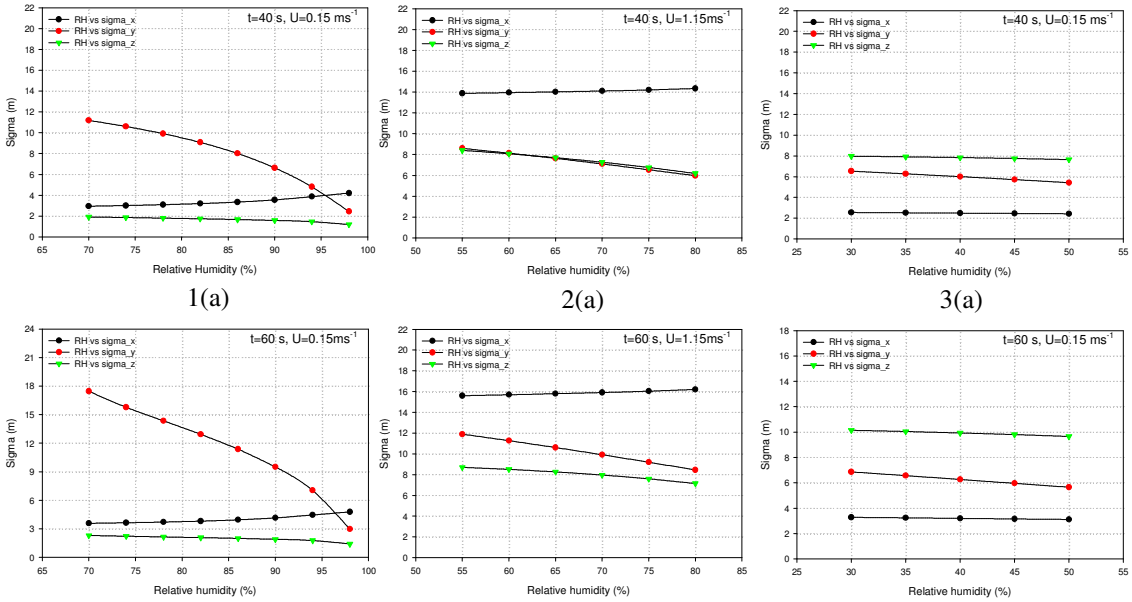
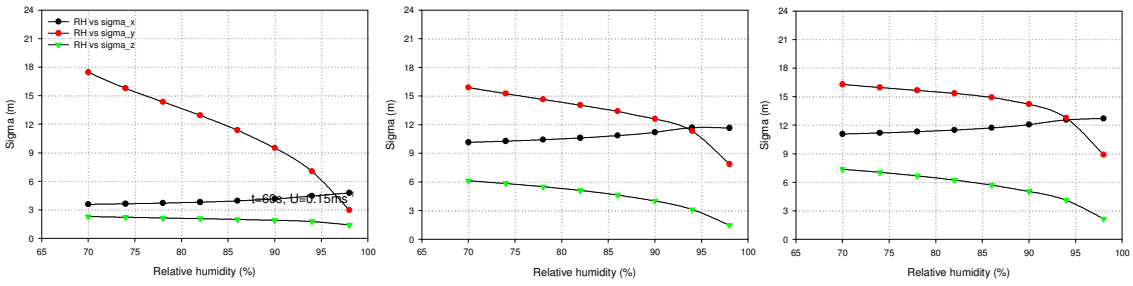


Figure 5.5 Influence of relative humidity on puff dispersion

Dispersion coefficients σ_y and σ_z decrease with relative humidity in all conditions, while σ_x increases slightly with RH in stable and neutral conditions, but decreases in unstable conditions. It is clearly shown in Figure 5.5 that RH shows a greater influence on puff dispersion in the crosswind direction than in the downwind and vertical directions, no matter what the stability category is. The greater the stability, the more influence shown on σ_y, as seen in Figure 5.5. Comparably, the RH influence on σ_x and σ_z was small and could be ignored.

The general trends of the relationship between RH and puff dispersion does not change with time and speed, but it does change in magnitude, as dicussed above about the relationship between puff dispersion and time and wind speed.

In stable conditions, the large change of σ_y with RH mainly appears when the wind is weak. When the wind velocity increases, the rate of change of σ_x and σ_y with RH decreases. However, the influence of RH on σ_z increases as wind velocity increases because increasing wind speed weakens the air stability. When wind speed is strong enough, there was little influence of RH on σ_y, as shown in Figure 5.6, which is simulation results starting from stable conditions at a dispersion time of 60 s.



(a) U=0.15 m/s (b) U=1.15 m/s (c) U=1.65 m/s

Figure 5.6 Influence of wind speed on influence of RH on puff dispersion

Figures 5.5 and 5.6 also suggest that stability increases with relative humidity, and therefore the vertical dispersion coefficient σ_z decreases with RH. When the wind speed increases under stable conditions, the stability decreases, therefore higher vertical dispersion results at higher speeds.

5.3.5 Influence of pressure

When it is sunny, there is usually high pressure, and cloudy weather usually accompanies low pressure, so pressure could be one of the parameters that reflects the air stability.

Figure 5.7 shows the influence of pressure on puff dispersion from stable to unstable conditions. In each condition, the pressure was changed in a small range to ensure the air condition would still stay in the same P-G stability category.

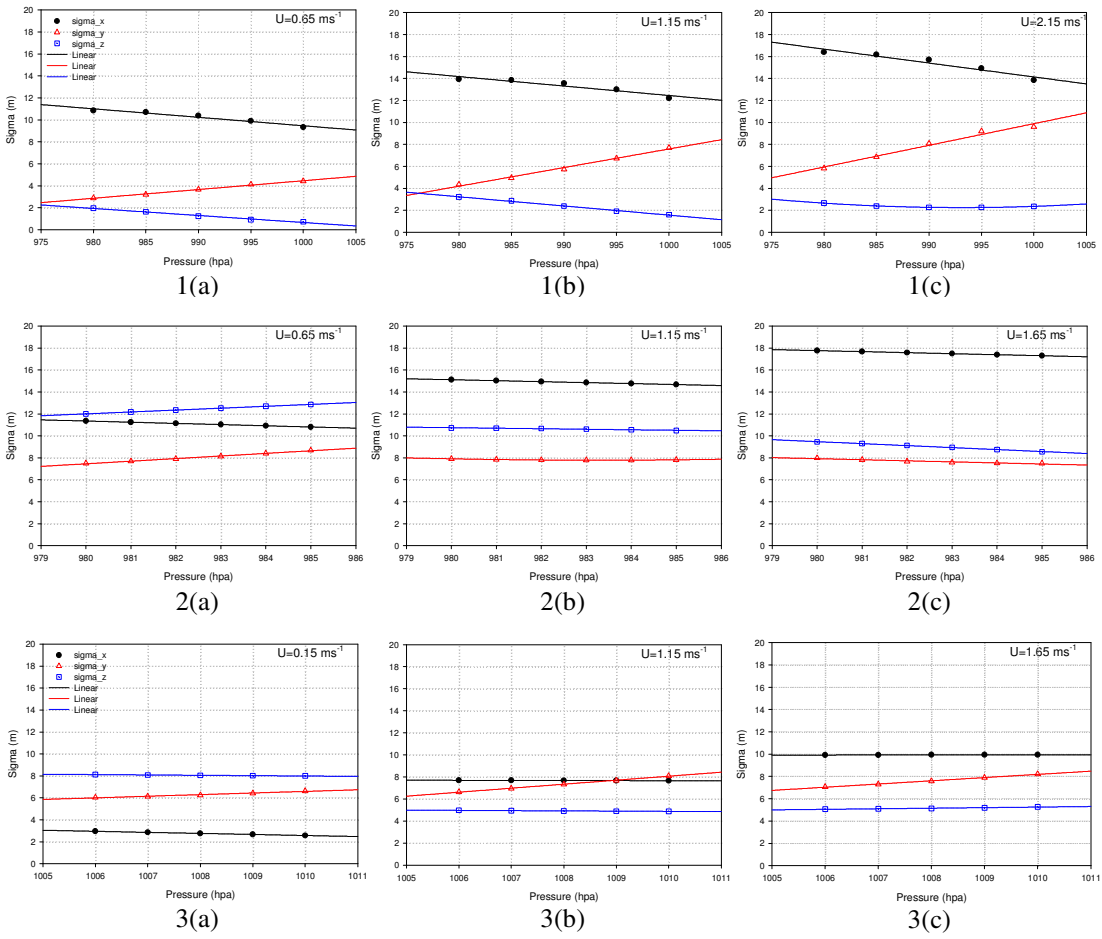


Figure 5.7 Influence of pressure on puff dispersion

Generally speaking, σ_x and σ_z decease and σ_y increases linearly with pressure as long as the pressure change does not cause a change in atmospheric stability. The pressure seems to have a relatively stronger influence on σ_y than on σ_x and σ_z. However, the wind change may affect the relationship between dispersion coefficients and air pressure because wind speed change will lead to a change in stability, as shown in Figure 5.7-2 of σ_y and σ_z.

From Figure 5.7-1 in stable conditions, it can be seen that σ_x and σ_z decrease with air pressure, and the rate of decrease increases with pressure, while for σ_y, it increases with pressure, but the rate of increase decreases with pressure and wind speed. When pressure is low, it might be related to a near neutral condition, and as pressure increases, the stability increases gradually, which explains why σ_z decreases with pressure. However, if wind speed increases are accompanied by increasing pressure, the stability may have changed from near neutral, through stable to neutral, and σ_z decreases initially due to the stability increase. When stability changed to neutral from stable (due to the wind speed increase), σ_z started to increase with pressure, as shown in Figure 5.7-1(c). The change of wind speed affects the magnitude of σ_x and σ_y, the stronger the wind, the more dispersion in longitudinal and lateral directions, which is in agreement with the influence of wind speed to puff dispersion, as discussed previously. The rate of decrease of σ_x and increase of σ_y with pressure increase with wind speed.

It is very interesting to notice that the relationship between σ_z and air pressure changes with wind speed. When the wind is weak (0.65 m/s), σ_z increases linearly approximately with pressure, while σ_z remains almost unchanged when the wind speed is 1.15 m/s. When the wind speed reaches 1.65 m/s, σ_z decreases as air pressure increases, similar for σ_y. From Table 5.6, it is known that case 2 was at 17:30 in March with moderate insolation. When the wind was weak, it might be related to a slightly unstable condition, and an increase of air pressure intensified the instability, thus σ_z increased as air pressure increased. When the wind got stronger, the atmospheric stability

turned to neutral, and air pressure increase may have weakened the instability, and resulted in a decrease of σ_z.

In unstable conditions, small air pressure changes seemed to have little influence on σ_x and σ_z, but σ_y increased with air pressure. Wind speed increase may just have changed stability slightly but still remained in same stability category, so only magnitude of dispersion coefficients changed but not the relationhship of dispersion coefficients and air pressure.

