NOAA Technical Memorandum ERL/ARL-218

TECHNICAL REVIEW OF THE VLSTRACK DISPERSION MODEL

W. R. Pendergrass
C. J. Nappo
K. S. Rao
R. M. Eckman

Air Resources Laboratory
Silver Spring, Maryland
November 1996

noaa NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / Environmental Research Laboratories

TECHNICAL REVIEW OF THE VLSTRACK DISPERSION MODEL

W. R. Pendergrass
C. J. Nappo
K. S. Rao
R. M. Eckman
Atmospheric Turbulence and Diffusion Division
Oak Ridge, TN

Air Resources Laboratory
Silver Spring, Maryland
November 1996

UNITED STATES
DEPARTMENT OF COMMERCE

Michael Kantor
Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION

D. JAMES BAKER
Under Secretary for Oceans
and Atmosphere/Administrator

Environmental Research
Laboratories

James L. Rasmussen
Director

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA/ERL. Use for publicity or advertising purposes, of information from this publication concerning proprietary products or the tests of such products, is not authorized.

ATDD Contribution File No. 96/03

For sale by the National Technical Information Service
5285 Port Royal Road, Springfield, VA 22161

TABLE OF CONTENTS

Page

LIST OF FIGURES v

LIST OF TABLES v

ABSTRACT vii

1.0 INTRODUCTION 1
- 1.1 Difficulties In Dispersion Modeling 1
- 1.2 VLSTRACK Background 1
- 1.3 The Model Validation Process 2
- 1.4 Report Organization 2

2.0 SPECIFIC TASKS 3

3.0 VLSTRACK MODEL DESCRIPTION 4

4.0 TECHNICAL COMMENTS 4
- 4.1 General Comments 4
- 4.2 Task 1. Methodology and Equations 6
  - 4.2.1 Computation of Monin-Obukhov length 6
  - 4.2.2 Temperature profiles 7
  - 4.2.3 Wind variability adjustment 7
  - 4.2.4 Concentration algorithms 8
  - 4.2.5 Deposition 9
- 4.3 Task 2. Empirical Parameters 10
  - 4.3.1 Mixing layer 10
  - 4.3.2 Parameterization of turbulent velocities within the boundary layer 11
  - 4.3.3 Parameterizations above the boundary layer 12
  - 4.3.4 Dispersion parameters 12
- 4.4 Task 3. Model Validation Studies 13
  - 4.4.1 Introduction 13
  - 4.4.2 Documentation 14
  - 4.4.3 Validation review 15
    - 4.4.3.1 Review of approach 15
    - 4.4.3.2 Final report 15
  - 4.4.4 Validation protocol 16
    - 4.4.4.1 Comparison data 16
    - 4.4.4.2 Biological comparison data 17
    - 4.4.4.3 Chemical comparison data 19
  - 4.4.5 Statistical Technique 20
  - 4.4.6 Parameter Sensitivity Study 21
  - 4.4.7 Concluding remarks 22

5.0 CONCLUSIONS AND RECOMMENDATIONS 23
- 5.1 Model Documentation 23

5.2 Model Formulation 23
5.3 Model Validation 24

6.0 REFERENCES 26

LIST OF FIGURES

Page

Figure 1 Conceptual view of atmospheric processes that can act upon effluents released into the atmosphere. 34

Figure 2 Conceptual model of shoreline fumigation based on field observations. From Lyons (1975) 35

Figure 3 Potential transport trajectories and plume concentrations for various averaging times. 36

Figure 4 Illustration of problem associated with sample size and mean difference factor. By selecting more sample points near the source (or beyond the intersection point of the two curves) weights the mean difference factor towards either under- or overprediction 37

Figure 5 Comparison between observed and predicted doses for chemical Trial 20. VLSTRACK typically underpredicts dose level near the source and overpredicts with distance. 38

Figure 6 Observed and predicted doses as a function of downwind distance for Trial 20. Both VLSTRACK and simple model predictions are plotted 39

LIST OF TABLES

Page

Table 1 Evaluation of Table 4 (Final Report) which estimates 0 to 10 min. wind direction and velocity for those chemical trials in which only 2 and 10 min. averaged data were reported 29

Table 2 Chemical data for Trials 2.28 and 2.29 30

Table 3 Observation data for chemical Trial 21 31

Table 4 Data for chemical Trials 19 and 20 32

Table 5 Power law exponent (p) values for different averaging times and stabilities (IAEA, 1989) 33

(blank page)

ABSTRACT

The United States Army Nuclear and Chemical Agency requested NOAA's Atmospheric Turbulence and Diffusion Division (ATDD) make an independent technical review of the Vapor, Liquid, and Solid Tracking (VLSTRACK) model as part of the overall Chemical and Biological Warfare Hazard Study. The specific purpose of this review was to provide an objective opinion on the approach, methodology, and technical formulas used in the VLSTRACK model. The Study Advisory Group will use the results and recommendations of the review to improve and add credibility to the hazard prediction software currently under development. The review by ATDD focused on the model physics and the meteorological equations used, and whether their use and application are in accordance with standard acceptable practices. The study also reviewed the empirical parameters in VLSTRACK, the techniques used to determine appropriate values for these parameters, and the model validation methods. This study was based on examinations of model descriptions and other supporting documents. Although access to the computational code was provided, the model was not actually executed. The purpose of this report is to indicate all problems with the model (subject to the limitations noted above) from both scientific and operational view points regardless of their overall significance. Because of the nature of the review, the report is concerned only with the discrepancies of the VLSTRACK model and not its positive aspects. Funding and response time for this study were quite limited. Consequently, the study was limited in scope to the most obvious model errors. A thorough examination of the VLSTRACK model including its execution and sensitivity testing was beyond the scope of this study.

Serious problems were discovered in many of VLSTRACK's physical elements, including the velocity standard deviations, concentration algorithms, atmospheric stability parameterization, surface deposition, and model tuning. On the basis of our examination of the VLSTRACK documentation, we concluded that the model is scientifically flawed. We note that the arbitrary tuning of model parameters has resulted in "better" model performance than would be the case otherwise. Since this tuning has no physical basis, the model gives the better or improved results for the wrong reasons, and thus violates one of the basic rules of model verification. Further, the tuning was based on several tracer studies performed at essentially the same location under similar meteorological conditions. Hence, we feel that the model's general applicability has not been demonstrated.

(blank page)

1.0 INTRODUCTION

1.1 Difficulties In Dispersion Modeling

The dispersion of pollutants in the atmosphere is a complex phenomenon that is influenced by many different processes. Some of the processes are illustrated in Figure 1, reproduced from Randerson (1984). Although this figure is concerned with power plant emissions, the same physics acts on all atmospheric releases. Plume dilution is achieved by the combined effects of dispersion due to atmospheric turbulence, dry and wet deposition, and chemical transformation. Atmospheric turbulence is a function of solar heating of the ground surface during the day, and its long-wave radiative cooling during the night. Turbulence and the resulting plume dispersion are strong during sunny days, and weak during cloudy days and at night. Wind speed also affects turbulence; generally, turbulence increases with increasing wind speed. Dry and wet deposition processes remove material from the polluted air and deposit it on the ground surface. Under certain conditions, the deposited material can be resuspended into the atmosphere. Chemical and physical processes can also affect plume dilution in several ways. If the chemical reaction is exothermic, then the released heat can add to plume buoyancy and plume rise. Because of chemical reactions, concentrations of the reacting species will decrease with downwind distance, but the products of these reactions will increase with distance.

Terrain effects add even more complexity to the atmospheric dispersion problem. Rough surfaces, for example, will generate more turbulence than smooth surfaces, for the same wind speed. Hilly and mountainous terrain affects both the transport winds and the turbulence levels, and can also lead to plume impacts on elevated terrain. The simple and commonly used air pollution models that assume flat terrain are not particularly useful in complex terrain.

In this report, we present the results of a technical review of the Vapor, Liquid, and Solid Tracking (VLSTRACK) atmospheric dispersion model. Many of the processes illustrated in Figure 1 are parameterized in the VLSTRACK model, and thus this model is highly complex. It is this complexity that confounds the model validation process. In a simple model, such as the Gaussian plume model, wind speed, diffusion, and source strength, are directly related to the predicted concentrations. However, in a model such as VLSTRACK, the "cause-and-effect" lines are not obvious. Small errors in one parameterization can lead to large errors in another. Relating model inaccuracy to a specific parameterization is generally not possible. However, if each parameterization is based on sound physics, then it is reasonable to assume that the model predictions are as accurate as possible. This is the underlying premise of the model review presented here. We have not executed the VLSTRACK model. Our results are based solely on the documents describing the model physics and the model verification and tuning procedures.

1.2 VLSTRACK Background

From 1990 to 1993, the Naval Surface Warfare Center (NSWC) in Dahlgren, VA, has worked on developing a transport and diffusion model for use in predicting the downwind hazards of chemical and biological agents. The VLSTRACK computer model, Version 1.51 and Version 2.0, (Bauer, 1993; Bauer and Gibbs, 1994) is intended to provide approximate downwind hazard predictions for many currently known or suspected chemical and biological agents and a wide variety of munitions capable of disseminating these agents. In addition, the model was designed to be used not only for operational hazard assessments, but also for research and development studies and training applications, with all operations being "user-friendly".

In 1992, the U.S. Army Nuclear and Chemical Agency (USANCA) began work on an advanced hazard prediction system. The proposed system would integrate the effects of variable meteorology and complex terrain with a transport and dispersion model so as to obtain a more realistic hazard prediction capability. In August 1993, the Study Advisory Group (SAG, a group formed and charged with overseeing the model development study) decided to use the VLSTRACK model in this advanced hazard prediction system. However, since some of the SAG members expressed reservations, USANCA requested an independent technical review of the VLSTRACK model as part of the overall Chemical and Biological Warfare (CBW) Hazard study. The SAG intends to use the results and recommendations of this review to improve and add credibility to the hazard prediction software currently under development.

The NOAA/Air Resources Laboratory's Atmospheric Turbulence and Diffusion Division (ATDD) has previously performed many transport and dispersion model evaluation studies. Because of their expertise, the ATDD was asked to perform a technical review of the VLSTRACK model. The purpose of this review was to provide an expert opinion on the approach, methodology, and technical formulas used in the VLSTRACK model. The review focused on the physics and meteorological equations, and whether their use and applications are in accordance with standard acceptable practices. The study also focused on the empirical parameters used in the model, the techniques used to determine appropriate values for these parameters, and model validation tests.

1.3 The Model Validation Process

Prediction of the consequences of releases of hazardous materials into the environment has required the recent development and application of numerical models that focus on "episodic" releases, as opposed to the more usual regulatory "long term average" models applicable to criteria pollutants. While the operational emphasis has shifted to the instantaneous event, the numerical techniques commonly employed are still those originally developed to treat long-term ensemble averages. Along with the development of event-driven models, it has been recognized that simply providing a concentration at some point in space and time is no longer sufficient. Models can only provide estimated concentrations, which must be accompanied with estimated levels of uncertainty. A model prediction without an estimate of its associated uncertainty is of little use to a decision maker; this is particularly true for application of ensemble-average models to episodic events.

The model validation process is designed to provide a measure of the model's level of uncertainty by means of statistical analysis of the model's predictive capability. The goal of the validation process is to provide a quantitative measure of model reliability in the form of confidence limits associated with some stated degree of accuracy (e.g., we have 95% confidence that the model predictions will be within some given factor of the observed concentration value). The validation process should address a number of issues including how well the model has been formulated to address the stated problem, model parameterizations, and its calculation methodology. The process of model validation also requires selection of a suitable comparison database and use of appropriate statistical techniques to evaluate the model's predictive capability.

1.4 Report Organization

In this report, we present the results of ATDD's technical review. Due to the limited support and time allocated to this review, the reviewers concentrated on those model formulations and validation efforts which were obviously scientifically flawed which unfortunately presents a "negative" review regardless of technical merits of the code. Though this review was fairly comprehensive, it did not include aspects of the model such as liquid droplet evaporation and transformation, computational grid generation schemes,

and VLSTRACK's interface with the meteorological prediction model. The ATDD review considered only those topics that were within ATDD's areas of expertise and were explicitly addressed in this report.