In all, air pressure shows more infuence on σ_y than on σ_x and σ_z as long as the pressure change does not affect the stability. In both neutral and unstable conditions, air pressure was changed in a very small range to ensure the stability category did not change. However, in case 1, pressure was changed from 980 to 1000 hPa, a range that was 4 times wider than for the neutral (2) and unstable (3) conditions. If the change of pressure were limited to the same small range as in unstable conditions, then it can be found that the influence of pressure on puff dispersions in all three directions is similar for all three conditions.

5.3.6 Influence of Insolation

Figure 5.8 shows simulations generated by ANN models by changing the received solar irradiation or insolation only. Similarly, the change of solar irradiation was controlled in a small range, to ensure each case stayed within the same P-G stability.

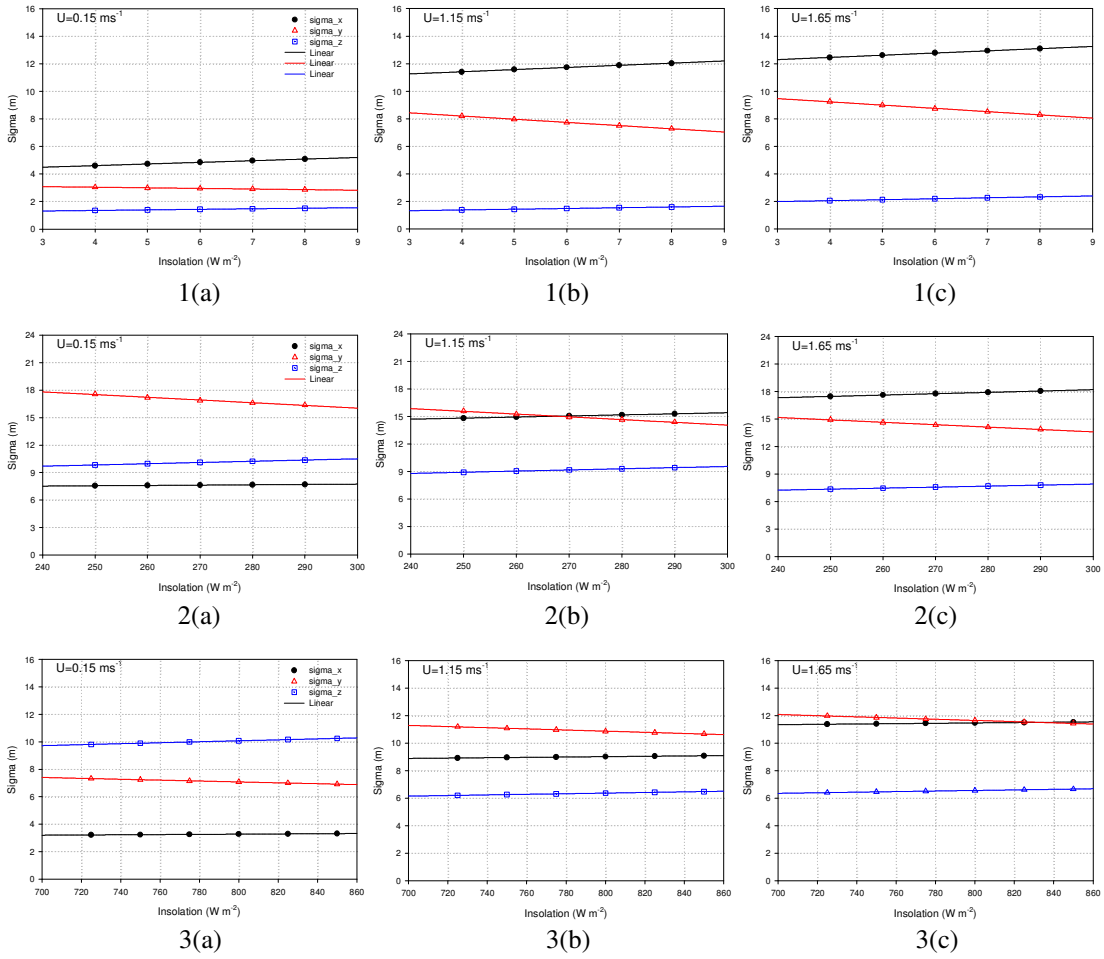


Figure 5.8 Influence of insolation on puff dispersion

Figure 5.8 shows that σ_x and σ_z increase, while σ_y decreases with insolation, in all the different stability conditions. The more insolation received, the higher the temperature on the surface, and the more air was warmed and moved upward. Due to the fact that only a small amount of insolation was changed and stability stayed in the same category, the change of σ_z caused by a solar irradiation increase was small, as shown in Figure 5.8. Under the same temperature and insolation, an increase in wind speed will cause a stability increase, and thus σ_z decreases with wind speed in neutral and unstable conditions, as shown in Figures 5.8-2 and 5.8-3. However, in stable conditions, an

increase of wind speed decreases the stability intensity, but still stays within the stable conditions, so σ_z increases slightly with wind speed, as shown in Figure 5.8-1.

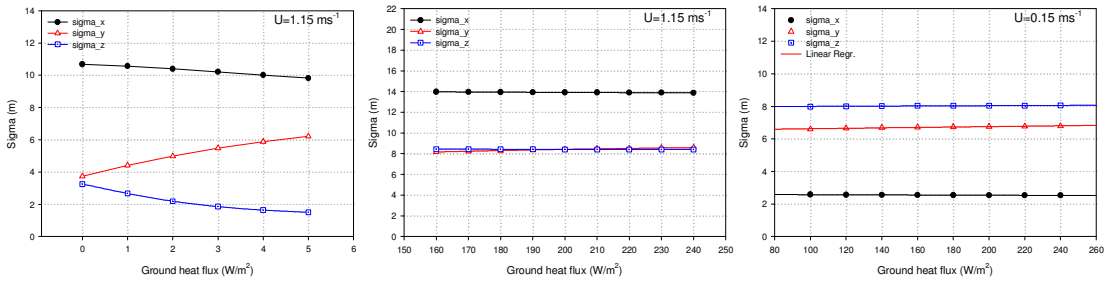
Similar to the air pressure influence on dispersion coefficients, the change of insolation shows a relatively stronger influence on σ_y than the other two dispersion coefficients. This influence increases with increasing stability. Linear regressions give good fits to the simulations of dispersion coefficients for varying amounts of solar irradiation. Wind speed seems to not have much influence on the solar irradiation-dispersion coefficient relationships, except in stable conditions where increasing wind enhances the drop of σ_y with insolation, which may due to the stability change from very stable to slightly stable.

5.3.7 Influence of ground heat flux

In the field experiments, both insolation and ground heat flux (radiation from the ground) close to the ground were measured, and their difference was taken as the net ground heat flux. The greater the ground heat flux, the more the adjacent air is warmed by radiant and convective heat transfer and becomes buoyant. Consequently, vertical dispersion should increase with increasing ground heat flux, and vice verse.

On sunny days with a clear sky, solar irradiation is a lot stronger than ground heat flux (case 3 in Table 5.6). While in neutral conditions in summer time with a cloudy sky (case 2), the received insolation and ground heat flux are close and the intensity of the insolation received should be much less than in sunny, unstable conditions but a lot stronger than at night, where only a little insolation remained (case 1). The ground heat flux was changed while keeping other parameters constant in each case as shown in Table

5.6. The influence of ground radiation on dispersion coefficients is shown in Figure 5.9, where (a) is under stable conditions, and (b) and (c) are under neutral and unstable conditions, respectively.



(a) F (b) D (c) A

Figure 5.9 Ground heat flux influence on puff dispersion

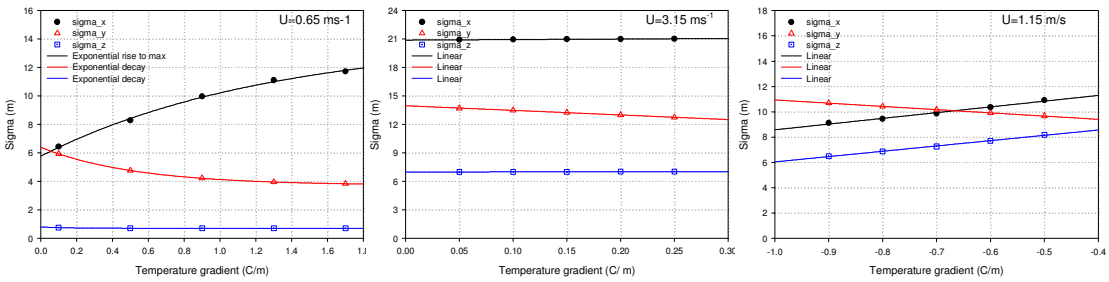
It is found that the change of ground heat flux shows great influence on puff dispersion in stable conditions, while little influence under neutral and unstable conditions. In stable conditions, when the ground heat flux increases, the net heat flux decreases and even drops to be negative, therefore surpressing the vertical dispersion. Consequently, σ_z decreases with increasing ground heat flux. In neutral and unstable conditions, due to strong insolation, the percentage of the change of net heat flux caused by the change of ground heat flux is much smaller than the corresponding change in stable conditions, and thus shows a smaller influence on puff dispersion. The influences of ground heat flux on σ_x and σ_y are opposite, in that σ_x decreases, while σ_y increases with increasing ground heat flux (similar to the influence of insolation).

5.3.8 Influence of temperature gradient (TG)

Unstable conditions are usually related to negative temperature gradients where temperature decreases with height, and vice verse in stable conditions. In neutral

conditions, there is not much temperature difference with height, the gradients should be around zero.

Figure 5.10 shows the simulation results of the influence of the temperature gradient (TG) on puff dispersion under different P-G stability categories, where (a) is in stable conditions, (b) is in neutral conditions and (c) is in moderately unstable conditions.



(a) F (Stable) (b) D (Neutral) (c) B (Moderately unstable)

Figure 5.10 Influence of temperature gradient to puff dispersion

As with the influence of insolation and ground heat flux on puff dispersion, σ_x and σ_y change with the temperature gradient in opposite directions, i.e., σ_x increases, while σ_y decreases as the temperature gradient increases. When in stable and neutral conditions, σ_z decreases slightly when the temperature gradient increases, which is reasonable, in that a positive temperature gradient corresponds to a negative heat flux and surpresses vertical dispersion. Since the initial vertical dispersion is so small, further supression will not cause much change in σ_z. However, in unstable conditions, based on simulations, the influence of TG on vertical dispersion is such that σ_z decreases as the negative temperature gradient increases, as shown in unstable conditions (Figure 5.10c). This may be another example where ANNs did not learn enough information of vertical

dispersion in unstable conditions due to the lack of data in the training set under those conditions. As a result, predictions of vertical dispersion in unstable conditions are questionable.

In summary, a change of TG has more of an influence on horizontal dispersion, i.e., σ_x and σ_y (except in unstable conditions) than on σ_z. The exponential expressions (rise to maximum or decay) are best fits for dispersion in stable conditions, while linear regressions are best for dispersion in neutral and unstable conditions.

5.4 Summary

Puff dispersion can be expressed quite effectively by simple Gaussian puff models if the appropriate dispersion coefficients and cloud centroid can be determined. The predictions of dispersion coefficients from ANNs are very close the observations. Generally speaking, Gaussian puff models with dispersion coefficients from ANNs outperform COMBIC and a Gaussian puff model using Slade's parameterizations. The F_2, F_{10} and NMSE measures of accuracy all support this conclusion, as shown in Table 5.5. Slade's Gaussian puff model is good when dealing with dispersion in near neutral and stable conditions.

The ANN models for dispersion coefficients use continuous meteorological variables which are easily measured as inputs to avoid defining the P-G stability category in advance, as used by most models. As already noted, each P-G stability category covers a wide range of meteorological conditions, therefore introducing relatively large errors when using these puff dispersion models. Further, it may be difficult to properly define the P-G stability by cloudiness, wind speed and time of day. This is the first time a neural

network has been used to predict the dispersion coefficients for Gaussian puff models. The results have provided very good fits to observations.

By using ANNs to generate simulations for dispersion coefficients, it has been found that dispersion coefficients increase with time. However, the relationship between dispersion coefficients and dispersion time is not linear, as stated by statistical theories (Van Ulden, 1992; Mikkelsen, 1987), but rather power law or exponential expressions better describe the relationship between dispersion coefficients and short dispersion times (60 s).

Wind speed was found to have a significant influence on puff dispersion, since the change of wind speed results in a change of stability category. When the wind is strong enough, and the stability stays in the same category, then the wind influence becomes minor and can be ignored.

The influence of temperature on dispersion coefficients shows up only in stable conditions and when the temperature is above the freezing point. In other conditions, the influence of temperature can be ignored.

Relative humidity has a direct relationship with temperature, as a higher temperature usually corresponds to a lower humidity, so its influence on puff dispersion is similar to that of temperature. An obvious influence was found only with σ_y in stable conditions. The more stable the atmospheric condition is, the greater the influence of relative humidity.

Air pressure changes may cause changes of wind speed, and further cause changes in temperature, relative humidity, and finally in atmospheric stability. If for a small change of air pressure, the stability remains in the same category, the influence of

air pressure on puff dispersion is minor. But at night, if pressure significantly drops, e.g., from 1005 hpa to 980 hPa, it may be related to a weather change from a clear sky to a cloudy sky, and therefore stability may change from stable to neutral, as shown in Figure 5.7-1.

The solar irradiation seems to affect σ_y far more than σ_x and σ_z, while the influence of ground heat flux on puff dispersion is much more obvious in stable conditions than in neutral and unstable conditions. It seems the percentage of net ground heat flux change rather than the absolute change of gound heat flux plays the role on the influence on dispersion coefficients. In unstable conditions, since the change of downward insolation caused small percentage change of heat flux, no obvious change was observed by changing downward insolation.

Intuitively, temperature gradient should be a key factor in puff dispersion. Under unstable conditions, the larger the temperature gradient, the more unstable the atmospheric conditions, and the greater the vertical dispersion is supposed to be. However, use of the ANN model in unstable conditions shows the opposite result, that vertical dispersion decreases with increasing temperature gradient. This may be one sign that the developed ANN is not good enough in predicting vertical dispersion in unstable conditions, due to less information in the training data.

In sum, the ANN model generated dispersion coefficients for Gaussian puff models that, in turn, provided insights into the effects of individual meteorological parameters on dispersion.

Chapter 6
Results and Conclusions

6.1 Summary and Conclusions

Four one-week-long periods of field trials were conducted over flat terrain at DRDC -Valcartier in November 2002 and March, May and August 2003. During these periods, three different non-buoyant aerosols (10 g of microspheres made of PVC, 20 g of glass spheres and 20 g of talc powders) were released from near ground level, to examine the influence of a variety of atmospheric and micrometeorological factors on aerosol dispersion in the surface boundary layer over flat terrain. Factors considered included wind velocity, air temperature, relative humidity and pressure, along with their associated gradients, ground heat flux and insolation, as well as diurnal and seasonal variations. The influence of derived turbulence statistics was also examined.

Transient instantaneous three-dimensional concentration maps were measured using a scanning lidar system, which provided over 160,000 useful concentration measurements to train the ANN. Since not enough parameters were available to be able to convert the lidar measurements to absolute concentrations, relative concentrations were determined, where signals after lidar inversion were normalized by the initial peak value of the corresponding cloud.

The field data were analyzed using a relative coordinate system, which followed the centre of mass of the puff as it diffused downwind, and thus removed the effect of puff meander from the diffusion process.

Several multi-layer feed-forward ANN models were constructed and trained by automatically pruning hidden nodes and using the extended delta-bar-delta backpropagation learning rule. The one with the best performance over the training, test, and validation sets was selected and compared with traditional Gaussian puff models using Slade's parameterizations for dispersion coefficients and COMBIC, the US Army Research Laboratory's Gaussian-based dispersion code.

Overall, the ANN model showed better performance than COMBIC and Slade's model. Though the ANN underestimated maximum concentrations in all conditions (as did the other two Gaussian puff models), this is in agreement with some reports (Yi and Prybutok, 1996; DeVito, 2000; Cappa and Anfossi, 2001). The ANN predictions of higher concentrations for all validation trials were in the range of a factor of two, as shown in Figure 3.3. For all 4 validation trials, the ANN was the best, based on the closeness of the predictions to measurements.

Several ANN models of absolute concentration were also developed by using some empirical values for those parameters necessary for the conversion from direct lidar inversion values into absolute concentrations. The average of predictions of those ANNs were compared with COMBIC and Slade's model. This time, ANNs were unequivocally the best among the three models, providing the lowest NMSE, FB and highest R, F_2 and F_{10}, even though they still slightly over-estimated the low concentrations. Also, the ANNs under-estimated the maximum concentrations, probably due to insufficient amount of data with high concentrations available in the training data. The model with Slade's dispersion coefficients showed the worst performance, with more than 90% of its predictions beyond a factor of two of the measured values.

The trained ANN for normalized concentrations was also used to simulate the puff concentration distribution in unstable, neutral and stable conditions. The model predicted very smoothly peaked concentration distributions that were well represented by Gaussian and Lorentzian distributions in the downwind and crosswind directions, though the predicted centre of mass tended toward the upwind direction as dispersion time was increased. Though Lorentzian curves showed better fits to the predictions, long tails predicted by Lorentzian regressions in downwind and crosswind directions did not reflect reality. Comparatively, Gaussian regressions are a better choice though they might underestimate peak concentrations.

A sensitivity analysis performed by dithering the inputs and checking the corresponding change in the output showed that dispersion time and relative positions are the most significant inputs for the ANN. In addition, insolation, temperature, air pressure, TKE and C_N^2 are also big contributors to the output of the ANN. By comparison, the puff initial size seemed to not have much influence on the output.

For the vertical distribution, the ANN predicted endless or indeterminate tails in the vertical locations far from the centre of mass. This may due to the missing zero-concentration points vertically outside the boundaries of the training data. The same problem was found in horizontal dispersion, when no zero-concentration points were included in the training data. The predictions became much more reasonable after some zero concentration points outside the horizontal boundaries were inserted into the training data (DeVito, 2000). Also, concentration predictions would be expected to be close to zero when vertical positions are far from the centre of mass, if zero-concentration points outside the vertical boundaries had been contained in the training data. If the long tails of

the vertical predictions are ignored, then Gaussian regressions gave good fits for vertical dispersions. However, the slow settling of aerosol particles in the puff might be a reason that predictions of points which are below the centre of mass are always higher than predictions at same distance above the centre of mass, especially when dispersion time increases, and causes the non-symmetrical contour shape in the *xz* (downwind-vertical) plane, as shown in Figure 4.9.

The ANN simulations also showed the influence of the wind shear effect on puff dispersion. The simulation of the local centroids of the cloud changed with height, with the higher the cloud slice, the more downwind the corresponding centroid. The local centroid linearly increased with height in stable conditions, while the change of centroid location with height was not obvious in unstable conditions (Figure 4.10).

To improve the performance of a Gaussian puff model, the selection of proper dispersion coefficients is the key factor. ANN models for dispersion coefficients were developed based on dispersion time and meteorological variables which avoided the selection of P-G stability as used in Slade's parameterizations and COMBIC. For each dispersion coefficient, five ANN models were trained. The average of the predictions from the five ANN models was taken as the final prediction of the dispersion coefficient.

A Gaussian puff model using dispersion coefficients predicted by ANNs was compared to a Gaussian puff model using Slade's dispersion coefficients and COMBIC which used two different schemes for dispersion coefficients when dispersion time was shorter and longer than 30 s. The ANN-based Gaussian puff model using dispersion coefficients predicted by ANNs for the validation sets did not perform as well as the ANN model developed directly for concentration distributions, but still outperformed the

other two Gaussian puff models (Slade and COMBIC) using different schemes for dispersion coefficients (Table 5.5).

The relationships of dispersion coefficients (the influence of puff initial size was removed) with dispersion time, wind speed and other meteorological variables were studied. Slade's power law was deemed suitable for relationships between σ_x, σ_y and dispersion time in stable and neutral conditions, while hyperbolic and exponential regressions were better expressions for σ_x and σ_y, respectively, in unstable conditions. The exponential rise to a maximum was a good fit for the vertical dispersion coefficients under all conditions. However, since far less data were used in unstable and neutral conditions than data in stable conditions, the results were felt to be more reliable for stable conditions.

Changing in the wind speed had a large influence on dispersion coefficients, when the wind was weak. However, when the wind speed was over 2.0 m/s, the change to the dispersion coefficients caused by changes of wind speed was minor.

As long as temperature change did not affect the stability much, then the change of dispersion coefficients caused by temperature change was so small that it could be ignored, except in stable conditions where a temperature increase above zero could cause a quick drop of σ_x and σ_y.

Relative humidity mainly showed its influence on the dispersion coefficient σ_y, where σ_y decreased with increasing RH. The more stable the condition, the more change was caused by a change of RH.

The influence of air pressure on dispersion coefficients was consistent, with pressure changes resulting in almost the same amount of change in dispersion coefficients

in all conditions. However, the influence of air pressure on dispersion coefficients also changed with wind speed, due to the stability change caused by the wind speed change.

A change in solar irradiation mainly caused a change of σ_y in all conditions. Although there was little influence on either σ_x or σ_z, increasing solar irradiance resulted in a decrease in σ_y, with a linear relationship giving a good approximation. However, the influence of solar irradiation on dispersion coefficients depended on the relative rather than the absolute change (Figure 5.8).

The influence of ground heat flux is similar to solar irradiation but in opposite directions, σ_x or σ_z linearly decreased, while σ_y increased with increasing ground heat flux. Similarly, the influence of ground heat flux on dispersion coefficients depended on the relative rather than the absolute change of net heat flux (difference between insolation and ground heat flux) and was mainly observed in stable conditions (Figure 5.9).