In Section 2, the specific tasks of the review are described as presented in the Statement of Work. Section 3 provides a brief description of the VLSTRACK model; Section 4 presents our detailed findings. Throughout Section 4, we present recommendations for many of the specific items discussed with regard to the VLSTRACK model formulation and its validation. In Section 5, we provide our conclusions and recommendations.

2.0 SPECIFIC TASKS

The specific tasks listed in the Statement of Work (SOW) issued to the ATDD by USANCA were:

Task 1. Methodology and Equations.

(1) Examine the transport, diffusion, and deposition methodology of the model. Determine whether the approach and procedures are reasonable.

(2) Examine the physics and meteorological equations used in the model. Verify that their form and use are in accord with standard acceptable practices in the transport and diffusion community.

(3) Comment on the impact of the transport and diffusion methodology and equations on the desired performance of the VLSTRACK model.

Task 2. Empirical Parameters.

(1) Identify the empirical parameters in the physics and meteorological equations used in the VLSTRACK model.

(2) Determine whether the approach and procedures used to set the parameters are reasonable.

(3) Comment on the determination of the empirical parameters as they relate to the methodology and as they relate to the performance of the model.

Task 3. Model Validation Studies.

(1) Comment on the approach used to evaluate/validate the VLSTRACK model as compared with the approach used to evaluate/validate similar types of models.

(2) Comment on the expected performance of the VLSTRACK model as compared with other models with similar methodology. Reference applicable field tests as appropriate.

Task 4. Recommendations.

(1) Provide recommendations and/or provide a strategy to improve the VLSTRACK model.

Task 5. Documentation.

(1) Provide in paper copy and electronic media a final report and progress reports.

3.0 VLSTRACK MODEL DESCRIPTION

VLSTRACK is a puff-trajectory dispersion model. The puff concentration is prescribed by a three-dimensional Gaussian distribution. The following model information is taken from Bauer and Gibbs (1994).

The VLSTRACK Version 1.5 computer model provides approximate downwind hazard predictions for a wide range of chemical and biological agents and munitions of military interest. The code is written in standard FORTRAN, and is intended to be "user-friendly" and highly portable. Output can be obtained either as a cumulative hazard from the time of the attack or as a periodic hazard. The model is designed to accommodate variable meteorology which can be provided by a meteorological forecast. This feature is considered very important by the model developers for biological and secondary evaporation computations. A forecasted height profile of horizontal winds can also be used for high-altitude releases. VLSTRACK Version 2.0 can also use a horizontal plane of meteorological data (2-D) or full spatially variable meteorological data (3-D), both as functions of time. Although the model normally runs quickly, the rigorous computations can take an hour or more for biological plumes, for large numbers of munitions, and when secondary evaporation is considered. For quick estimates, the model features a rapid approximations option for each of these attack situations which can be used for preliminary hazard evaluation; the rigorous computations can then be done if a more accurate hazard estimate is required and time permits.

Many of the computations used in the VLSTRACK model come directly from the Non-Uniform Simple Surface Evaporation, Version 4 (NUSSE4) model developed at the U.S. Army Chemical Research, Development and Engineering Center (CRDEC) in Edgewood, Maryland. Modifications were made to some of these NUSSE4 computations to increase their range of applicability. The computations from NUSSE4 used with little or no modifications are for meteorological condition estimation of the Monin-Obukov length; surface roughness length and friction velocity; Lagrangian time scale and turbulence standard deviations; rate of cloud expansion with time, droplet shape, and terminal velocity; droplet Sherwood number; vertical velocity profile near the ground; ground temperature; log-normal droplet size distribution approximation; and overlapping point source approximation of a line source.

4.0 TECHNICAL COMMENTS

4.1 General Comments

The VLSTRACK 2.0 draft Software Design Document (Bauer, 1993) reads more like a military document than a model user's guide; it is not very readable or user-friendly. Part of the problem can be attributed to attempts to restate or rewrite the various equations taken from NUSSE3 (Saucier, 1987), which is a clearer and more readable document. In this process, the intermediate steps, assumptions, and explanations were lost; their references were omitted or lumped together at the end with those belonging to other steps; and the equations became unrecognizable from those generally given in the literature. The various assumptions and empirical constants were not clearly stated. Much of the mathematical notation is also different from previous reports and that used in the meteorological literature. It would be useful to keep the familiar notation used in previous reports in the various equations, to list intermediate steps, and to provide more information and applicable references at each step to improve clarity of presentation.

Another problem is the interspersing of the text on the model formulations with the many Tables for I/O variables of the various Computation Functions. These tables only list the computer code names,

and not the corresponding mathematical variables used in the equations. This being the case, it is puzzling why they were allowed to intrude throughout the text and make it unreadable. It is highly desirable to separate the report into two clear sections: the first section consisting only of the model formulations and related technical discussions; and the second section consisting of the computer code User's Guide, which includes the tables for the I/O variables (clearly identifying the related equations and variables) and all other information that would explain the model implementation and assist the model users.

After studying the documentation for the VLSTRACK model as described in detail in Bauer (1993), we have concluded that the model, in its present form, has numerous technical flaws. Compared with other puff models, VLSTRACK contains empirical parameters that are either undocumented or difficult to justify. For example, Eq. (21) in Bauer (1993) contains numerical factors of -2.5 and 0.4 that do not appear in the original reference (Holtslag and van Ulden, 1983). As another example, for diffusion over water, the velocity standard deviations σ_u and σ_w are arbitrarily multiplied by a factor of 0.75. Since the reduced turbulence levels over water are already accounted for by using a smaller roughness length z_0, this factor seems unjustified. On p. 3-96 of Bauer (1993), an unexplained factor of 2000 appears in the equation for the wind variability adjustment distance.

The "tunings" performed by Bauer and Wolski (1993) of the empirical equations for σ_u, σ_w, and the diabatic adjustment to the wind profile (denoted by η in the documentation) are especially problematic. Basically, the empirical relations for these variables were adjusted to produce a better fit between the VLSTRACK output and field measurements of biological dosage and deposition. Since the dosage and deposition patterns are affected by just about every subroutine in the model, there is no obvious reason to believe that any discrepancies between the observations and model estimates are solely due to these variables. This tuning is not in accordance with accepted practices in the dispersion community, since any number of problems in a model can be swept under the rug using this approach.

The documentation for VLSTRACK contains a number of apparent misunderstandings about atmospheric structure and nomenclature. For example, velocity standard deviations are incorrectly called turbulence intensities, and the standard tropospheric lapse rate of 0.0065°C/m is incorrectly called the adiabatic lapse rate, which is actually 0.0098°C/m. The nomenclature of the report's description of the mixing and boundary layers needs improvement; for example, it does not make much sense to talk about temperature profiles changing direction. It is also incorrectly stated that the free troposphere above the boundary layer has neutral stability (it is on average stable).

All but one of the chemical agent trials used by Bauer and Wolski (1993) in their validation of VLSTRACK involved detonation of an artillery shell or rocket, either on the ground or within several meters above ground. This restriction does not really demonstrate the general applicability of the model. Further, the model results are likely to be very sensitive to the split between primary and secondary vapor sources, the later resulting from evaporation of volatile agents. Setting the height of the secondary vapor cloud centers to 0.5 m is also arbitrary. There is no accurate way to determine these emission characteristics or their uncertainties for various agents.

Bauer and Wolski (1993) used in their initial validation analysis only those experimental data points that had a corresponding non-zero VLSTRACK 1.2 prediction value. In other words, zero predictions by the model were not considered for comparison with data. This is likely to bias the results. They also used only the so-called "difference factor" (a scaled logarithm of the ratio of predicted to observed concentration) in the evaluation, and claim that "use of difference factors lends a symmetry that the other methods lack". A popular and better-known statistic such as the fractional bias (see, e.g., Hanna, 1993; Rao and Hosker, 1993) also exhibits the same type of symmetry that the authors gave in their

example. In any case, the authors should use a number of statistical measures to evaluate the model performance (see, e.g., Tangirala et al., 1992; Ciolek, 1994) in order to have greater confidence in their conclusions.

4.2 Task 1. Methodology and Equations

The equations and page numbers given below refer to the VLSTRACK 2.0 Draft Software Design Document (Bauer, 1993).

4.2.1 Computation of Monin-Obukhov length

VLSTRACK uses the inverse Monin-Obukhov length Γ to calculate cloud dispersion parameters. If an estimate of the Pasquill stability category is available, then Γ can be evaluated from a simple equation, i.e., p. 3-72, Eq. (8). It should be clearly stated in the users's guide that Eqs. (8) and (9) are approximations given by Shir and Shieh (1974) to the graphical relations of Golder (1972) between surface roughness and Monin-Obukov length as a function of the Pasquill stability category. If an estimate of the Pasquill stability category is not available, then the VLSTRACK model calculates Γ using its mathematical definition, p. 3-74 Eq. (24), which is a function of the sensible heat flux **(H)** at the ground surface and the surface friction velocity. The calculation method uses a technique based on the work of Holtslag and van Ulden (1983) to estimate **H** at the surface; see p. 3-73, Eq. (21). Holtslag and van Ulden adapted a simplified approach to the Penman-Monteith equation for the partitioning of available energy (net radiation minus soil heat flux) into sensible and latent heat fluxes. Their equation was obtained for normal summer conditions in the Netherlands (roughly 52° N latitude) for a grass-covered surface, supplied with enough water to evaporate. These limiting assumptions are important and should be clearly stated in the user's guide. Next, their soil heat flux is typically about 10% of the net radiation (Holtslag and van Ulden, 1983), as also used in NUSSE3 (Saucier, 1987). However, VLSTRACK assumes the available energy to be 0.4 times the net radiation, which requires the soil heat flux to be 60% of the net radiation; this is very high and cannot be justified for any type of surface. This equation is therefore likely to significantly underestimate the surface heat flux.

There are a number of other problems associated with the VLSTRACK implementation of Holtslag and van Ulden's (1983) technique. One problem is the arbitrary assumption of a temperature difference $\Delta\mathbf{T}$ between the surface and the air; sand, for example, is assumed to have $\Delta\mathbf{T} = +10°C$ during the day and -10°C at night. But as shown by Louis (1979), both **H** and **L** are uniquely determined if $\Delta\mathbf{T}$, wind speed **U**, and surface roughness length $\mathbf{z_0}$ are specified. The use of Holtslag and van Ulden's technique therefore becomes superfluous if $\Delta\mathbf{T}$ is specified.

A second problem is that VLSTRACK applies Holtslag and van Ulden's technique over any surface and at any time of the day or year, whereas the technique has been demonstrated to work only in relatively limited situations. Holtslag and van Ulden clearly state that their technique only applies during daytime when the solar heating is strong. They also state that the technique may not work over rough surfaces such as a forest, and during the winter. More general formulations for surface fluxes, such as those by Jarvis and McNaughton (1986), indicate that the Holtslag and van Ulden approach may not work well over rough surfaces, dry areas, and in windy conditions.

Even if Holtslag and van Ulden's technique is assumed to be accurate, there still is the problem of estimating the fractional surface moisture, denoted by χ in the document. VLSTRACK assumes that χ equals 1.0 for water, 0.3 for sand, and 0.65 for everything else. With these rough estimates of χ, it is unlikely that the estimated heat fluxes will be very accurate. Can the grass in, say, Great Britain be expected to have the same χ as the grass in the American Midwest? Or can both wet sand and dry sand

be expected to have the same χ? Given such large uncertainties, the resulting estimates of atmospheric stability may not be any better than those obtained by using simple quasi-qualitative tables (e.g., Panofsky and Dutton, 1984, p. 242) to estimate Pasquill stability categories.