As with the influence of insolation and ground heat flux on puff dispersion, σ_x and σ_y changed with the temperature gradient, albeit in opposite directions; σ_x increased, while σ_y decreased as the temperature gradient increased. However, for simulations in unstable conditions, σ_z decreased as the temperature gradient increased (Figure 5.10c). This may be another sign that the ANNs did not learn enough information of vertical dispersion in unstable conditions due to a lack of data in the training set under those conditions. Consequently, predictions of vertical dispersion in unstable conditions are of questionable value.

To conclude, a comprehensive ANN model for puff concentration distribution has been developed that incorporates explicit values of a number of meteorological and

turbulence parameters over a broad range of seasonal conditions. This single model provides better concentration and dispersion predictions than conventional Gaussian puff models.

6.2 Recommendations and future work

A more practical and robust ANN model could be developed if a more detailed description of the character of the flow field was measured. The wind used in the ANN models was the wind measured at heights of about 2.0 m, while it is more conventional to consider wind speeds at 10 m heights. In addition, since the effects of wind shear were determined to be very influential in the process of puff dispersion near the surface, vertical profiles of mean velocity should also be measured.

When ANN models are deployed in practice, the prediction of absolute concentrations is the final concern. Predictions of the ANN model for normalized concentrations can be converted into absolute concentrations if the initial maximum concentration is known. However, this initial maximum value may have to be estimated using other models, and therefore will probably introduce more errors into the final results. To get the absolute concentration measurements, visibility, size distribution of aerosols and lidar ratios of air background and aerosols have to be measured, which requires that some other experiments be done before the puff releases.

Given that an ANN model is empirically based, it is reliable when used for interpolation. When input conditions are out of the range of the ANN training data, extrapolation will be required. Therefore, the predictive capability of any ANN model is restricted to the conditions under which the data were collected. Though the ANN is based on a wide range of conditions, however, due to the scanning volume restrictions,

only a few diffusion data were available in unstable and neutral conditions (due to clouds moving quickly downwind in neutral conditions and vertical expansion in unstable conditions). The result was that clouds quickly moved out the scanning range of the lidar, with usually only one or at most two scans being available. Consequently, not enough information of diffusion with time could be caught with those limited data, and the use of the ANN in unstable and neutral conditions remains somewhat suspect.

The current model is based on data of very short diffusion time for unstable and neutral conditions due to the fixed lidar scanning space. To make the model more useful, longer range dispersion data is needed. Also, only small amounts of aerosols were released and the artificially generated puff cloud quickly mixed with the background air, and thus slipped below the detection limit of the lidar system. Increasing the aerosol amount will lengthen the time period of the puff staying in the air before being mixed, thus more diffusion data could be measured (although this might also increase the optical density, and a more effective lidar inversion technique will be required).

In the development of the ANN model, the first useful scan was taken as the source of a puff, therefore strictly speaking, the developed ANN model is not for puffs from point sources, but rather from a volumetric source. And in neutral and unstable conditions, the size of puff source may be significantly larger than that in stable conditions. To be able to study the dispersion from point sources, a better dissemination technique should be employed, or at least, make sure all puff diffusion starts from same source (with the same small dimensions).

Aside from considering improvements in measuring techniques and more variables, including clear-air zero-concentration data points outside the vertical boundary

in the training data set without greatly skewing the output frequency distribution should be considered. Including clear-air zero-concentration points outside the horizontal boundaries of a puff has been proved to be very helpful in predicting the concentration of points away from the central portion of the puff in the downwind and crosswind directions. It should be helpful in the vertical predictions too.

The future work should include developing the appropriate functions to describe the relationship of dispersion coefficients to meteorological inputs, to make it available for practical use. In addition, as source information is very important to the accuracy of model predictions and the developed ANN model for puff concentration distribution is for a short time application (under 100 s), it would be very useful to try to link the developed model to other more complex models for longer dispersion times.

References

Andrews, W.S., S.D. Roney and G. Roy (1996) An examination of modelling approaches for predicting the dispersion of aerosols from obscurant grenades. Paper presented at the 1996 Battlespace Atmospherics Conference, Naval Command Control and Ocean Surveillance Center, RDT&E Division, San Diego, CA, 3-5 December 1996

Andrews, W.S., J.R. Costa and G. Roy (1997) Measuring and modelling the influence of atmospheric effects on the concentration distributions within transient aerosol plumes. Paper presented at the 1997 Battlespace Atmospherics Conference, Space and Naval Warfare Systems Center, San Diego, CA, 2-4 December 1997

Andrews, W.S., J.R. Costa and G. Roy (1998) Modelling concentration distributions within transient aerosol plumes. Paper presented at the 1998 Battlespace Atmospheric and Cloud Impacts on Military Operations Conference, Air Force Research Laboratory, Hanscom Air Force Base, MA, 1-3 December 1998

Andrews, W.S., B.J. Lewis and D.S. Cox (1999) Artificial neural network models for volatile fission product release during severe accident conditions. Journal of Nuclear Materials 270:74-86

Andrews, W.S., X. Cao, G. Roy, L. Forand and G. Potvin (2003) Field data collection in support of dispersion modelling, Proceedings of Battlespace Atmospheric and Cloud Impacts on Military Operations Conference 2003, Monterey CA, NRL, September 2003.

Andrews, W.S., X. Cao, G. Roy, L. Forand and G. Potvin (2005) Modelling the dispersion of aerosols from point sources. Proceedings of Battlespace Atmospheric and Cloud Impacts on Military Operations Conference 2005, Monterey CA, NRL, October, 2005

Arya, S.P. (1999) Air pollution meteorology and dispersion. Oxford University Press

Ayres, S.D. and S. DeSutter (1995) Combined Obscuration Model for Battlefield Induced Contaminants (COMBIC92) Model documentation. U.S. Army Research Laboratory, Battlefield Environment Directorate, White Sands Missile Range, NM 88002-5501, 1995

Batchelor, G.K. (1952) Diffusion in a field of homogeneous turbulence II. The relative motion of particles. Proc. Camb. Philos. Soc., 48: 345-362

Baughman, D.R. and Y.A. Liu (1995) Artificial neural networks: Neural networks in Bioprocessing and Chemical Engineering. Academic Press

Bennett, M., S. Sutton and D.R.C. Gardiner (1992) An analysis of lidar measurements of buoyant plume rise and dispersion at five power stations. Atmos. Environ. 26A(18): 3249-3263

Bennett, M. (1995) A lidar study of the limits to buoyant plume rise in a well-mixed boundary layer. Atmos. Environ. 29(17): 2275-2288

Bissonnette, L.R. (1996) Lidar inversion methods: an introduction. Proc. 8th Int. Workshop on Multiple scattering lidar experiments. 102

Bisonnette, L.R., G. Roy, S. Cober and G.A. Isaac (2002) Lidar monitoring of clouds: Remote sensing of clouds and the atmosphere. Proc. SPIE 4539: 8-17

Bissonnette, L.R., G. Roy, G. Vallée and S. Cantin (2003) Measures lidar d'un panache d'insecticide èmanant d'une source continue. RDDC Valcartier ECR 2003–273

Cappa, C. and D. Anfossi (2001) Short-term prediction of urban NO2 pollution by means of articial neural networks. Int. J. Environment and Pollution, 15(5): 483-496

Chatwin, P.C. (1968) The dispersion of a puff of passive contaminant in the constant stress region. Q. J. R. Meteo. Soc. 94: 350-360

Chatwin, P.C. (1990) Statistical methods for assessing hazards due to dispersing gases. Environmetrics 1: 143–162.

de Hann, P. and M.W. Rotach (1998a) A novel approach to atmospheric dispersion modeling: the Puff-Particle Model. Quart. J. Roy. Meteo. Soc. 124: 2771-2792

de Hann, P. and M.W. Rotach (1998b) The treatment of relative dispersion within a combined puff-particle model (PPM). In: Air Pollution Modelling and its Application XII, S.-E.Gryning and N. Chaumerliac(eds), Plenum Press, New York, 389-396

de Haan, P. (1999) On the use of density kernels for concentration estimates within particle and puff dispersion models. Atmos. Environ. 33: 2007–2021.

DeVito T.J., W.S. Andrews and G. Roy (2000) Modeling aerosol puff concentration distributions from point sources using artificial neural networks. Paper presented at the 2000 Battlespace Atmospheric and Cloud Impacts on Military Operations Conference, Cooperative Institute for Research in the Atmosphere and Army Research Laboratory, Fort Collins, CO, 24-27 April 2000

DeVito T.J., (2000) Modelling aerosol puff concentration distributions from point sources using srtificial neural networks. Master thesis at Royal Millitary College, Canada

Draxler, R.R. (1979) Some observations of the along-wind dispersion parameter. Preprint Volume. Fourth Symposium on Turbulence, Diffusion and Air pollution. Am. Met. Soc., 1979, 5-8

Dwinnell W. (1998) Modeling methodology 2: model input selection. PCAI, January/February 1998, 12(1): 23-26.

Ermak, D.L. and J.S. Nasstrom (2000) A Lagrangian stochastic diffusion method for inhomogeneous turbulence. Atmos. Environ. 34: 1059–1068.

Evans, B. T. N. (1984) On the inversion of the lidar equation. Report DREV R-4343/84, Valcartier, 1984

Evans, B. T. N. (1988) Lidar signal interpretation and processing with consideration for military obscurants. Report DREV R-4477/88, Valcartier, 1988

Forand, J.L. (1999) The L(W)WKD marine boundary layer model. Ver. 7.09, DREV-TR-999-099, Valcartier

Forand, J.L. (2003) Data analysis from fast sonic anemometer and hygrometer, meteorological & micrometeorological. DRDC Valcartier TR 2003-048, Valcartier, June

Frenkiel, F.N. and I. Katz (1956) Studies of small-scale turbulent diffusion in the atmosphere. J. Meteo. 13: 388-394

Gardner M.W. and S.R. Dorling (1999) Neural network modelling and prediction of hourly NOx and NO2 concentrations in urban air in London. Atmos Environ 33(5):709–719

Guardani R. and C.A.O. Nascimento (2004) Neural network-based study for prediction ground-level ozone concentration in large urban areas, applied to the Sao Paulo metropolitan area. Int. J. Environment and Pollution. 22(4): 441-459

Gifford, F.A. Jr. (1955) A simultantaneous Lagrangian-Eulerian turbulence experiment. Mon. Weath. Rev., 83: 293-301

Gifford, F.A. Jr. (1957) Relative atmospheric diffusion of smoke puffs. J. Meteolo. 14: 410-414

Gifford, F.A. (1982) Horizontal diffusion in the atmosphere: a Lagrangian–dynamical theory. Atmos. Environ. 16: 505-512 or JOSA A, 18, Issue 2: 392-398

Hanna, S.R., G.A. Briggs and P.H. Rayford, Jr.(1982) Handbook on atmospheric diffusion. National Technical Information Center U.S. Dept. of Energy, Springfield

Hurley, P. and W. Physick (1993) A skewed homogeneous Lagrangian particle model for convective conditions. Atmos. Environ. 27A: 619–624.