There is a problem with the value of the constant in Eq. (26) on p. 3-74, Eq. (53) on p. 3-96, and Eq. (57) on p. 3-98. The constant 1.2 in these equations is too low by about a factor of 4; this empirical constant should be 4.7 (Businger et al., 1971), as widely used in the micrometeorological research. It appears that the value of this constant was "adjusted" in VLSTRACK to fit the dosage and deposition data (Bauer and Wolski, 1993), which is not an acceptable approach. We believe it is inappropriate to adjust well-founded descriptions of atmospheric variables to force model outputs to agree with observed concentrations. The derivation of Eqs. (32) and (33) on p. 3-75 is not clear. Even if we accept Eq. (32) as correct, we believe the constant in the denominator should be 8.1 instead of 2.4.

Finally, paragraph 2 on p. 3-76 is poorly written, and requires careful editing and rewriting. The atmospheric stability class, Pasquill stability category, and Monin-Obukhov length are related to each other, as are the bulk-Richardson number ($\mathbf{Ri_B}$), temperature difference ($\Delta\mathbf{T}$), and horizontal wind direction fluctuation (σ_θ) over homogeneous flat terrain (e.g., Slade, 1968). When $\Delta\mathbf{T}$ and the wind speed are known, for example, the $\mathbf{Ri_B}$, P-G stability class and other parameters are all uniquely determined. The Golder (1972) nomograms also show that Monin-Obukhov length is related to the Pasquill stability class and the roughness length. Such relationships and dependencies among these various parameters should be recognized and clearly stated.

4.2.2 Temperature profiles

Equations (59) and (60) on p. 3-98 use 0.0065 C/m as the adiabatic lapse rate of temperature (as explicitly stated on p. 3-134) instead of 0.0098 C/m. This leads to wrong temperature profiles. Also, it is difficult to justify the use of the index **m** (which depends on the Pasquill stability category near the surface) for determining the temperature profile above the mixing layer height.

4.2.3 Wind variability adjustment

When the observed wind direction varies in space, the VLSTRACK model uses a "wind variability adjustment distance" to increase the cloud diffusion over complex terrain. We believe this adjustment is a bad idea. In a puff model such as VLSTRACK, a clear separation is made between small-scale atmospheric eddies whose effects are parameterized through the growth of the individual puffs, and larger-scale eddies that appear explicitly in the wind field as spatial variations in speed and direction. Observed variations in wind direction at different meteorology locations should only be used to produce variations in the wind field. By using the wind variability adjustment, the VLSTRACK model is basically counting the same eddies twice, once in the wind field and once in the puff diffusion. Since wind-direction measurements are usually averaged over ten minutes or more, there is no reason to believe that variations in the wind direction at different meteorology locations are somehow representative of the general level of turbulence. Instead, these variations may result from wind deflections around complex terrain, differences in elevation of the meteorology locations, or simply errors in the alignment of the wind instruments (in the field, it is difficult to align wind vanes to an accuracy of better than $\pm 5°$ to true North). A more reasonable way to account for increased diffusion in complex terrain is to use larger values of the roughness length $\mathbf{z_0}$, since this will tend to increase the computed friction velocity $\mathbf{u_*}$.

It is not clear why this wind variability adjustment factor has units of length or how it is used in the model. The report should clearly discuss why this parameter is necessary, and exactly how it is used in the model. Also, what is the constant 2000, and how was this value determined?

4.2.4 Concentration algorithms

The concentration algorithms are difficult to understand, partly because of the nonphysical distinction between the mixing layer and the boundary layer (see Section 4.3.1 below). For the case of the input and cloud center heights within the mixing layer, 50% of the cloud is "stripped off" at the ground (i.e., 50% of the mass is deposited on the ground), and the other 50% is reflected back into the mixing layer. This is accomplished in the model by multiplying the image source contribution term by a "reflection coefficient" which has been assigned a value of 0.5, because this "has been found to best agree with experimental data". Similarly, when the cloud penetrates the elevated inversion above the mixing layer, 50% of the material is lost and 50% is reflected back into the mixing layer. This was accomplished by multiplying the real source term by a reflection coefficient, again set to 0.5 "to best agree with experimental data". In this connection, we note that:

(a) it is difficult to justify scientifically such modification of the concentration algorithms by arbitrarily assigning large values for the reflection coefficient, just to obtain agreement with the diffusion data.

(b) it is not clear if and how the material deposited or lost (as discussed above) is accounted for by the model, and there was no mention of this in the report. Considering the toxic nature of the contaminants being modeled, the location and amount of such deposition will have important consequences.

(c) it is mystifying that, in spite of the losses discussed above, the report says (p. 3-110) that "each cloud retains its original mass at all times," because "no depletion terms are used". It is well-known that reflection coefficients are used to represent depletion of material through deposition (at the ground) and ventilation (at the top).

(d) only one reflection each from the ground and the mixing layer are used in Eqs. (73) and (74); this does not guarantee a uniform concentration distribution at longer downwind distances from the release location in the mixing layer under convective and neutral conditions. Multiple eddy reflections are normally required to ensure this uniformity.

(e) when the cloud is above the boundary layer under stable conditions, it usually does not contribute to the concentration in the mixing layer. Instead, in Eq. (72) on p. 3-109, this concentration was effectively set to 0.25 times the real source contribution. Once a portion of a cloud penetrates through the mixing layer, that portion will be transported and diffused differently than the rest of the cloud. It should thus be treated as a separate puff. The partial-reflection concept becomes even less tenable when a cloud penetrates both the mixing and boundary layers.

(f) the boundary layer height of 800 m assumed for stable conditions (p. 3-108) is too high. Also, it is not physical to assume that the mixing layer height is greater than the boundary layer height.

The treatment of the surface and mixing layers used in other puff models such as INPUFF (Petersen and Lavdas, 1986) is more realistic than that used in VLSTRACK. Typically, full reflections are used both at the surface and the top of the mixing layer. To avoid the use of multiple eddy reflections, the puff is assumed to become well mixed in the vertical (i.e., uniform concentration with height) when the vertical standard deviation σ_z of the puff reaches a size comparable to the mixing-layer depth.

4.2.5 Deposition

Despite its importance, there is no clear discussion of deposition methodology in the VLSTRACK model. The deposition properties of gases and aerosols differ from those of liquid droplets and particles. The deposition mechanisms at the earth's surface include gravitational settling, turbulent and Brownian diffusion, chemical adsorption, inertial impaction, and thermal and electrical effects. Gases and aerosols are primarily deposited by turbulent transfer, while large particles are deposited by gravitational fall.

It appears (see p. 3-191) that clouds consisting of gases and droplets/particles of 10 μm or less in diameter are not allowed to deposit in VLSTRACK. The minimum height above the ground surface for these clouds is arbitrarily set to 0.5 m. Only clouds having larger droplets or particles with non-zero terminal velocities are assumed to impact the ground. Their impact locations and impact times are computed to estimate the deposition and dosage; how this is done is not clear. Thus, the surface deposition of gases and aerosols by the various mechanisms given above does not seem to be taken into account. Also, on p. 3-215, the last paragraph begins with, "For deposition, the hazard at a grid point is the sum of area concentrations from all of the cloud records". We do not understand what this sentence means.

In models such as VLSTRACK, one is primarily concerned with estimating the concentrations and dosage of the hazardous contaminants in the mixing layer, and their ground deposition patterns. Any material that is "stripped off" from the cloud in the mixing layer affects these estimates, and should be clearly accounted for. Although Bauer (1993) mentions material being "stripped off" from vapor clouds, the VLSTRACK model apparently ignores the deposition of vapor clouds to the ground surface. This is likely to be a reasonable assumption at short downwind distances from the source, but will result in increasing errors farther downwind. Most puff models use specified deposition velocities to represent the effects of dry deposition of vapor or gas clouds.

The parameterized dry removal of gaseous material or small particles from a plume or puff is generally accomplished by appealing to the so-called dry deposition velocity (e.g., Chamberlain, 1953) which relates the turbulent flux of airborne material delivered to a surface to the difference between the local airborne concentration and the concentration at the particular receptor surface. The proportionality factor is called the dry deposition velocity because it has units of length per unit time. Formally, the flux = $\mathbf{V_d}$ ($\mathbf{C_A}$ - $\mathbf{C_S}$), where $\mathbf{V_d}$ is the effective overall mass transfer or dry deposition velocity, and $\mathbf{C_A}$ and $\mathbf{C_S}$ are the concentrations in the air and at the receptor surface, respectively. The deposition velocity is not a constant, and depends on a host of meteorological, chemical, and receptor-specific conditions; it is generally difficult to evaluate without experimental effort. The usage of $\mathbf{V_d}$ within a dispersion model depends on the model type; see Hosker (1986) and Rao (1982) for a survey of schemes. Puff models generally remove a portion of the puff mass at each time step in the calculation, as the puff moves along a trajectory determined by the mean wind; see, e.g., the INPUFF 2.0 model (Petersen and Lavdas, 1986) which adapts Rao's (1982) analytical plume concentration algorithms for dry deposition and gravitational settling of pollutants.

The most difficult aspect of applying the dry deposition velocity approach to gaseous plume or puff depletion is the specification of reasonable values for $\mathbf{V_d}$. An organized approach has been developed over the years from Thom's (1975) suggestion of a resistance analogy, where the mass transfer process is considered to occur through a series of layers, within each of which certain transfer processes dominate, while others are less important or negligible. Mass transfer is then held to be limited by a resistance characteristic of each layer, and the overall deposition velocity is calculated as the reciprocal of the sum of these various series resistances. The formalism has been summarized by Slinn et al.,

(1978). The advantages of this approach are that some of the layer resistances can be estimated in a straightforward way, and that rate-limiting steps and important physical processes can often be identified. The transfer resistance right at the receptor surface is usually the most poorly known term. Some of the literature on small particle and gas dry transfer at surfaces is briefly described in Hosker (1986).

Extensive research on dry deposition was performed under the National Acid Precipitation Assessment Program throughout the 1980's; see NAPAP (1991) for a summary. One product was the NOAA dry deposition inferential method, in which meteorological data, pollutant physical and chemical properties, and observations of local surface conditions and vegetation are used to infer values for $\mathbf{V_d}$; these estimates in turn are combined with measured concentrations to calculate the fluxes of selected pollutants to particular surfaces at given times. The procedure is described by Hicks et al. (1985, 1987). Most of the recent improvements to this method have centered on the transfer of particular chemical species to particular vegetative surfaces, and involve a multi-disciplinary approach that incorporates vegetative response to local conditions and stresses in its estimates of uptake through stomatal and other pathways. It is possible that a similar inferential approach could be developed for use in VLSTRACK, that would allow quantitative estimates of the turbulent transfer of gases and small particles to various receptors. This in turn would provide more realistic depletion of a puff as it moves along its trajectory.

For liquid droplet clouds, VLSTRACK does not allow deposition to occur before the cloud center comes very close to the ground, at which time the cloud "impacts" the ground. This is an oversimplification of liquid droplet deposition. By assuming that all the droplet deposition occurs instantaneously when the cloud center reaches the ground, VLSTRACK will most likely produce a deposition "footprint" that is smaller and more concentrated than the actual footprint. Total dosages from an attack will thus be overestimated at the location where the cloud center reaches the ground, but will be underestimated at points upwind and downwind of this point (although the dosages at downwind locations may be affected by secondary vapor clouds). This error is likely to be most significant for small droplets, since their gravitational settling velocity is small.

The "washout" (rain falls through a puff and carries material to the surface) or "rainout" (pollutant materials are incorporated in a precipitating cloud and carried down in the droplets) are very efficient at removing pollutants from the atmosphere. Techniques to deal with wet removal of aerosols and soluble gases are summarized in Hosker (1986); the required washout ratios and wet deposition velocities can in some cases be estimated or inferred from measurements for specific species, although measurements are preferred. The results depend on rainfall rate, droplet size distributions, and other terms characteristic of the precipitation event. While there are still many uncertainties in this procedure, it does provide a means to at least estimate the puff deletion and deposition that will occur when a puff encounters precipitation; this is presently not accounted for in VLSTRACK. Though chemical and biological releases are unlikely to occur during precipitation conditions, the complexity of the sources and the large, varied spatial domain likely to be involved cannot always guarantee dry conditions.

4.3 Task 2. Empirical Parameters

4.3.1 Mixing layer

The distinction made between the mixing layer and the boundary layer (pp. 3-76 and 3-108) is not physical. The mixing layer height in the VLSTRACK model is defined in the report as "the height at which the vertical temperature profile changes direction". This is technically incorrect since

temperature is a scalar quantity and does not have a direction; we assume that the author is referring to a change in slope of the profile.

The mixing-layer depth is usually defined (e.g., Panofsky and Dutton, 1984) as the depth of the turbulent region near the ground. More importantly, and strictly speaking, the mixing layer is actually the height to which the turbulent mixing of pollutants in the atmosphere extends. In the VLSTRACK model, the atmospheric boundary layer height is also described as "the height where the atmospheric dispersion behavior changes"; this is also inappropriate. Several definitions exist in literature for the atmospheric boundary layer. During convective conditions, the boundary-layer depth is often used as a synonym for the mixing-layer depth. Alternatively, the boundary layer is defined as either the height of the lowest temperature inversion, or as that part of the atmosphere where the influence of surface friction and heating or cooling is felt (Fleagle and Businger, 1980; Stull, 1988). During the nighttime stable conditions, the atmospheric boundary layer is typically an order of magnitude smaller than its maximum daytime value over land. At night, the mixing layer height is also smaller than the height to which the surface radiation inversion extends, i.e., the boundary layer height. Above the sea, the diurnal variation of the mixing layer height is much smaller than over land.

4.3.2 Parameterization of turbulent velocities within the boundary layer

We first note an error in nomenclature that is used consistently throughout the model document: turbulence intensity is not the same as the standard deviation of the wind velocity fluctuations; rather, it is the ratio of the latter quantity to the mean wind speed, and is therefore a dimensionless quantity (see, for example, Hanna et al., 1982; Stull, 1988).

The VLSTRACK model uses parameterizations for the horizontal (σ_u) and vertical (σ_w) velocity standard deviations that are not in agreement with field measurements. For neutral and stable conditions, the model assumes that $\sigma_u = \mathbf{b}\mathbf{u}_*$ and $\sigma_w = \mathbf{B}\mathbf{u}_*$, where **b** and **B** are empirical constants. Based on turbulence field measurements (e.g., Panofsky and Dutton, 1984), typical values of **b** are 1.9 to 2.4, whereas typical values of **B** are about 1.25. There is no evidence in these field measurements that **b** and **B** should decrease as the stability increases from class D to class G (the decreased turbulence in stable conditions results from a smaller $\mathbf{u}_*$). VLSTRACK uses values of **b** and **B** that are too small, with **b** ranging between 1.3 and 1.7, and **B** ranging between 0.50 and 0.65. The values of **B** used in Eqs. (38) to (41) on p. 3-77 are too low for all (A to G) stability conditions compared to the widely-used empirical values in the literature. For unstable conditions, the **B** value used in VLSTRACK is roughly one-half to one-sixth of the correct value.

According to Bauer and Wolski (1993), **b** and **B** were adjusted so that the VLSTRACK model would closely fit a set of biological dosage and deposition measurements. The factors **b** and **B** represent a relationship between the velocity standard deviations and the friction velocity $\mathbf{u}_*$, and not a relation between the velocity standard deviations and the dosage and deposition of a contaminant cloud. Many other factors in the model affect the dosage and deposition, so it is simply not appropriate to adjust **b** and **B** to mask problems that may lie elsewhere in the model. The only justifiable method for adjusting these parameters is to compare measurements of σ_u and σ_w with simultaneous measurements of $\mathbf{u}_*$. Otherwise, one should not arbitrarily change physical properties of the atmospheric turbulence to fit diffusion data.

For unstable conditions, VLSTRACK uses a formula of the form $\sigma_w = \mathbf{B}'\mathbf{w}_*$, where $\mathbf{w}_*=(-Z_b\Gamma/k)^{1/3}$ is the convective velocity scale and $\mathbf{B}'$ is a constant, which is varied as a function of stability. In the formula for $\mathbf{w}_*$, $\mathbf{Z}_b$ is the mixing layer height, and k is the von Karman constant. This formula is not a good representation of the vertical velocity fluctuations, because it assumes that the ratio $\sigma_w/\mathbf{w}_*$ is

constant with height in the boundary layer. A number of studies (e.g., Panofsky and Dutton, 1984) indicate that this ratio is in fact a function of the height **z** above the ground. Since height-dependent parameterizations of σ_w are available (Panofsky and Dutton, 1984; Stull, 1988), it is not clear why σ_w/w_* was assumed to be constant in VLSTRACK. Available data (e.g., Holtslag and Moeng, 1984) show that σ_w in the Convective Boundary Layer depends also on the normalized height z/z_b and the friction velocity u_*.

4.3.3 Parameterizations above the boundary layer

When a puff is above the boundary layer, the model assumes that $\sigma_u = 1.7u_*$ and $\sigma_w = 0.65u_*$ (Eqs. (42) and (43) on p. 3-79). The reasoning behind these formulas is the belief that the troposphere above the boundary layer has neutral stability. However, the free troposphere is in fact stably stratified on the average. The formulas used for σ_u and σ_w at high altitudes have no physical basis whatever, and will generally produce values of σ_u and σ_w that are far too large. There is only patchy or sporadic turbulence above the mixing layer, and this does not scale with the surface friction velocity u_*. The free-air turbulence is provided by occasional convective penetrations from the mixing layer, gravity wave breakdowns, shear-generated turbulence, cumulus clouds, etc. Pollutants released above the mixing layer at night can often travel and meander across large distances, causing horizontal spreading with little or no vertical diffusion. Since the friction velocity u_* is a measure of the momentum transfer that takes place right at the earth's surface, it is only useful for describing the level of turbulence within a relatively shallow layer (roughly 100 m deep during the day) near the surface. There is no physical relation at all between the turbulence that exists above the boundary layer and u_*. A simple and more defensible approach for estimating σ_u and σ_w above the atmospheric boundary layer would be to set them to small constant values. For example, Petersen and Lavdas (1986) use $\sigma_w = 0.01$ m/s.

4.3.4 Dispersion parameters

The incremental changes in cloud dispersion parameters, Eqs. (143) and (144) on p. 3-179, do not come from the work of G. I. Taylor, so this reference should be removed. Instead, reference should be made to NUSSE3 (Saucier, 1987), where they are derived as approximations to Gifford's (1982) expression. The droplet diffusion factor should be defined. The term in the first set of parentheses is not clear, and should be explained.

The dispersion parameters are generally not used in the incremental formulations of the type given in Eqs. (143) and (144). When using these equations with temporally varying meteorological conditions, the concept of virtual travel distance (Ludwig, 1982) should be used to provide a realistic representation of plume growth after a change in atmospheric stability; otherwise, large errors might occur. It should be clearly discussed if and how this is being implemented in VLSTRACK 2.0.

In Eqs. (115) and (116) on p. 3-156, the total alongwind and crosswind dispersion coefficients are each calculated as the sum of the sigma obtained from Eqs. (143) or (144), and a wind shearing factor with an empirically tuned coefficient of 0.3. No reference or justification was given for this formulation. In this connection, it should be noted that while variances are additive, standard deviations (as used in VLSTRACK) are not. Hence, Eqs. (115) and (116) will give incorrect values for dispersion coefficients.

4.4 Task 3. Model Validation Studies

4.4.1 Introduction

Atmospheric dispersion models are mathematical approximations to "real-world" atmospheric transport and diffusion, and as such there are uncertainties associated with their predictions. There is no "one" general atmospheric dispersion model; each dispersion model is designed to address specific issues or a specific application. The current state of the science limits models to temporal and spatial scales defined in the design of the dispersion model. Due to necessarily imperfect simulations of complex atmospheric processes, a model's prediction without an estimate of the associated uncertainty is of only limited value to a decision-maker. This measure of uncertainty can be either qualitative or quantitative. Qualitative measures assign a description such as "realistic" or "conservative" to the model output. Quantitative measures try to bound a model's predictions with a statement on the level of confidence, such as "accuracy within a factor-of-three", which scales the model's output. The quantitative measure of "confidence" is particularly relevant to the emergency management scenario, in which individual episodic events are often modeled with standard ensemble-average techniques such as the Gaussian plume dispersion model.

The principal factors affecting the reliability and validity of atmospheric dispersion simulations are: (1) specification of the problem, (2) formulation and verification of the model, (3) evaluation of parameter values, and (4) validation of model results. Specification of the problem requires a description of the scenarios to be simulated by the dispersion model. The scenario specification should include the intended use of the model (emergency planning, emergency response, long term assessments), temporal and spatial resolution required in the output (near-field, regional, short term, etc.), the various processes which need to be resolved (heavy gas, chemical reactions, evaporation, buoyancy, etc.), and the character of the releases (continuous, episodic, instantaneous, etc.).

Model verification is the process of assessing the scientific accuracy of a model. The model formulation should be based on sound physics, and give good predictions for the "right" reasons. This requires a scientific review of model formulations, as well as a review of the code to verify that the model is properly implemented. Emergency management models have an implied need for verification, to give confidence in model predictions beyond the range of available data, and for modeling new situations with different dispersion conditions.

Model parameters refer to the constants and independent variables of the equations used in the simulation. Model parameters include varying degrees of uncertainty. The degree to which model parameters influence the level of model accuracy is typically addressed through a model sensitivity study. For practical purposes, it is useful to identify through model sensitivity analyses those components with the greatest potential for contributing to the uncertainty in the predicted results, and those which have negligible influence. Model validation trials then only need to be designed to test important components. Errors in documenting the model's processes, including input parameters and calculation techniques, directly translate to model error and create doubts in the reliability of the model.

It must be recognized that regardless of how a model has been designed and verified, some level of uncertainty will be associated with its predictions. The process of quantifying the level of uncertainty involves testing the model against independent data sets. Model evaluation consists of assessing the performance of a model by comparing its predictions to the concentrations measured in a field study. Over the past decade or so, sophisticated statistical and other techniques have been advanced for the evaluation of air dispersion models from scientific and operational viewpoints. Adequate

meteorological and concentration data are required for model evaluation. It is important that residuals, expressed as the difference between predicted end observed concentrations, not show any trend when plotted against independent variables such as wind speed, stability, or downwind distance. A residual analysis (Hanna, 1993; Rao and Hosker, 1993) is performed to identify and correct the observed biases, and the modified model is re-evaluated in an interactive process.

Model validation involves multiple evaluations for testing various model components to ensure a known level of accuracy. The level of accuracy that is acceptable is subjective, and it will vary from case-to-case. The process of model validation requires a suitable database which is collected from different field and laboratory experiments to address a specific issue. Quantitative uncertainty associated with these data must be indicated. Care must also be taken to ensure that the validation data are independent from those used to develop the model. Temporal and spatial scales of the data must be relevant to the model. Specific guidance on model calibration is provided by the U.S. EPA (1986):

> "Calibration of short term models is not common practice and is subject to much greater error and misunderstanding. ... This approach is severely limited by uncertainties in both source and meteorological data and therefore it is difficult to precisely estimate the concentration at an exact location for a specific increment of time. Such uncertainties make calibration of short term models .... unacceptable."

After quantifying uncertainties in the test data, one can undertake a direct comparison with model predictions. The validation process is carried out as necessary to encompass the range of conditions over which the model may be applied (stabilities, complex terrain, land use types, wind regimes, etc.). Model validation involves a substantial investment in time and resources, and requires careful planning and consultation between the modelers, managers, and primary model users.

Irrespective of whether the model has been subjected to the validation process or not, some uncertainty will be associated with the dispersion model's calculations for all applications. Consequently, it is desirable to perform an uncertainty analysis (Rao and Hosker, 1993; Hanna, 1993) which provides 95% confidence intervals on the model predictions. This facilitates improved decision-making based on model uncertainty and risk assessment.

4.4.2 Documentation

Material initially provided by NSWC related to the validation of the VLSTRACK model included:

1) Wolski, M.G.,"Final Validation of VLSTRACK 1.2", Draft Report, 1993.
2) "Planned Validation and Verification Effort for the CBD-IMPACT Software," Draft Statement, 3 May 1994, no author indicated.