Hurley, P. (1994) PARTPUFF, a Lagrangian particle-puff approach for plume dispersion modelling applications. J. Appl. Meteorol., 33: 285-294

Jørgensen, H.E., T. Mikkelsen, J. Streicher, H. Herrmann, C. Werner and E. Lyck (1997) Lidar calibration experiments. Appl. Phys. B 64: 355-361

Klett, J.D. (1981) Stable analytical inversion solution for processing lidar returns. Applied Optics 20(2): 211-220

Klett, J.D. (1985) Lidar inversion with variable bckscatter/extinction ratios. Applied Optics 24(11): 1638-1643

Kovalev, V.A. (2003) Near-end solution for lidar signals that includes a multiple-scattering component. Applied Optics 42(36): 7215-7224

Luhar, A.K. and S.A. Young (2002) Dispersion moments of fumigating plumes-lidar estimates and pdf model simulations. Boundary-Layer Meteorol. 104: 411-444

Mangia, C., U. Rizza, U. Giostra, M. Cassiani and G.A. Degrazia (2001) A Lagrangian puff dispersion model: evaluation against tracer data. Int. J. Environ. Pollu. 16(1/2/3/4/5/6): 28-35

Measures, R.M. (1984) Laser remote sensing: fundamentals and applications. John Wiley & Sons

Middleton, D.R. (2005) Evaluation of dispersion model parameters by dual Doppler lidars over West London, England. Int. J. Environment and Pollution 25(1-4): 80-94

Mikkelsen T., S.E. Larsen and H.L. Pecseli (1987) Diffusion of Gaussian puffs. Q. J. R. Meteorol. Soc., 113: 81-105

Mikkelsen, T., H.E. Jørgensen, S. Thykier-Nielsen, S.W. Lund and J.M. Santabarbara (1995) Final data and analysis report on: high-resolution in plume concentration fluctuations measurements using lidar remote sensing technique. Technical report, No. Risø-R-852(EN), Risø National Laboratory, Roskilde, Denmark, 1995

Minai, A.A., and R.D. Williams (1990) Acceleration of back-propagation through learning rate and momentum adaptation. International Joint Conference on Neural Networks, 1, January

Mohan, M. and T.A. Siddiqui (1997) An evaluation of dispersion coefficients for use in air quality models. Boundary-Layer Meteorology. 84: 177-206

Mungiole, M. and A. Wetmore (2001) COMBIC Modifications to determine aerosol cloud densities for multiple obscurant input sources. US Army Research Labboratory Adelphi MD, Computational and Information Sciences Directorate

Munro, R.J., P.C. Chatwin and N. Mole (2003) A concentration pdf for the relative dispersion of a contaminant plume in the atmosphere. Boundary-Layer Meteorol. 107: 253-271

Nasstrom, J.S., G. Sugiyama, J.M. Leone, Jr. and D.L. Ermak (2000) A real-time atmospheric dispersion modelling system. Paper submitted to 11th Joint Conference on the Applications of Air Pollution Meteorology with the Air and Waste Management Association Long Beach, CA, January 9-14, 2000

NEURALWARE (2003) Reference Guide: NeuralWorks Professional II/Plus and NeuralWorks Explorer, Pittsburgh, PA

Nickola, P.W. (1971) Measurements of the movement, concentration and dimensions of clouds resulting from instantaneous point sources. J. Appl. Meteo. 10: 962-973

Pasquill, F. and F.B. Smith (1983) Atmospheric diffusion. 3rd ed. New York, Halsted Press

Patterson, D.W. (1996) Artificial neural networks: theory and applications. Simon & Schuste

Petersen, W.B. and L.G. Lavdas (1986) INPUFF 2.0 A multiple source Gaussian puff dispersion algorithm — user's guide. EPA/600/8-86/024. August 1986.

Rao, K.S., R.M. Eckman and R.P. Hosker Jr. (1989) Simulation of tracer concentration data in the Brush Creek drainage flow using an integrated puff model. Appl. Met. 28: 609-616

Rizza, U., M. Cassiani, U. Giostra and C. Mangia (2000) An advanced puff model based on a mixed Eulerian/Lagrangian approach for turbulent disperison in the convective boundary layer. Boundary-Layer Meteorol., 95: 319-339

Roy, G., G. Vallée and M. Jean (1993) Lidar-inversion technique based on total integrated backscatter calibrated curves. Applied Optics, 32(33): 6754-6763

Saffman, P.G. (1962) The effect of wind shear on horizontal spread from an instantaneous ground source. Q. J. R. Meteo. Soc., 113(81): 382-393

Sato, J. (1995) An analytical study on longitudinal diffusion in the atmospheric boundary layer. The Geophysical Magazine Series 2, 1(2): 105-151

Slade, D.H. (1968) Meteorological and atomic energy: Diffusion from instantaneous sources. Atomic energy commission, division of technical information. TID-24190, USAEC, 163-175

Smith, F.B. and J.S. Hay (1961) Expansion of clusters of particles in the atmosphere. Quart. J. Roy. Meteorol. Soc., 87:82-101

Smith, F. B. (1965) The role of wind shear in horizontal diffusion of ambient particles. Q. Jl. Roy. Met. Soc. 91: 318-329

Smith, F.B. (1968) Conditional particle motions in a homogeneous turbulent field. Atmos. Environ., 2: 491-508

Sykes, R. I., D. S. Henn, S. F. Parker and R. S. Gabruk (1996). SCIPUFF - A generalized hazard dispersion model. Ninth Joint Conference on the Applications of Air Pollution Meteorology with A&WMA, American Met. Soc.

Turner D.B. (1994) Workbook of atmospheric dispersion estimates – An introduction to dispersion modelling. 2nd Ed., CRC Press Inc

Van Ulden, A.P. (1992) A surface-layer similarity model for the dispersion of a skewed passive puff near the ground. Atmos. Environ. 26A(4): 681-692

Walcek, C.J. (2002) Effects of wind shear on pollution dispersion. Atmos. Environ. 36: 511-517

Wasserman, P.D. (1993) Advanced methods in neural computing. New York, NY: Van nostrand Reinhold

Wei, He, R. Koga, K. Iokibe, O. Wada and Y. Toyota (2001) Stable inversion method for a polarized-lidar: analysis and simulation. J. of the Optical Society of America A 18(2): 392-398

Wilson, D.J. (1981) Along-wind diffusion of source transients. Atmos. Environ. 15(4): 489-495

Wilson, J.D. and B.L. Sawford (1996) Review of Lagrangian stochastic models for trajectories in the turbulent atmosphere. Boundary-Layer Meteorol. 78: 191-210

Yee, E. (1998) Turbulent diffusion of instantaneous clouds in the atmospheric surface layer: measurements and comparison with Lagrangian similarity theory. Physics Letters A242: 51-62

Yi, J. and R. Prybutok (1996) A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialised urban area, Environmental Pollution. 92(30), 349-357

Yu, L., S. Wang, and K.K. Lai (2006) An integrated data preparation scheme for neural data analysis. IEEE Trans. Knowledge and Data Engineering, 18 (2):217-230

Zuev, V. E. and I. E. Naats (1983) Inverse problems of lidar sensing of the atmosphere. Spring-Verlag, Berlin

Appendix A

The developed ANN model for relative concentration

The ANN model for relative concentration is trained with feedfoward network structure 21-62-1. i.e., one input layer with 21 inputs, one hidden layer with 62 hidden nodes and one output node. Architecture of a simple 3-layer BP feedforward ANN network with one output node is shown in Figure A1, where X_i is input node i, Y_j is hidden node j and Z_k is output node k, and x_i, y_j and z_k are their corresponding outputs.

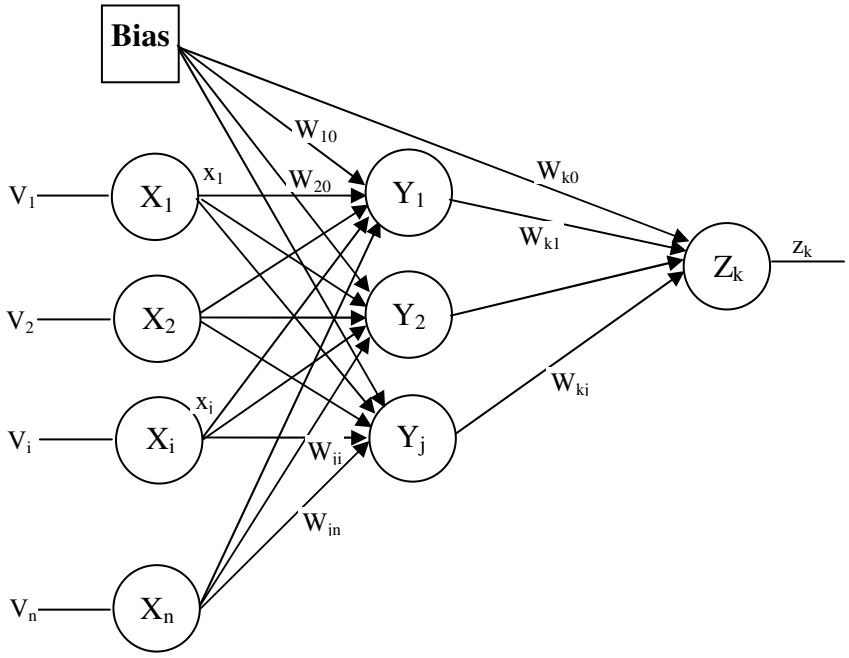


Figure A1 Architecture of a 3-layer ANN

The bias node in Figure A1 is with a fixed value of 1, and is connected to each hidden and output nodes. The bias node serves to offset the transfer function and tends to cause the network to converge more quickly.

Each of the input vectors is mapped or scaled, variable by variable, over the range of each particular variable, as shown below

$$x_i = \frac{(x_{i\max} - x_{i\min})v_i + (v_{i\max}x_{i\min} - v_{i\min}x_{i\max})}{v_{i\max} - v_{i\min}}$$

where x_i is the scaled ith input value, $v_{i\max}$ and $v_{i\min}$ are the maximum and minimum values, respectively, of the range of the un-scaled ith input variable, v_i, and $x_{i\max}$ and $x_{i\min}$ are maximum and minimum values, respectively, of the range of the scaled ith input variable, x_i. (Non)linear function is applied to each of the scaled input variable and make them more uniformly distributed. Table A1 lists the corresponding maximum and minimum and (non)linear function used for each input variable.