After an initial review of "Final Validation of VLSTRACK 1.2" (hereafter referred to as the Final Report), several reports were identified from its References for inclusion in the review process. The requested material included data reports for Dugway Proving Ground Trials of chemical releases (references 16, 17, 20 of Final Report), and material from the biological simulations (references 21, 22, 23, 24 of Final Report). In addition, the "Initial Validation of VLSTRACK 1.2" (hereafter referred to as the Initial Report) was requested. Of the seven reports requested, the following four were provided:

1) Trial Report DPGTR 213 F-3, Dugway Proving Ground
2) Trial Report DPGTR 213 A-3, Dugway Proving Ground

3) Trial Report DPGTR 213 F-5, Dugway Proving Ground
4) NSWCDD/TR-92/647, by Bauer and Wolski (1993).

Also, two biological trial reports,

1) Shoreline Diffusion Program, Oceanside, California, Vol. 1, Technical Report, by T. B. Smith and B. L. Nieman, 1969, Final Report, MR169 FR-860, Meteorological Research, Inc. 230 pp.
2) Aerosol Diffusion over Woodlot Complexes, Vol. 1, Summary and Results, by G. R. Hilst, 1969, TRC 343, Travlers Research Corporation, 165 pp.

were obtained from Library Services of Dugway Proving Ground.

The output from the VLSTRACK 1.2 model was also requested. Model output was provided for chemical trial runs 21 and 22 (DPGTR 213 A-3 and DPGTR 213 F-5), referenced in the Final Report. VLSTRACK 1.2 output included vapor dosage maps and tabular model input summaries. Similar data output was provided for biological trials 1, 27, 98, and 99 referred to in the Final Report.

4.4.3 Validation review

4.4.3.1 Review of approach

All of the materials indicated above relate to VLSTRACK version 1.2. While the Statement of Work specifically names "Draft Report, February 1994, Initial Validation of VLSTRACK versions 1.5-2.0," this report was not provided in time for this review. Both the Initial and Final Reports followed similar methods for validation of VLSTRACK. The following discussion will first address problems identified within the Final Report, and then discuss the overall issue related to model validation.

4.4.3.2 Final report

The Final Report provides a summary of the comparison effort in which experimental data from 131 biological trials and 43 chemical trials were compared to predictions from VLSTRACK 1.2. A logarithmic ratio (defined as the "difference factor") of the predicted to experimental values was used as the comparison statistic. The stated purpose of the Final Report was "to provide a better estimate of the accuracy of VLSTRACK 1.2 over a wide range of chemical and biological agents, attack configurations, and meteorological conditions". In general, in both the Initial and Final Reports, the presentation of the validation effort is rather cursory. Much of the evaluation material, such as descriptions of the comparison data, parameter sensitivity studies, etc. was not provided. The reports rely heavily on the quantity of the comparison data, but lack a discussion of their data quality, leading to many unresolved issues. Specific problems within the Final Report include:

p. 4: The required VLSTRACK 1.2 input parameters that were not included in all of the biological trial reports are described as being developed from an "educated guess". There is no discussion of this process or how the initial lateral and vertical sigmas and stability were defined in the absence of data.

p. 12: While there is a fairly detailed discussion of the sample grid used for the Dugway Proving Ground trials, there is no description of the individual trials. For example, it is only by referring to DPGTR 213 A-3 that it became clear that the source for this trial was actually a ground burst of a 750 pound toxic bomb dropped from 12,000 feet with an airspeed of approximately 220 knots. There is no discussion of associated impact spills, liquid/vapor splits, or initial release sigmas for what is obviously a unique release event.

p. 12: The report indicates that, during the chemical trials, the grid points at Dugway were equipped with both liquid and vapor sampling devices within the inner grid; however, none of the supplied Dugway reports contained liquid spill data. These spill data are particularly important in estimating the strength of the secondary source developed from the evaporating liquid. As shown later, VLSTRACK 1.2 overestimates contributions from the secondary sources.

p. 14: In Table 3 (Final Report), a relationship is mentioned between the Pasquill stability category and the vertical temperature gradient; however no specific reference is provided.

p. 14: Table 4 (Final Report) is a table for converting 2 min and 10 min average data into an algorithm for estimating both wind speed and wind direction for averaging times between 0 and 10 min. No reference is provided to document the technique. Table 1 in the present report gives an evaluation of Table 4 (Final Report); it can be seen that the latter allows the travel direction of the release to rotate 55 degrees beyond the 10-min value, while the material undergoes a constant acceleration beyond this value.

p. 16: Figure 2 (Final Report) provides a list of common parameters which VLSTRACK 1.2 requires. This Figure defines "fill weight" as E times Q, where E is dissemination efficiency and Q is total fill weight. However, in the same figure, the initial lateral and vertical cloud sigmas (which have units of length) are also defined as functions of Q. No explanation or reference for this extrapolation is provided; the form of the sigma relationship appears to be an empirical fit to observed data, but this is only our guess.

p. 16: The initial horizontal and vertical sigmas are not provided in Table 5 or in the Appendices of the Final Report.

p. 19: Equation 2 (Final Report) is defined as the difference factor (F). In the Initial Report, the difference factor is cited as a resolution to a perceived problem with small concentrations. While this is true, the difference in logarithms can be associated with the fact that concentrations tend to be log-normally distributed, which means that the logarithm of the concentration is normally distributed or "symmetrical" as stated in the report.

p. 21: Reference is made to scatter on the S-shaped curves but not explained. There is a mention of "considerable scatter" in Figure 8 (Final Report); but this figure looks similar to Figure 5 (Final Report). Other figures are described as "scattered", "accurate as overall", and "within standard deviations", but without companion statistics or documentation.

4.4.4 Validation protocol

The following discussion provides general comments on the VLSTRACK validation effort. The issues addressed include comparison data, statistical validation technique, and parameter sensitivity.

4.4.4.1 Comparison data

There is no explanation or justification within either the Initial or Final Reports for the selection of the comparison data. Emphasis is placed on the number of data points, but not on the quality of the data. Neither report provides more than a cursory overview of both the chemical and biological data. Model validation requires data suitable for the issue under study (e.g., dispersion models developed to address complex terrain issues should use data from complex terrain; heavy/dense gas models should use dispersion data for dense gases, etc.). In many cases, the model cannot be tested directly with field

trials; in this situation the various components of the model such as transport, dispersion, and evaporation must be independently evaluated. In any case, the quality of the comparison data must be known in order to justify assigning a score to the model's predictive ability.

4.4.4.2 Biological comparison data

Paragraph 1, Summary, in the report "Shoreline Diffusion Program, Oceanside, California" (hereafter referred to as the Shoreline Study) by Smith and Nieman (1969) describes the intent of the field measurement program:

> "...to investigate the diffusion characteristics in the vicinity of a shoreline under conditions of onshore flow. Specific objectives of the program were to develop an understanding of the physical process occurring in the shoreline environment..."

Results of the Shoreline Study indicated:

- "A stable layer exists at a height of 150-400 m above the ocean surface which restricts upward diffusion growth..."
- "... growth in the cloud widths... which accelerated with distance ... this effect was dependent on location along the coastline."
- "Vertical cloud growth appeared to be restricted at large downwind distances... "
- "... resulting effect was a less rapid decrease in dosage at large downwind distances than would be expected from normal model calculations."
- "... line source showed evidence of rapid vertical diffusion between the source and the coastline."

The results of the Shoreline Study highlight the complexity of flow and dispersion within a shoreline environment. Lyons (1975) presents (see Figure 2) a conceptual summary of observations of a shoreline fumigating plume. An internal boundary layer (IBL) develops downwind of the land/water interface as a consequence of the step change in surface roughness and thermal conditions. Dispersion of material released in this environment will be a function of release location relative to the IBL; as indicated in Figure 2, material released upwind of the IBL will rapidly disperse within the IBL. There are no provisions within VLSTRACK 1.2 to model the land/water interface and developing IBL. Without a mechanism to account for the various turbulence and stability regimes associated with the IBL and fumigation of material into the IBL, any correlation between model and field data is fortuitous.

Paragraph 1, Summary, "Aerosol Diffusion Over Woodlot Complexes," by Hilst (1969) describes the project as:

> "... designed specifically to determine the manner in which complex vegetation patterns and terrain usage affect the transport and diffusion of an aerosol cloud..."

Specific findings of the study indicated:

- "Terrain and vegetation effects were more significant with increasing thermal stability... The effect of the trees ... might well be expressed as that of a coarse filter... "

- "When a tracer cloud entered a wooded area either vertically or horizontally greater than the dimension of the cloud, the reduced advection velocity and enhanced turbulence caused the cloud to expand rapidly until it attained a size proportional to the obstacle."

- "A deficit of material was found immediately in the lee of a wooded area... "

- "Releases made in the rolling terrain northeast of the grid center did not follow the observed wind pattern, but instead reflected the effect of the local topography."

- "Under stable conditions, the tracer cloud followed valleys and flowed around hills. These motions were not reflected in the wind field, because the towers near the area were located on high land areas."

Techniques to incorporate these processes are not provided within VLSTRACK 1.2. Without mechanisms to account for the increased turbulence or terrain steering of the released material, comparisons between the model and field data are suspect. Both the Shoreline Study and Wooded Complex reports provide transport and dispersion modeling techniques which address the complexities identified in each study.

In general, the biological trials, based on information inferred from the Final Report Appendices, demonstrate a large degree of uncertainty in the observations. Data reported in Appendix C, p. C-5 to C-8, indicate particular problems with the observations used in the validation. For example, Table 2 (in this report) lists the data for Trials 2.28 and 2.29. The data for Trial 2.29 appear inconsistent at 1.5 km when compared to Trial 2.28 at the same downwind distance. Trial 2.28 indicates a large underprediction while trial 2.29 shows a large overprediction, yet the Final Report shows identical release conditions and identical release quantities for both Trials. Trials 2.26 and 2.27, p. C-5 (Final Report), have similar release conditions, but at 1.5 km downwind the Trial data differ by 70.60 (Trial 2.26) to 694.40 (Trial 2.27) units. A scan through Appendix C (Final Report) shows roughly 40% of the reported biological trials have two or fewer data points. The lack of observation points is unfortunate, since as indicated in Trial 2.36 and Trial 2.37 (Shoreline Trials 21G and 21Y), the ratio of predicted to observed values is maintained, indicating that VLSTRACK 1.2 perhaps modeled the decay of concentration with distance correctly for this stability, wind speed, and release scenario, but underpredicted the concentration level by a factor of five.

The above agreement may however be fortuitous and simply attributable to the conditions of the study. The Shoreline Report indicates that the trials with one or two observations were conducted to determine plume horizontal and vertical dispersion parameters (σ_y and σ_z) for both surface and elevated releases. This series of trials would probably be better utilized to validate VLSTRACK 1.2 dispersion parameterizations instead of the accuracy of concentration predictions.

It should be noted here that both the Woodlot and the Shoreline dispersion experiments used fluorescent particles (FP) as tracer. The FP tracer is inherently non-conservative; significant material losses can occur by dry deposition of the FP on to vegetation between the source and the receptor point (Hosker *et al.*, 1993). There is also a tendency for FP to lose their fluorescence (and hence become non-detectable) after exposure to strong sunlight and high humidity. There is uncertainty associated with the measured concentrations because of these shortcomings.

4.4.4.3 Chemical comparison data

Very little site information or project data are provided in the various Dugway Proving Ground Trial Reports. Of the six Trial Reports provided, five showed wind direction at release from the north-northwest and A stability. The dependence of the VLSTRACK 1.2 validation of chemical releases on the Dugway data biases the validation to those conditions associated with Dugway Proving Ground. For the near-field trials at Dugway, the site appears to be simple with relatively flat terrain, uniform surface conditions and steady winds. These conditions are appropriate for Gaussian models. While VLSTRACK 1.2 may test well against Dugway Proving Ground data, it has not been demonstrated that the conclusions of this chemical release validation can be applied to other surface and atmospheric conditions.