Table A1 corresponding maximum and minimum and transformation functions for input and output

	Input	v_{min}	v_{max}	x_{min}	x_{max}	f(x)
1.	Rx_0	4.32474	42.3715	-0.7	5	tanh(x)
2.	Ry0	4.91793	29.565	0.95	7	1.0/sqrt(x)
3.	Rz0	1.4979	18.7354	1.273	18	log10(x)
4.	*t	-130.23	0	0.3162	1.407	(-x)^2.0
5.	U	0.14839	3.70996	-0.4	2.75	tanh(x)
6.	α	-1.1632	0.36293	0.65	0.85	log(x/(1-x))
7.	shape	0	1	0	1	linear
8.	size	6.06	27	1.934	2.391	(x)^4.0
9.	RelX	-66.371	96/7128	-5.685	8.25	tanh(x)
10.	RelY	-89.275	79.2288	-7.35	6.65	tanh(x)
11.	RelZ	-32.045	42.5536	-5.05	6.71	tanh(x)
12.	Inso	0	856.2	0.01832	10.37	log10(x)
13.	d_inso	-2.8	742.9	0.02	20.09	log10(x)
14.	T	-20.24	28.11	-5	2.344	tanh(x)
15.	*dT/dz	-1.7047	0.98	1.05	3	(-x)^4.0
16.	*RH	-97.95	-27.5	0.006	1.275	(-x)^0.25
17.	dRH/dz	-3.9286	4.38	-2.35	2.2	tanh(x)
18.	P	980.7	1010.54	0.005	0.9441	log(x/(1-x))
19.	MOL	-6.0791	2.6751	-9	4	tanh(x)
20.	TKE	4.51	446	0.02	1.5	x^0.25
21.	C_N^2	0.0103	183	0.01832	8	log10(x)
Output **		-6	0	0.5623	3	x^4.0

* v_{min} and v_{max} are minimum and maximum values of (-x).
** v_{min} and v_{max} are minimum and maximum values of log10(RelConc)

After linear and nonlinear mapping, all inputs are linearly mapped into [-1,1] or [0,1]. and Table A2 is the mapping coefficients for each input variable.

Table A2 Coefficients of internal mapping to [-1,1] of inputs

Input (x)	Coefficient of transformation x*=ax+b	
	a	**b**
Bias node	1	0
1	1.246668	-0.2465541
2	3.086344	-2.166527
3	1.738462	-1.18224
4	1.064018	-1.106383
5	1.457929	-0.4460614
6	1.79282	-2.109826
7	1	0
8	0.1069954	-2.496897
9	1.000012	-1.150383e-005
10	1.000002	9.834786e-007
11	1.000042	-3.951956e-005
12	0.7265198	0.2620157
13	0.6662336	0.1319109
14	1.009251	0.009159453
15	0.02506753	-1.03047
16	2.550031	-1.709715
17	1.021599	0.003182278
18	0.2463063	0.3037733
19	1.000336	0.0003356272
20	2.737401	-2.029427
21	0.7575299	0.3158824
Output	50.56252	40.54999

The normalized i^{th} input value, x_i, connected to the j^{th} hidden node, has a weight, $w_{ji.,}$ applied to it. As a result, the hidden node receives the value $w_{ji}x_i$ as an input from the i^{th} node. With the exception of the input layer, nodes sum the weighted inputs. Thus, for the j^{th} hidden node, the summed inputs can be represented by I_j, as shown in Figure A2, and hyperbolic tangent function is applied as the transfer function for all hidden and output nodes.

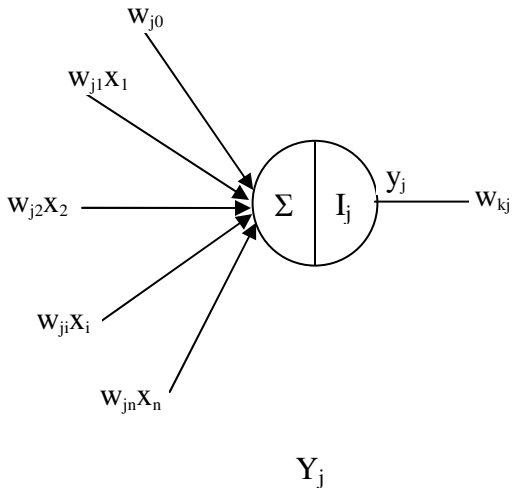


Figure A2 Architecture of the j^{th} node in the hidden layer

So $$I_j = \sum_{i=0}^{21} w_{ji} x_i ,$$

and $$y_j = \frac{e^{I_j} - e^{-I_j}}{e^{I_j} + e^{-I_j}} = \tanh(I_j) .$$

The connections or weights of the ANN model is shown in Table A3.

Table A3 Weights or connections between inputs, hidden nodes and output

Input	Hidden nodes					
	1	**2**	**3**	**4**	**5**	**6**
Bias node	-0.2157502	-0.2178508	0.03958238	-0.2644971	0.4143459	0.6259806
1	0.1943717	-0.02716723	0.0756191	0.2330781	-0.4904777	-0.08512886
2	-0.1100507	-0.3927591	-0.1353807	-0.2065222	0.2842189	-0.1267999
3	-0.2005071	-0.2112705	-0.2183321	-0.2412927	-0.3768625	-0.08292078
4	-0.4123725	-0.01128652	0.1624323	0.2080373	0.1177	0.7707281
5	-0.3017018	-0.05654617	0.1647525	0.0465951	0.4229668	0.3031128
6	-0.5498158	0.3713478	-0.3797355	-0.2500623	-0.1056138	-0.3542728
7	0.3341917	0.005254213	-0.289725	0.2028558	0.08059305	0.5526771
8	-0.1376869	-0.04373707	0.1047858	-0.05165071	-0.1360485	-0.05030367
9	0.6176309	-0.1642844	-0.3131033	0.08351656	-0.08649668	-0.1068206
10	0.2393937	0.3597145	-0.3826337	0.07922127	-0.3289927	0.008599514
11	-0.02105703	-0.06988984	0.10785	0.07147758	-.0001545266	-0.1322541
12	-0.08036762	-0.01454886	-0.1750097	-0.05384137	0.1217475	0.3040862
13	-0.2456671	-0.02499704	-0.1083657	0.3788695	0.002031285	0.1639292
14	-0.6245255	0.007713555	0.4115998	0.5057113	0.3963481	-0.07299408
15	-0.3676081	0.001480732	-0.1937222	0.09038228	0.1182805	-0.126945
16	0.3234687	0.1826617	0.1272519	0.4831422	-0.3418439	0.2985591
17	0.1729272	0.1191787	-0.1404401	0.01010134	-0.3486672	0.3104022
18	0.5103945	0.1477879	-0.23705	0.05588979	-0.2282034	-0.3300927
19	-0.09758022	-0.247706	0.08219323	-0.4223496	-0.1160952	0.0759436
20	0.2009542	0.1347538	-0.1805495	0.09121042	0.3149568	0.3124126
21	0.6468438	0.1140091	-0.0989505	0.4605693	-0.2115101	0.1809292
Output	-0.08069207	0.02713241	0.03863226	0.2335511	-0.07277268	-0.2520019

Table A3 Weights or connections between inputs, hidden nodes and output **Cont.-1**

	Hidden nodes					
Input	**7**	**8**	**9**	**10**	**11**	**12**
Bias node	-0.02907554	0.07112975	0.3280505	-0.1542062	-0.4893863	-0.489559
1	-0.3669612	0.1573894	-0.151002	0.3302394	-0.420909	0.04587847
2	-0.4063432	-0.04612966	-0.1762318	-0.1541951	0.1727313	-0.2383998
3	0.3117501	0.03184282	0.0667374	-0.2030166	0.0317875	-0.2410618
4	0.1724782	0.2082342	-0.4700755	0.1137135	0.1560288	-0.4294662
5	-0.2710556	0.100251	-0.1651353	0.2762873	-0.09052295	0.1770451
6	0.2638234	-0.3437465	0.05222166	-0.08454378	-0.4048814	-0.2607721
7	0.2118235	0.09798755	0.1808764	-0.1808116	-0.3453395	-0.3454785
8	0.2783717	-0.2637194	0.1172293	-0.4986396	-0.7917814	0.4562229
9	0.1827933	0.6983398	-0.08643807	-0.06788371	-0.6395409	0.1018323
10	0.1706335	0.007280208	0.1612034	-0.1914193	-0.8763486	0.01044369
11	-0.3162737	0.2705333	0.1452478	-0.2436531	0.2931657	-0.2799253
12	-0.0612052	0.2274131	-0.2187688	0.2178409	-0.2792906	-0.4824236
13	0.1255783	0.08707684	0.09042569	-0.174633	0.322454	-0.3776844
14	0.4391898	-0.09396896	-0.3149302	0.1632355	-0.4764182	-0.1093156
15	0.01423171	-0.06458029	0.01847658	0.09902739	-0.4439897	0.0636371
16	0.2878263	-0.02894794	-0.1949711	0.0164578	0.3240896	0.07273295
17	0.09334919	0.2611827	0.1157466	0.08973476	0.07579752	0.212619
18	0.1317369	-0.3890211	0.2711927	-0.3219469	0.8929539	0.3314594
19	-0.5506486	0.3623723	-0.02015743	-0.3401417	0.08476397	0.03492257
20	-0.2364507	-0.5755886	0.3009897	-0.217878	0.3398294	-0.1302211
21	-0.1102529	0.1590496	0.4137577	-0.436239	0.0658158	0.3581955
Output	-0.03851756	0.07272935	-0.07514427	0.1422737	0.2594632	0.2395521

Table A3 Weights or connections between inputs, hidden nodes and output **Cont.-2**

	Hidden nodes					
Input	**13**	**14**	**15**	**16**	**17**	**18**
Bias node	0.08840536	-0.01674143	0.2555992	-0.566063	0.1119347	-0.3335799
1	0.1209987	-0.07730278	0.06559597	-0.1196674	-0.07325107	0.320608
2	0.02414667	-0.2368461	0.1425041	0.2205414	-0.02111892	-0.5167263
3	-0.2468926	-0.4893889	-0.16695	-0.4236897	-0.1930615	0.267066
4	-0.03564541	0.1184961	-0.5177307	0.2717851	0.7608885	-0.109436
5	-0.3370334	-0.3719267	-0.2659364	-0.1823258	0.0845364	-0.2822405
6	-0.1310828	-0.08563955	-0.1672816	-0.07285403	-0.08013319	-0.3280114
7	-0.2254697	0.1486676	-0.1041616	-0.03291336	0.1614045	-0.0253848
8	0.05464769	0.08333179	0.4094965	0.03305428	-0.03366699	0.03290086
9	0.1461473	0.1874842	0.6947723	0.4353334	-1.791632	0.09546155
10	-0.2158426	-0.1053581	0.02509179	-0.602362	0.9298823	0.3071053
11	0.05502864	-0.03424111	0.3338603	0.3463861	-0.5621065	0.09075208
12	0.2275486	0.2488642	0.07176482	-0.232965	-0.5321333	-0.01371531
13	0.1774577	-0.09305287	0.1811955	0.3242536	0.2177006	-0.2415204
14	-0.3467484	0.01110101	0.2634607	-0.2715881	-0.3997873	0.1640448
15	-0.08849021	0.02650147	0.2867638	-0.1547718	0.2470999	0.04753669
16	-0.09965088	0.1962875	-0.3149294	-0.03369097	0.139103	-0.3274323
17	-0.1635922	-0.1989825	0.004053556	0.0001623268	-0.05413973	0.07029188
18	-0.2350738	0.02274453	0.3560613	0.3046423	0.1647533	0.2413605
19	0.04555789	-0.2556796	-0.01501346	0.4187854	0.1109433	0.09703121
20	0.04552112	0.2718435	0.1146549	-0.4477859	0.2158506	0.07270873
21	-0.1032361	-0.2108703	-0.3551191	-1.184664	0.09993166	0.1633009
Output	0.01117459	-0.0588536	-0.1073852	-0.2674082	0.2670166	0.03936384