The individual reports indicate unique source conditions. The impact detonation of toxic bombs or rockets generates complex source conditions and subsequent downwind deposition and concentration patterns. The Final Report provides a general explanation of the procedure used to determine ground-level centerline concentrations for each trial. However, there is no discussion on resolving the complex surface dose patterns (refer to DPGTR 242 D-3) in which the maximum surface dose does not follow the prevailing wind direction as indicated in Figure 5, DPGTR 242 D-3.

The chemical data base reported in Appendix C (Final Report) has several errors which are propagated through the validation analyses. Data for chemical Trial 21 are listed in Table 3 (this report). The DPGTR 213 A-3 report shows the dose value of 500 mg-min/m^3 listed at time 0-10 min (at x = 0 and at x = 7.6 m) to be extrapolated values and significantly below the 0-2 min values at the same downwind distances. According to DPGTR 213 A-3, the dose value at x = 15.2 m for 0-10 min is >500, which indicates this to be an estimated value. The value listed in Appendix C for this Trial at 15.2 m is 964 which, according to DPGTR 213 A-3, is associated with the measurement at a grid point adjacent to the estimated value. All three points are therefore suspect, and should be eliminated from the model validation.

The above questionable data from both chemical and biological data were identified after only a cursory review of the comparison data. The comparison data base must have a known uncertainty. Neither the Initial nor the Final report provides any indication of the reliability of the comparison data. With regard to both the biological and chemical trials, no data are provided in either report related to the initial puff/plume horizontal or vertical cloud sigmas. As indicated above, DPGTR Trial 213 A-3 reports a source consisting of a ground burst of a 750 pound toxic bomb, released at 12,000 feet at 220 knots; Project Windsoc reports a 100 km line source; and Project Shoreline uses both point and line sources. While these release conditions may be routine in CBW events, these conditions are far different from the standard stack emissions for which most Gaussian plume/puff dispersion models are applicable. Downwind concentration/dose levels will be particularly sensitive to the initial dilution at the source; the sensitivity of VLSTRACK to the initial source condition is not addressed in either the Initial or Final Validation Reports.

The data as presented in the Final Report are given as dosage, which implies some averaging or exposure time. Data reported for the Dugway Trials are reported for 0-2, 0-10, and 0-30 min intervals. Using dosage as the predicted value requires a model to not only predict downwind concentrations correctly but also the exposure time. In the Dugway Trials, the sampler simply accumulated concentrations over a prescribed time period; data are reported with units of mg-min/m^3, and one can extract only an average concentration for the reporting interval from the data.

The VLSTRACK 1.2 Software Design Document states that the hazard at each grid point is the sum of concentrations multiplied by the time step. Using dosage data presents the classical problem of equivalent dose, where short-term exposure to high concentrations gives the same dose as exposure to lower concentrations for a longer time. While the computed dose may be the same, the physical consequence is not. For example, National Institute for Occupational Safety and Health lists 30 ppm as the IDLH (Immediately Dangerous to Life and Health) concentration and 1 ppm as the STEL (Short Term Exposure Limit) concentration for chlorine. While a one minute exposure to 30 ppm concentration gives an equivalent dose as 30 min exposure to 1 ppm concentration, one would not expect the same physical consequences. The Dugway Proving Ground observations do not permit the validation process to evaluate the model's ability to correctly predict both concentrations and exposure times. However, an estimate of the model's ability to predict average concentration levels can be extracted from the 0-2 min and 0-10 min dose values. An average concentration over the time interval can be estimated from the difference between the two dose values. Table 4 (this report) tabulates Trial and model observations for chemical Trials 19 and 20. For Trial 19, the model score changes slightly from 1.58 to 2.14; for Trial 20, the ratio increases from 41.2 to 101.6. Thus, while the incremental change can be used to estimate an average concentration, it is impossible to determine how the concentration varied over the exposure period. Neither short-term concentrations nor exposure times can be extracted from the data; while the model predicts both quantities, neither can be individually evaluated from the Dugway Trials. Inclusion of dose as the primary predictor prohibits validating the model for its concentration prediction capability.

This has particular relevance to an emergency response scenario in which the instantaneous or short-term concentration is more significant than a long term exposure. This is also relevant to the use of ensemble-average dispersion techniques, in particular the Gaussian dispersion model, to predict instantaneous concentrations or the results of an individual release event. The Gaussian model is based on a prediction of the mean concentration from an ensemble of instantaneous puffs or plumes of material (Slade, 1968). As illustrated in Figure 3, individual puffs have an infinite number of potential trajectories from the source; also indicated are the instantaneous concentrations. As material is swept downwind, a receptor at a fixed location may see extreme variations in concentration levels. As the averaging time increases, the concentration distribution is assumed to approach a Gaussian distribution, in which the horizontal and vertical sigmas define the mean concentration. The influence of averaging time is illustrated in Table 5 which gives the exponent **p** in the power-law expression (IAEA, 1989) used to convert the concentration, C_1 with a given sampling time $\mathbf{t_1}$ to concentration C_2 with different sampling time $\mathbf{t_2}$:

$$\frac{C_2}{C_1} = \left(\frac{t_2}{t_1}\right)^{-p}$$

4.4.5 Statistical Technique

Both the Initial and Final Reports use a difference factor **F** defined as

$$\mathbf{F = 10 \log (d_p/d_e)}$$

where $\mathbf{d_p}$ is predicted value, and $\mathbf{d_e}$ is experimental value. While this statistic does have symmetry, there are difficulties with using it as the single validation statistic.

Particular problems exist when the predicted concentration profile intersects the observed profile. A quick review of the data presented in Appendix C of the Final Report indicates numerous examples where VLSTRACK underpredicts near the source and overpredicts with distance, or conversely (Appendix C, Trials 4.2, 4.9, 4.10, 4.14, 21, 22, 23, 24.1, 24.2, 24.3, etc.). One problem with the single mean statistic is illustrated in the summary of biological trials (p. B-5, Final Report). While the mean difference factor **F** is -0.10, the mean of the absolute difference is 3.83. The former gives an almost perfect ratio $\mathbf{d_p/d_e}$ of 0.98, while the latter gives 2.4. Using only the difference factor also skews the validation results by sample size. Those trials in which there were numerous points weight the distribution towards these trials. Figure 4 illustrates the problem associated with sample size and relative position of sample points. In this example, the two sample sets show mean values either slightly underpredicting or overpredicting based on sample size.

While there is no specific regulatory protocol on model validation, guidelines have been provided by the U.S. Environmental Protection Agency and the American Meteorological Society. A good review of current practices is provided by Weil *et al.* (1992). Recent compliance interest at the U.S. DOE's Rocky Flats Nuclear facility has produced a cooperative program between the DOE, EPA, and the State of Colorado to develop an evaluation methodology and protocol for evaluating atmospheric dispersion models used to simulate radioactive releases, both emergency and historical, from the Rocky Flats Facility. Ciolek (1994) provides a summary of the statistical measures used in this protocol. These include true difference, absolute difference, absolute fractional bias, variance of true difference, variance of fractional bias, mean square error, and correlation coefficient. Tangirala *et al.* (1992) and Hanna (1993) described simplified model evaluation methods based on fewer statistical measures, such as fractional bias, normalized mean square error, and correlation coefficient. They also demonstrated the use of non-parametric bootstrap resampling methods for estimating confidence limits on these performance measures.

4.4.6 Parameter Sensitivity Study

This category of model evaluation is not contained within the Initial or Final Reports, but should be. Parameter analyses are included in the Initial Report, but only summarized in the form of model adjustments. Documentation has not been provided which details model sensitivity (model performance as function of stability class, wind speed, source conditions, etc.), and critical elements are not detailed (role of liquid/vapor initial split, secondary evaporation, wind shear, etc.).

A look at the data in Appendix C (Final Report) indicates potential problems with secondary source evaporation routines. Figure 5 (this report) plots VLSTRACK-predicted dose curves for 0-2 min, 0-10 min, and 0-30 min for chemical Trial 20 and the observed data. The plots clearly indicate the overprediction of VLSTRACK 1.2 for the 0-10 min and 0-30 min periods. Since the indicated wind speed (3.6 m/s) would clearly place the initial vapor puff well beyond the sampling grid after two minutes, the increase in predicted dose must be attributed to secondary evaporation. While the source for Trial 20 is a surface release, the model maximum dose develops downwind of the release point, but this is not observed in the Trial data. Identification of potential model sensitivities and impacts on model performance is a key role of model validation. In this case the model should be exercised with varying vapor/liquid splits and evaporation rates to define the model performance as related to secondary evaporation.

There are dozens of atmospheric dispersion models (U.S. DOC, 1993) used in the emergency response as well as regulatory applications; each purports to be unique to its application. A comparison of a model against other models with similar conceptual schemes provides the justification for developing a

"new and improved" model. Figure 6 is a plot of observed versus predicted dose for Trial 20, VLSTRACK 1.2, and a simple model of the form:

$$C = \frac{Q}{2\sigma_\theta \sigma_\phi x^2 U}$$

where **U** is mean wind speed, **x** is downwind distance, σ_θ is horizontal dispersion parameter, and σ_ϕ is vertical dispersion parameter. Assumptions are made that the initial vapor/liquid split is 60/40; that initial horizontal and vertical cloud sigmas are reflected in the width of the 0-15 sec sample grid; and that the liquid spill completely evaporates over 15 min. The agreement between the simple model and observed dose is quite surprising. While the simple model does appear to have a better score here than VLSTRACK 1.2, the simple model would not be expected to adequately simulate chemical/biological releases. This highlights the need for an expanded validation effort for VLSTRACK to clearly judge its performance. Increased model complexity need not necessarily translate into better model performance.

4.4.7 Concluding remarks

Much of the confusion regarding validation efforts for VLSTRACK 1.2 can be attributed to a lack of information. Only a cursory overview of the validation effort is provided in both the Initial and Final Reports. Problems in the validation effort are identified with the comparison data, the validation statistical technique, and the lack of parameter sensitivity analyses.

Many of the biological and chemical comparison data are inappropriate for this validation effort. No confidence level is provided for the comparison data, which prevents evaluating uncertainties with the model predictions. The two biological trials discussed in this report were conducted over highly wooded terrain and complex surface conditions. VLSTRACK 1.2 does not provide parameterization models for dispersion within shoreline environments or over wooded complexes. Using the Dugway data validates VLSTRACK only for site conditions similar to those at the Dugway Proving Ground. The unique source conditions of the Dugway Trials, extrapolation of surface dose contours, and the use of "educated guess" for undefined model parameters, etc. makes the validation effort suspect. Reliance on dose as the predicted quantity implicitly assumes the model not only can predict concentration levels correctly, but also can predict the exposure time. This key factor was not demonstrated in the validation effort.

Using a single statistical measure such as the difference factor (**F**) provides limited information on the model's predictive capability. As demonstrated, this difference factor is subject to large uncertainties due to sample size and the relative position of observed and predicted concentration curves.

Though roughly a factor-of-three is quoted within the Final Report as the performance score for VLSTRACK, this factor has little relevance for the model's application. Is the factor-of-three appropriate for complex terrain, varying surface conditions, all stabilities, shoreline environments, etc? The role of model validation is to provide confidence limits for the decision-maker in operational use of VLSTRACK. This has not been performed in either the Initial or the Final Validation Reports.

5.0 CONCLUSIONS AND RECOMMENDATIONS

We presented recommendations in Section 4 for the specific items discussed with regard to the VLSTRACK model formulations and its validation. In this section, we provide a summary of recommendations and conclusions.

Every dispersion model is limited by the assumptions it uses to predict atmospheric concentrations. The VLSTRACK is no exception; there are many unsubstantiated assumptions, and we comment on only some important ones below. It would be highly desirable to clearly list capabilities as well as limitations of the model in the VLSTRACK report, to increase credibility and better inform potential users.

Tuning of the VLSTRACK model should be eliminated. As indicated, tuning is not an acceptable modeling technique and confuses interpretation of model accuracy and sensitivity. The Final VLSTRACK model should be reevaluated with the existing comparison data to define variability in model performance and establish the areas of model sensitivity.