Table A3 Weights or connections between inputs, hidden nodes and output **Cont.-3**

	Hidden nodes					
Input	**19**	**20**	**21**	**22**	**23**	**24**
Bias node	0.08584581	0.1243483	-0.02234053	-0.05986464	-0.1845183	0.04965009
1	-0.1121786	0.2646293	0.1625935	0.1214402	0.03052167	-0.1563556
2	0.1394639	0.02576391	0.01051477	0.3310955	0.006579227	-0.4239371
3	0.0530801	-0.04539276	-0.153732	-0.1717058	0.1882178	0.1466145
4	-0.4092653	-0.1073288	-0.02045662	-0.4637154	-0.3711769	-0.1884108
5	0.1516389	0.1584874	-0.05297685	-0.01709982	-0.08583638	-0.4457475
6	-0.03165682	0.1790209	0.04852888	-0.1726884	-0.09242931	-0.03143305
7	-0.05176743	0.1178332	0.04982506	0.1980895	-0.04109903	0.04334866
8	-0.2384306	0.1937521	0.08310107	-0.0826579	0.2337932	-0.1031733
9	0.002516559	-0.0672032	-0.1885795	-0.2878733	-1.463154	-0.2907307
10	-0.06628494	0.2400831	0.06524978	-0.1048979	0.4768025	0.2322064
11	0.0226045	0.1973555	0.1350772	0.1178879	0.310793	0.277182
12	-0.01887084	-0.3549605	-0.1341304	-0.562274	0.2029816	-0.02991663
13	-0.2950478	0.4196094	-0.1517647	0.3762508	-0.09299771	-0.04090174
14	-0.5807862	-0.1397318	0.2798608	-0.07614637	-0.2989197	-0.01480066
15	0.2411328	0.01943525	-0.08368584	0.0005741047	-0.01988591	-0.07729054
16	-0.4540434	-0.2125352	0.4028892	-0.1832668	0.4221692	-0.3757408
17	0.09416708	0.02190033	0.1666061	-0.03713343	0.2677509	-0.3002317
18	0.1605258	-0.351274	0.005338179	-0.3035541	-0.1758282	-0.0301113
19	0.1494842	-0.1878613	0.02268256	-0.3664077	-0.122474	0.340541
20	-0.3741336	-0.08922915	0.01071079	-0.08486085	-0.07114223	0.2804206
21	0.05953203	-0.001112021	-0.188725	-0.4274724	0.3681357	0.1002251
Output	-0.09515741	-0.06506345	-0.02706111	0.01155625	-0.2716323	-0.01537478

Table A3 Weights or connections between inputs, hidden nodes and output **Cont.-4**

	Hidden nodes					
Input	**25**	**26**	**27**	**28**	**29**	**30**
Bias node	-0.1429496	-0.1192754	-0.05863185	0.1410667	-0.05129825	-0.06844173
1	0.1397846	0.442761	-0.2982509	-0.210014	-0.06175139	0.3635992
2	-0.09871813	-0.3135684	0.1111997	-0.0397209	0.1252317	-0.3329602
3	0.09264795	-0.08841824	-0.02528405	0.3168457	0.1520922	0.1643082
4	0.4738354	0.97594	-0.2767795	-0.2212452	0.2347015	0.2573489
5	-0.3862199	0.1970153	0.1657217	0.04664168	-0.4082733	-0.1474861
6	0.06435598	0.1097391	0.2024295	-0.1167798	0.4673467	-0.018484
7	-0.3350072	-0.4431129	-0.2469054	0.02391823	0.1458947	0.2093275
8	0.2502055	0.3140481	0.3538669	0.4175343	0.1571677	-0.202
9	0.09308682	0.2475504	0.9025722	0.006756761	0.1464525	0.0987551
10	0.07613741	0.0221862	0.8632655	-0.149032	0.05164863	-0.09330665
11	-0.05854018	0.3229609	-0.5932267	0.06681001	-0.05702825	-0.3980624
12	0.3955078	0.1097511	-0.3948796	0.323947	0.194149	0.2254201
13	-0.1835991	0.08795334	0.1791442	0.3333724	-0.0178038	0.2708406
14	0.4092117	0.4067489	-0.6910704	-0.003897951	0.190485	-0.2286844
15	0.2286459	-0.2018591	0.1622194	-0.1066861	0.2208299	0.08915935
16	0.1575738	0.04992457	0.09474633	-0.2996542	0.05937169	0.2763729
17	-0.190293	-0.3061233	-0.2932158	-0.1262429	-0.2779028	0.03276665
18	-0.9575033	-0.5106862	0.3136124	-0.07760495	0.1832739	0.3522944
19	-0.05682746	0.004289	-0.03799717	0.001228485	0.2272659	0.1706874
20	0.08756431	-0.2191494	0.07032362	-0.262199	0.1443691	0.1124723
21	-0.2289999	-0.1197127	0.3649592	-0.2575034	-0.2140495	0.06218997
Output	0.1560965	0.1635464	-0.1744417	-0.06318633	-0.02719446	-0.1539386

Table A3 Weights or connections between inputs, hidden nodes and output **Cont-5**

	Hidden nodes					
Input	**31**	**32**	**33**	**34**	**35**	**36**
Bias node	0.3514586	0.1733897	0.4220018	0.3234284	0.5567493	0.208932
1	0.09102236	0.2977746	-0.03471154	0.4243974	0.03273095	0.288314
2	0.1795582	-0.001334896	-0.1272895	-0.2984175	0.1289636	-0.01242756
3	-0.2269806	-0.2400785	0.06364882	0.07579957	-.0009181055	0.05875435
4	0.1845667	-.0005847945	0.04764383	0.1034925	0.2381215	0.5379001
5	0.3885275	0.4711095	0.04742411	0.008169858	-0.1197631	-0.1512871
6	-0.2160834	-0.007680825	0.2845152	-.0008977438	0.3333493	-0.4593
7	0.1044482	-0.1292439	0.2540763	-0.3282228	-0.07347839	0.2643384
8	0.4761531	-0.1963271	0.05471276	-0.3829658	0.381963	-0.09336281
9	0.1156238	0.0106438	0.6961717	0.1853617	1.019239	-0.1473108
10	-0.09319075	0.1940297	-1.555359	0.3475871	1.620501	0.123946
11	-0.1102014	0.1333903	1.232898	-0.2167428	-0.3406953	-0.09645357
12	0.01864402	0.2256263	-0.1714638	-0.1168836	-0.05198577	0.1371778
13	-0.3400288	0.2285052	0.1490249	0.4173501	0.09924134	0.05758775
14	-0.2235084	-0.2203324	-0.2956777	-0.09369265	-0.3097342	-0.2538799
15	0.2056943	0.001959526	0.01182012	0.1171089	0.3913536	-0.2093054
16	0.123883	-0.2524897	0.2237914	0.1995755	-0.0825698	0.5515017
17	0.2145022	-0.1030453	0.3377087	-0.3329197	-0.006302873	0.2326195
18	0.009336689	-0.1364666	0.274386	0.2708172	-0.1015428	0.5878126
19	-0.05415836	-0.04421069	0.03896356	0.05323379	-0.02571043	-0.2623921
20	0.1944528	-0.05426801	-0.1536645	-0.2599445	-0.0559502	-0.1209301
21	-0.2378369	-0.4573187	0.06299928	0.2355399	-0.02261765	-0.07839889
Output	-0.03214107	0.03410877	0.2874691	-0.05646313	0.3492595	-0.07450539

Table A3 Weights or connections between inputs, hidden nodes and output **Cont-6**

	Hidden nodes					
Input	**37**	**38**	**39**	**40**	**41**	**42**
Bias node	0.01091707	-0.1469539	-0.1640032	0.1865424	0.3975089	0.2567771
1	-0.1696563	-0.009419162	0.1231273	-0.3522134	0.3947432	-0.252128
2	-0.2801153	0.3074552	-0.1122509	-0.1874623	0.2594727	0.06874792
3	-0.191442	0.05491044	-0.1960489	0.1522526	0.09227718	0.08869352
4	0.4954261	0.05320991	-0.0175485	-0.3086458	-0.2833367	-0.2517042
5	-0.4698385	-0.5558584	0.007350664	-0.151342	-0.0864132	-0.01742521
6	0.3732417	0.4337604	0.5408698	-0.171924	-0.06320605	-0.1545171
7	0.1528294	0.2372282	0.03770661	-0.1202398	0.01063974	0.1868102
8	-0.337082	-0.1449734	0.1078971	0.299151	0.02352034	-0.03086248
9	-0.6278818	0.3588083	0.1517482	0.04108832	0.1535784	-0.07754054
10	0.6808332	-0.1552298	0.1711394	0.1443528	0.2042402	0.2272982
11	0.09333429	-0.5503617	0.1102305	-0.169465	0.09154314	0.511315
12	-0.06951788	-0.1348674	0.2687145	0.1544334	-0.2013501	-0.1216527
13	-0.1007578	-0.04161844	-0.2132023	-0.08072167	0.06011536	-0.1592374
14	0.009445232	0.3655342	0.08835927	-0.1434218	-0.01697447	-0.1105163
15	0.0187955	0.2741618	0.1431806	-0.1953572	-0.3071392	0.1159215
16	-0.1843634	0.1580892	-0.03398314	0.00308113	0.2639525	-0.2576469
17	0.08456842	-0.03381879	-0.05837958	-0.3747206	-0.3326216	-0.238019
18	0.1804759	-0.09264485	0.30673	0.04100203	-0.1168516	-0.05358911
19	0.2293523	0.06784963	0.1737613	-0.06172907	-0.03264432	-0.5410492
20	0.4620198	-0.3225258	0.2367803	-0.02806245	-0.1840097	0.2442598
21	-0.206596	0.2059972	0.116368	0.7425524	0.3963484	0.1999023
Output	0.08905096	0.04471385	-0.01243046	-0.05614969	-0.05915197	-0.06647453

Table A3 Weights or connections between inputs, hidden nodes and output **Cont-7**