There should be a clear statement of the uniqueness of the VLSTRACK model with those parameterization schemes which define the applicability of the model. Hundreds of puff-trajectory dispersion models are currently in use and applied to potential biological and chemical release scenarios. Development of yet another puff-trajectory model requires a statement regarding its particular parameterization which address the stated application.

5.1 Model Documentation

The VLSTRACK software documentation should be rewritten to improve the clarity of presentation, and to provide more information and applicable references at each step. The document should clearly list capabilities as well as limitations of the VLSTRACK model to aid potential users. It would be particularly useful to state the key improvements and differences from the previous versions of the model, such as NUSSE3 and NUSSE4.

A User's Guide is suggested which includes both a model overview, and a description of theoretical model formulations, and input and output requirements. The User's Guide should incorporate a unified scientific/technical documentation, and a discussion of uncertainties associated with the model's use.

5.2 Model Formulation

The temperature difference ΔT between the ground and the air should not be specified if the method of Holtslag and van Ulden (1983) is used to calculate surface heat flux. When ΔT, the wind speed, and surface roughness are known, the Golder (1972) monograms can be used to relate the Monin-Obukhov length to the Pasquill stability class. The atmospheric stability class can be estimated by using simple qualitative tables (e.g., Panofsky and Dutton, 1984). Established descriptions of atmospheric variables should not be arbitrarily modified to force dispersion model predictions to agree with observed concentrations (model tuning).

The confusion in the documentation between "mixing layer" and "boundary layer" should be corrected. It must be recognized that at night, the mixing layer height is generally smaller than the height to which the surface radiation inversion extends, i.e., the boundary layer height.

The parameterizations for the horizontal (σ_u) and vertical (σ_w) velocity standard deviations used in the VLSTRACK model are not in agreement with field observations. It is strongly recommended that established formulations and or values for the constants be used. Above the atmospheric boundary layer, σ_u and σ_w can be set to small constant values instead of relating them to the surface friction velocity.

The use of a "wind variability adjustment distance" to increase cloud diffusion over complex terrain is a bad idea. A more realistic way to account for the increased diffusion in complex terrain is to use larger values of the roughness length $\mathbf{z_0}$, since this will tend to increase the computed friction velocity $\mathbf{u_*}$.

VLSTRACK should make use of the realistic treatments of the surface and mixing layers used in other puff models, e.g., INPUFF (Petersen and Lavdas, 1986). Full reflections of plume material should be used at both the surface and top of the mixing layer. To avoid the use of multiple reflections, the puff should be assumed to become well mixed in the vertical (i.e., uniform concentration with height) when the vertical standard deviation σ_z of the puff reaches a size comparable to the mixing-layer depth. Once a portion of a contaminant cloud penetrates through the mixing layer, it should be treated as a separate puff. The concept of virtual travel distance (Ludwig, 1982) should be used to provide realistic representation of plume growth after a change in atmospheric stability. In the equations for alongwind and crosswind dispersion coefficients, the addition of standard deviations is wrong, and should be replaced by the square root of the added variances. The wind shearing factor term should be justified.

A clear discussion is needed of deposition methodology for gases, and for large and small droplets or particles. Rao's (1982) analytical concentration algorithms with dry deposition and gravitational settling of pollutants, as used in the INPUFF 2.0 model, may be utilized in VLSTRACK, at least for some applications with known velocities. Wet removal processes should be included, in order to estimate the puff depletion and deposition that will occur when a puff encounters precipitation.

5.3 Model Validation

A first step should be a re-examination of the VLSTRACK model to identify its intended operating environment, range of source conditions, output requirements, and intended use. This will define the level of complexity required for model validation and the level of uncertainty acceptable in the model's predictions.

Once this theoretical assessment of the model is completed, a thorough review of all available field measurements should be conducted to establish a high quality comparison data base which encompasses the range of intended application. As with most toxic release models, there is limited opportunity to test directly against field data. While the specific nature of the toxic material may not be tested directly, the formulations for atmospheric processes contained in the dispersion model can be tested for accuracy and uncertainty. With the exception of secondary evaporation (which should be species-specific) and empirically determined source conditions applicable to munitions releases, the dispersion routines within VLSTRACK 1.2 are not unique to chemical or biological components. There are numerous dispersion field data sets such as those from the DOE's ASCOT program, the EPA's Complex Terrain Program, the EPRI's Plume Model Validation program, and USAF's Mt. Iron Tracer Study. These data sets provide an opportunity to test and validate VLSTRACK for a broad range of atmospheric and terrain conditions.

Our third recommendation is to conduct a model parameter analysis. Model validation is more than a performance score; its role is to define operating characteristics of the model. Model performance should be a function of how accurately the model simulates transport and dispersion over a range of

conditions. The intent is to provide the decision-maker with performance bounds for the model's predictive capability. While the Final Report partitions performance scores by source, this is only appropriate for common evaluation trials, and only after uncertainties in the model's predictions are evaluated for model parameterizations.

In its current form, the VLSTRACK 1.2 model validation, as detailed in the Final Report, is too limited in scope to be useful. Given the limitations of the comparison data and validation statistics, the stated model accuracy of "factor-of-three" may not be a realistic estimate of the model's uncertainty. Due to the so-called "initial calibration" of VLSTRACK 1.2, uncertainties associated with VLSTRACK predictions cannot be separated from those of the comparison data or attributed to any particular model formulation. VLSTRACK in its present form should not be used operationally without assuming large error bounds until a well-designed and expanded validation effort has been completed for the final version of the model.

6.0 REFERENCES

Bauer, T. J., 1993: Software Design Document for the Chemical/Biological Agent Vapor, Liquid, and Solid Tracking (VLSTRACK) Computer Model, Version 2.0. NSWCDD/MP-93/453, Naval Surface Warfare Center, Dahlgren, VA.

Bauer, T. J. and R. L. Gibbs, 1994: Software User's Manual for the Chemical/Biological Agent Vapor, Liquid, and Solid Tracking (VLSTRACK) Computer Model, Versions 1.5.1 and 2.0. NSWCDD/MP-94/111, Naval Surface Warfare Center, Dahlgren, VA.

Bauer, T. J. and M. G. Wolski, 1993: Initial Validation of VLSTRACK Version 1.2. NSWCDD/TR-92/647, Naval Surface Warfare Center, Dahlgren, VA.

Businger, J. A., J. C. Wyngaard, Y. Izumi and E. F. Bradley, 1971: Flux-profile relationships in the atmospheric surface layer. *J. Atmos. Sci.* **28**, 181-189.

Chamberlain, A. C., 1953: Aspects of travel and deposition of aerosol and vapour clouds. A.E.R.E. report H.P./1261. Atomic Energy Research Establishment, Harwell, Berks., England, 32 pp.

Ciolek, J. T., 1994: Results of an emergency response atmospheric dispersion model comparison using a state accepted statistical protocol. 8th Joint Conf. on Applications of Air Pollution Meteorology, Nashville, TN. Amer. Met. Soc., Boston, MA, 235-242.

Fleagle, R. G., and J, A. Businger, 1980: *An Introduction to Atmospheric Physics (2nd ed.)*. Volume 25 in International Geophysics Series, J. Van Mieghem and A. L. Hales, eds., Academic Press, New York, 432 pp.

Gifford, F.A., 1982: Horizontal diffusion in the atmosphere: A Lagrangian dynamical theory. *Atmos. Environ.* 21, 1371-1374.

Golder, D., 1972: Relations among stability parameters in the surface layer. *Boundary-Layer Meteorol.* **3**, 47-58.

Hanna, S. R., G. A. Briggs, and R. P. Hosker, Jr., 1982: *Handbook on Atmospheric Diffusion*. DOE/TIC-11223, Technical Information Center, U. S. Department of Energy, Oak Ridge, TN. Avail. National Technical Information Service as DE82002045, Springfield VA 22161, 102 pp.

Hanna, S. R., 1993: Uncertainties in air quality model predictions. *Boundary-Layer Meteorol.* **62**, 3-20.

Hicks, B. B., D. D. Baldocchi, R. P. Hosker, Jr., B. A. Hutchison, D. R. Matt, R. T. McMillen, and L. C. Satterfield, 1985: On the Use of Monitored Air Concentrations to Infer Dry Deposition. NOAA Technical Memorandum ERL-ARL-141.

Hicks, B. B., D. D. Baldocchi, T. P. Meyers, R. P. Hosker, Jr., and D. R. Matt, 1987: A preliminary multiple resistance routine for deriving dry deposition velocities from measured quantities. *Water, Air, and Soil Poll.* **36**, 311-330.

Hilst, Glenn R., 1969. Aerosol Diffusion Over Woodlot Complexes, Vol. 1, Summary and Results. TRC 343. Travlers Research Corp., Hartford CT. 165 pp.

Holtslag, A. A. M. and C. H. Moeng, 1991: Eddy diffusivity and countergradient transport in the convective atmospheric boundary layer. *J. Atmos. Sci.* **48**, 1690-1698.

Holtslag, A. A. M. and A. P. Van Ulden, 1983: A simple scheme for daytime estimates of surface fluxes from routine weather data. *J. Clim. Appl. Met.* **22**, 517-529.

Hosker, R. P., K. S. Rao, R. M. Eckman, J. T. McQueen, and G. E. Start, 1993: An Assessment of the Dispersion Models in the MARSS System used at the Kennedy Space Center. NOAA Tech. Memo. ERL ARL-205, Air Resources Lab., Silver Spring, MD, 91 pp.

Hosker, R. P., 1986: Practical application of air pollutant deposition models -- Current status, data requirements, and research needs. In *Air Pollutants and Their Effects on the Terrestrial Ecosystem*, A. H. Legge and S. V. Krupa, eds. Volume 18 in Wiley Series in Advances in Environmental Science and Technology, John Wiley and Sons, New York, pp. 505-567.

IAEA, 1989: *Evaluating the Reliability of Predictions made Using Environmental Transfer Models*. Safety Series No. 100, International Atomic Energy Agency, Vienna.

Jarvis, P. G. and K. G. McNaughton, 1986: Stomatal control of transpiration: scaling up from leaf to region. *Adv. Ecological Res.* **15**, 1-48.

Louis, J.-F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteorol.* **17**, 187-202.

Ludwig, F. L., 1982: Effect of a change of atmospheric stability on the growth rate of puffs used in plume simulation models. *J. Appl. Met.* **21**, 1371-1374.

Lyons, W. A., 1975: Turbulent diffusion and pollutant transport in shoreline environments. In *Lectures on Air Pollution and Environmental Impact Analyses*, Amer. Meteorol. Soc., Boston, MA, 136-208.

NAPAP, 1991: The U. S. National Acid Precipitation Assessment Program 1990 Integrated Assessment Report. NAPAP Office of the Director, Washington, DC. 520 pp.

Panofsky, H. A. and J. A. Dutton, 1984: *Atmospheric Turbulence, Models and Methods for Engineering Applications*. John Wiley & Sons, New York, 397 pp.

Petersen, W. B. and L. G. Lavdas, 1986: INPUFF 2.0 - A Multiple Source Gaussian Puff Dispersion Algorithm. EPA-600/8-86-011. U. S. Environmental Protection Agency, Research Triangle Park, NC, 105 pp.

Randerson, D., 1984: *Atmospheric Science and Power Production*. DOE/TIC-27601, Technical Information Center, Oak Ridge, TN, 850 pp.

Rao, K. S., 1982: Analytical solutions of a gradient-transfer model for plume deposition and sedimentation. EPA-600/3-82-079, U.S. Environmental Protection Agency, Research Triangle Park, NC; available as PB 82-215 153, NTIS, Springfield, VA.

Rao, K. S. and R. P. Hosker, 1993: Uncertainty in the assessment of atmospheric concentrations of toxic contaminants from an accidental release. *Radiation Protection Dosimetry*, **50**, Nos 2-4, 281-288.

Saucier, R., 1987: NUSSE3 Model Description. CRDEC-TR-87046, Aberdeen Proving Ground, MD.
Shir, C. C. and L. J. Shieh, 1974: A generalized urban air pollution model and its application to the study of SO2 distributions in the St. Louis Metropolitan Area. *J. Appl. Meteor.* **13**, 185-204.