	Hidden nodes					
Input	**43**	**44**	**45**	**46**	**47**	**48**
Bias node	-0.9076536	-0.2398529	-0.3081558	-0.1131299	-0.1801369	0.1111174
1	0.1015136	0.1455361	-0.2319213	0.04840022	0.1368264	0.02802893
2	0.1995923	0.2320718	0.3474172	0.128001	0.01692289	0.3627034
3	-0.1090407	0.03245578	-0.04894279	-0.2128158	0.2282056	-0.3088135
4	0.0414014	0.2307375	-0.8526269	-0.2461337	-0.3472747	-0.1505708
5	-0.03861913	0.3296937	0.03853192	-0.07077989	0.1560263	-0.04482055
6	0.1075116	0.09631141	-0.5790787	0.2691324	-0.08340344	0.09130825
7	-0.3384977	0.1744242	0.1555211	-0.08636696	0.0007659465	0.219349
8	-0.03409442	0.1766465	0.07883389	-0.1874392	-0.1432049	0.1846649
9	0.8278966	0.06200631	-0.01150028	0.03511063	-1.228926	0.3197109
10	-2.296747	0.03305441	-0.07000794	0.9560238	-0.293322	-0.4517408
11	-1.032114	-0.05011322	-0.1020639	-0.06183407	-1.105301	-0.08374374
12	-0.04084425	-0.1399569	0.3786881	0.1804466	0.03436669	0.2335551
13	0.05866638	0.2019725	-0.07839477	0.3835366	-0.1288069	-0.2589995
14	0.0363731	-0.04121115	-0.08776636	0.115942	0.1244129	-0.1408371
15	-0.004732447	-0.03683817	0.02256382	-0.08827572	0.02535248	0.3701559
16	0.1398767	-0.2275034	-0.08674665	-0.1115357	-0.1283959	0.08168368
17	0.05820399	-0.03904632	-0.8624134	-0.003588579	-0.03902706	-0.1064855
18	0.06875566	-0.2164527	0.321089	-0.3418404	0.3588492	0.1704323
19	-0.1758053	0.5983115	0.07497753	0.082609	0.195039	0.06242352
20	-.0003775914	-0.2301843	0.9713785	-0.01687809	0.2071267	-0.1670057
21	-0.1776008	0.2012305	-0.3213136	-0.2015512	0.1357068	0.3758746
Output	-0.4088761	0.003516561	0.2722255	-0.04218724	-0.2415035	-0.04055905

Table A3 Weights or connections between inputs, hidden nodes and output **Cont-8**

	Hidden nodes					
Input	**49**	**50**	**51**	**52**	**53**	**54**
Bias node	0.7029582	0.1336593	-0.01316882	-0.1610887	-0.6524774	1.148925
1	0.0754749	-0.1761206	0.1192613	-0.5883685	-0.02420491	0.2995353
2	0.2927596	0.283415	-0.5170856	0.03523048	0.07682834	-0.08801506
3	-0.02897758	-0.006489809	0.4173642	0.1329181	0.02089692	-0.09926859
4	-0.3288638	0.4895605	-0.1782553	0.1765337	-0.05245219	-0.1418193
5	-0.24263	-0.13059	-0.3502538	-0.08865563	-0.07431578	-0.3428421
6	0.371498	-0.07532045	0.4720257	0.1298303	-0.01461464	-0.02499953
7	0.2569417	0.5718542	0.04514321	-0.1863626	0.4266718	0.3592507
8	0.07232363	-0.05396434	0.3815588	0.4950414	0.5271809	0.08443551
9	-1.246807	-1.044651	0.1380863	-0.4517323	0.2181381	2.143632
10	0.7878147	-1.375451	-0.2215345	-0.3487799	-0.1924145	0.005278419
11	-1.741257	-0.5913573	-0.1175489	-0.1071157	-0.004843754	-0.7115466
12	0.22481	0.01809514	-0.2162864	-0.1939506	-0.468166	-0.07518509
13	0.0886329	0.2472545	-0.1718321	0.05511134	-0.1218468	0.2625022
14	-0.2625388	-0.1565296	-0.3973616	-0.1787707	-0.1418914	-0.501178
15	0.193447	-0.2774159	0.03433248	0.2558587	-0.07340495	0.09611694
16	-0.3481868	-0.03601308	0.5239779	-0.1662123	-0.2496414	-0.009535003
17	0.4060482	0.08005638	0.4280148	-0.001550885	0.2966492	0.1381219
18	0.2907229	-0.0322107	-0.02530702	0.1130399	0.5328767	0.06054716
19	0.258614	0.1178056	0.3574032	-0.08688555	-0.01927163	-0.2211788
20	0.1717872	-0.08080833	0.1253623	-0.03747219	-0.1896331	0.1497116
21	0.1157063	0.08538339	-0.12973	0.06646191	-0.2554773	0.09602892
Output	0.2590824	0.3162172	0.02378339	0.1071596	0.1359412	0.421989

Table A3 Weights or connections between inputs, hidden nodes and output **Cont-9**

	Hidden nodes					
Input	**55**	**56**	**57**	**58**	**59**	**60**
Bias node	-0.3660806	0.3837737	0.102364	-0.08262884	0.04738584	-0.2331962
1	-0.5054355	0.08524472	0.239901	-0.3236169	-0.6155861	-0.3133575
2	-0.2862814	-0.0455575	0.4714568	-0.1141468	-0.02642108	-0.2015774
3	-0.4445609	-0.1157679	-0.01566625	0.1605424	-0.08938028	0.5313287
4	-0.01636017	0.3304388	0.4165463	0.9172	0.1447988	-0.0806
5	0.1579828	-0.0121168	0.14175	-0.02177271	0.1039467	-0.1739028
6	-0.2455575	-0.08213586	-0.1940005	-0.1324463	0.2816758	-0.2149546
7	-0.04227765	-0.1979544	-0.1879202	-0.02599976	0.03961884	0.06175005
8	-0.07024531	-0.001444842	-0.036112	0.1401799	0.05286179	0.09763706
9	-0.1093562	0.006742487	0.1570502	0.3149869	0.2001956	0.1408415
10	0.06263973	0.1985009	0.3839115	0.2919152	-0.1163149	-0.1117699
11	0.3091238	0.1464648	0.6081826	-0.5517871	0.07178653	-0.1009575
12	-0.01142897	0.08216016	-0.4780216	-0.1334808	-0.06586625	0.3030162
13	-0.0189111	0.08165526	-0.2428648	-0.1419238	-0.174893	-0.1706632
14	0.1067545	-0.1401771	0.3240831	0.156642	0.06616341	-0.03144213
15	-0.02435729	0.2818825	0.6870058	0.1881082	0.4649261	-0.01224387
16	-0.05390855	0.2192083	-0.116127	-0.3383742	0.4822431	-0.14177
17	-0.1864572	-0.1577585	0.008883149	0.4274206	-0.07290415	0.2261055
18	0.07196519	0.07340334	0.178941	0.1166863	0.07432626	0.2686896
19	-0.04076903	0.3578089	0.1254043	-0.5577849	0.05762296	0.07122948
20	-0.01504896	-0.26467	0.02038623	-0.0293799	-0.3999664	0.08513831
21	0.1767193	0.008003578	0.5132089	-0.1276944	0.4434417	0.003063897
Output	0.0269067	-0.07082888	-0.1412271	-0.1806263	-0.05117295	-0.001564818

Table A3 Weights or connections between inputs, hidden nodes and output **Cont-10**

	Hidden nodes		
Input	**61**	**62**	**Output**
Bias node	0.1321102	0.1376745	-0.3328806
1	-0.1048353	-0.04540023	
2	-0.1563369	0.03283948	
3	0.04225446	-0.1215337	
4	0.1334621	0.1382665	
5	-0.06361276	-0.2185612	
6	-0.2957467	0.2930357	
7	-0.1554609	0.03310549	
8	-0.1732534	0.1041571	
9	-0.08988244	-0.1474887	
10	0.3136398	-0.2487977	
11	0.1480444	-0.1193671	
12	-0.1258541	0.07707145	
13	-0.1583689	0.1438563	
14	-0.02064096	0.0623262	
15	0.3393562	0.2747854	
16	-0.0504894	0.08620544	
17	-0.1081671	-0.06590912	
18	-0.1181983	-0.09719059	
19	0.06223637	0.2149623	
20	0.2163232	0.1821264	
21	0.02277104	-0.3145674	
Output	-0.0183004	0.0210209	

For output, after internal calculation, the output was de-scaled linearly back to the value of log10(RelC) with de-scaling coefficients showed in the last line of table A2.

Appendix B

P-G stability definition

Table B1 Definitions of Pasquill Stability Categories (Pasquill and Smith, 1983)

Surface Wind speed (m/s)	Daytime solar insolation			Night	
	Strong	Moderate	Slight	Thinly overcast, or ≥ 1/2 low cloud	≤ 3/8 cloud
<2	A	A-B	B	-	-
2-3	A-B	B	C	E	F
3-5	B	B-C	C	D	E
5-6	C	C-D	D	D	D
>6	C	D	D	D	D
For A-B take average of values for A and B etc.					

A: Extremely unstable	D: Neutral
B: Moderately unstable	E: Slightly stable
C: Slightly unstable	F: Moderately stable

where strong insolation corresponds to sunny midday in midsummer in England, slight insolation to similar conditions in midwinter; Night refers to the period from 1 hr before sunset to 1 h after dawn. The neutral category D should also be used , regardless of wind speed, for overcast conditions during day or night, or for any sky conditions during the hour preceding or following the night as defined above. (Pasquill and Smith, 1983, pg. 336).

Appendix C

COMBIC Input card (Validation set 1)

```
WAVL           0.532  0.000000  0.000000
VIS           15.000
COMBIC
PHAS             1.0       3.0       6.0      12.0       9.0       0.0
1.0
FILE             9.0 H_sg09.out
NAME
cloud: s010815sg09
MET1         97.8300    0.2600       5.6      9.89   1002.69       0.0
0.0
MET2          2.5600      0.00       0.0      0.00      0.00       0.0
TERA          0.0500      0.00       0.0      0.0
MUNT          1.0000    0.0441       0.0     24.00    100.00       1.0
1.0
EXTC         24.0000    0.3700       1.0      1.00      1.00       1.0
1.0
CLOU          1.0000
SUBA          1.0000    1.0000       0.0      1.00      2.00      24.0
0.0
SUBB         6.92701   7.89262   2.49195      0.01      9.89       0.0
0.0
SUBC          5.5000    0.0000       1.0      0.00      1.00       0.0
0.0
DONE
END
CONTINUE
WAVL           0.532  0.000000  0.000000
VIS           15.000
COMBIC
PHAS             2.0       3.0       6.0      12.0       9.0       0.0
1.0
FILE             9.0 H_sg09.out
FILE            12.0s010sg0902.txt
NAME
Cloud: s010815sg0902, at 22.95s
ORIG             0.0       0.0       0.0      90.0    270.00       0.0
LIST             2.0     22.95     22.95      0.00
SLOC             1.0      1.00      0.00    150.00      0.00      0.00
5.5000
OLOC             1.0      14.0       0.0      6.00      0.00    150.00
TLOC             1.0      14.0       0.0      5.00       1.0
EXTC            24.0       1.0       1.0       1.0       1.0       1.0
1.0
VIEW             1.0       1.0      27.0      27.0      27.0      27.0
90.0
```

```
GREY              9.0       0.05       0.95       1.0        1.0        1.0
0.0
TPOS              0.0        0.0        0.0       0.0        0.0       11.0
11.0
DONE
END
STOP
```