Slade, D. H., 1968: Meteorology and Atomic Energy - 1968. U.S. Atomic Energy Commission Publication TID-24190, 445 pp.

Slinn, W. G. N., L. Hasse, B. B. Hicks, A. W. Hogan, D. Lal, P. S. Liss, K. O. Munnich, G. A. Sehmel, and O. Vittori, 1978: Some aspects of the transfer of atmospheric trace constituents past the air-sea interface. *Atmos. Environ.* **12** (11), 2055-2087.

Smith, T. B. and B.L. Nieman, 1969. Shoreline Diffusion Program, Oceanside California, November 1969. MR169 FR-860, Meteorology Research Inc. 230pp.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 666 pp.

Tangirala, R. S., K. S. Rao, and R. P. Hosker, 1992: A puff model simulation of tracer concentrations in the nocturnal drainage flow in a deep valley. *Atmos. Environ.* **26**, 299-309.

Thom, A. S., 1975: Momentum, mass, and heat exchange of plant communities. In *Vegetation and the Atmosphere*, vol. 1, J. L. Monteith, ed. Academic Press, London, pp. 57-109.

Weil J. C., R. I. Sykes, and A. Venkatram, 1992: Evaluating air quality models: Review and outlook. *J. Appl. Meteorol.* **31**, 1121-1145.

U.S. DOC, 1993: Directory of Atmospheric Transport and Diffusion Models, Equipment, and Projects. FCM-I3-1993, Office of the Federal Coordinator, Rockville, MD.

U.S. EPA, 1986. Guideline on Air Quality Models (Revised). EPA-450/2-78-027R, Office of Air Quality Planning and Standards. Research triangle Park, NC.

Wolski, M. G., 1993: Final Validation of VLSTRACK 1.2. Naval Surface Warfare Center, Dahlgren, VA.

Table 1

	Time (min)		Direction (degrees)		Velocity (km/hr)
Z	= 1	D2	= 45.0	V2	= 10.0
Z+1	= 2	D2	= 45.0	V2	= 10.0
Z+2	= 3	D2+d	= 57.5	V2+v	= 10.55
Z+3	= 4	D2+2d	= 70.0	V2+2v	= 11.10
Z+4	= 5	D2+3d	= 82.5	V2+3v	= 11.65
Z+5	= 6	D2+4d	= 95.0	V2+4v	= 12.20
Z+6	= 7	D2+5d	= 107.5	V2+5v	= 12.75
Z+7	= 8	D2+6d	= 120.0	V2+6v	= 13.30
Z+8	= 9	D2+7d	= 132.5	V2+7v	= 13.85
Z+9	= 10	D2+8d	= 145.0	V2+8v	= 14.45

D2	= 45 deg	V2	= 10 km/hr
D10	= 90 deg	V10	= 12 km/hr
d	= (90-45)/3.6 deg	v	= (12-10)/3.6 km/hr
d	= 12.5 deg	v	= 0.55 km/hr

Table 1. Evaluation of Table 4 (Final Report) which estimates 0 to 10 min wind direction and velocity for those chemical trials in which only 2 and 10 min averaged data were reported. Table 4 (Final Report) is used to convert the data to 0 to 10 min data with a smooth transition in wind speed and direction. However, as illustrated above, this scheme gives a change in wind direction well beyond the 10 min value, and a constant acceleration of the wind.

Table 2

Trial	Distance (km)	Trial Data (10^3 units-min/m^3)	VLSTRACK (10^3 units-min/m^3)	Factor F Difference
2.28	1.50	465.70	69.22	-8.28
2.28	8.50	4.40	3.02	-1.64
2.29	1.50	5.50	62.23	10.59
2.29	8.50	4.50	2.65	-2.30

Table 2. Chemical data for Trials 2.28 and 2.29. The data for Trial 2.29 appear inconsistent at 1.5 km when compared to Trial 2.28 at the same downwind distance, yet the Final Report indicates identical release quantities and conditions for both trials. Trial 2.28 shows a large underprediction while Trial 2.29 shows a large overprediction. The average **F** factor (see Section 4.4.5) of the two cases suggests an agreement of about a factor of 1.3 between model and field data.

Table 3

Time	Distance (m)	Trial Data (mg-min/m^3)
0-2 min	0	1044
	7.6	7500
	15.2	1456
0-10 min	0	500
	7.6	500
	15.2	964

Table 3. Observation data for chemical Trial 21. There is an apparent conflict with reported values for the 0-10 minute period. The 500 mg-min/m^3 dose at x=0 and x=7.6 m is an estimate, which is significantly below the value for the previous measurement period. Since the reported values represent accumulated doses, the dose values of 500 are suspect. The 0-10 min observation at x=15.2 m (964 mg-min/m^3) is also significantly below the corresponding 0-2 min value; the data report actually listed the 964 observation for an adjacent grid point for the 0-10 min period. All three observation points above for the 0-10 min period are therefore suspect and should be eliminated from the model validation.

Table 4

Trial 19

Time (min)	Distance (m)	Trial (mg-min/m^3)	VLSTRACK (10^3 mg-min/m^3)	Factor F	Dose Ratio
0-2	110	149	223.64	1.76	1.50
0-10	110	170	268.77	1.99	1.58

Incremental dosage change	= 21	= 45.13 (mg-min/m^3)
Average Concentration	= 2.625	= 5.64 (mg/m^3)
Factor F	= 3.32	
Ratio	= 2.14	

Trial 20

Time (min)	Distance (m)	Trial (mg-min/m^3)	VLSTRACK (mg-min/m^3)	Factor F	Dose Ratio
0-2	114	38	174.47	6.62	4.5
0-10	114	61	2514.38	16.15	41.2

Incremental dosage change	= 23	= 2339.91 (mg-min/m^3)
Average concentration	= 2.875	= 292.48 (mg/m^3)
Factor F	= 20.07	
Ratio	= 101.6	

Table 4. Data for chemical Trials 19 and 20. Each trial is evaluated to determine an incremental dose, (0-10) dose minus (0-2) dose, from which an average concentration can be estimated for the 8 min period and **Factor F** can be calculated from the incremental change. As shown for Trial 19, the ratio of observed to predicted **F** for the increment is only slightly higher; however the incremental ratio for Trial 20 is roughly 2.5 times higher.

Table 5

Atmospheric Stability	Very Stable	Stable	Neutral	Unstable
Averaging Time				
3 in to 15 min	0.12/0.6	0.52	0.35	0.4
15 min to 1 h	0.12			
1 h to 4 h	0.43			
4 h to 1 day	0.86			

Table 5. Power law exponent (p) values for different averaging times and stabilities (IAEA, 1989).

POWER PRODUCTION AND THE ATMOSPHERE

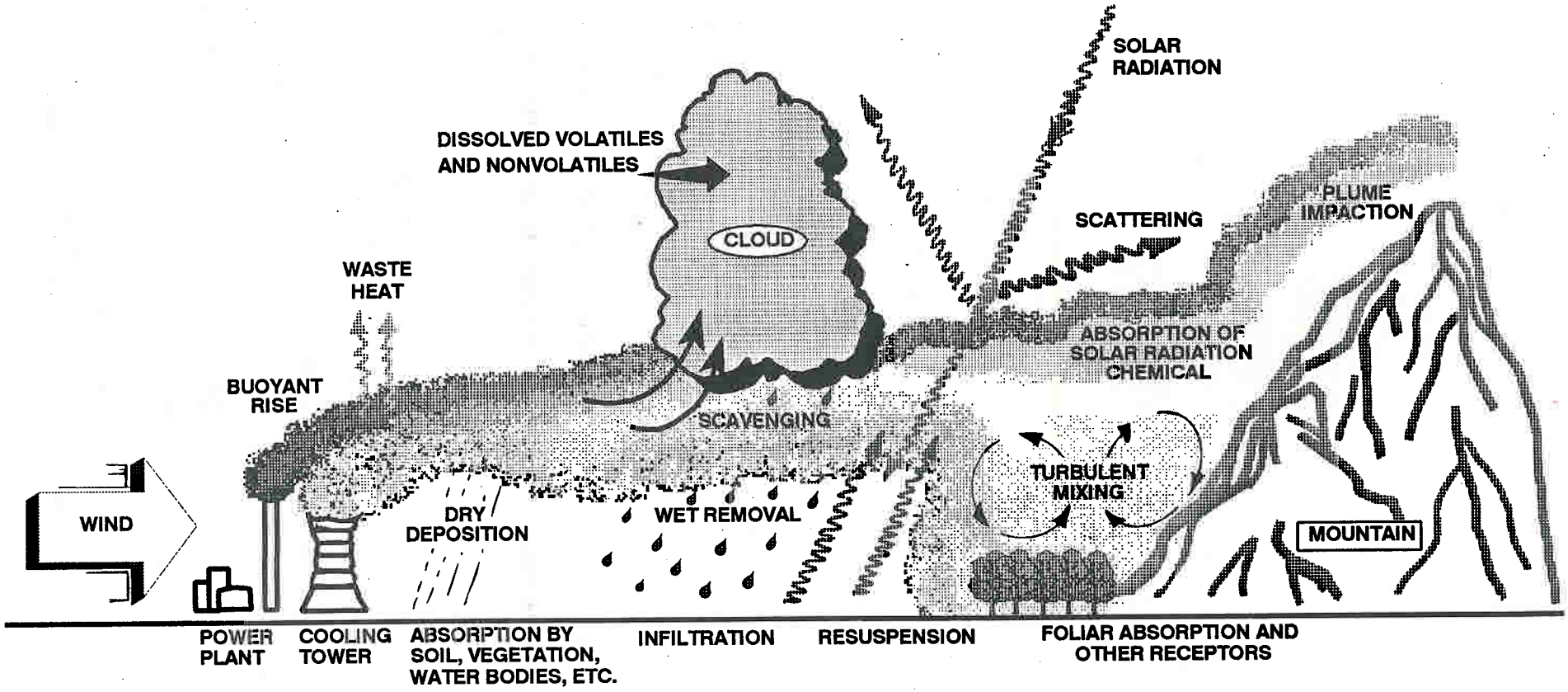


Figure 1.
Conceptual view of atmospheric processes that can act upon effluents released into the atmosphere.

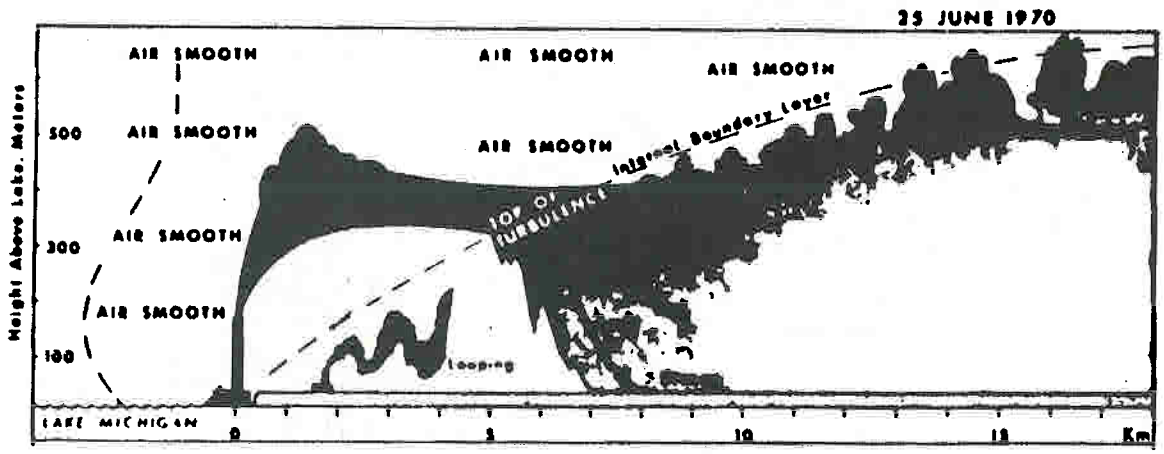


Figure 2. Conceptual model of shoreline fumigation based on field observations. From Lyons (1975).

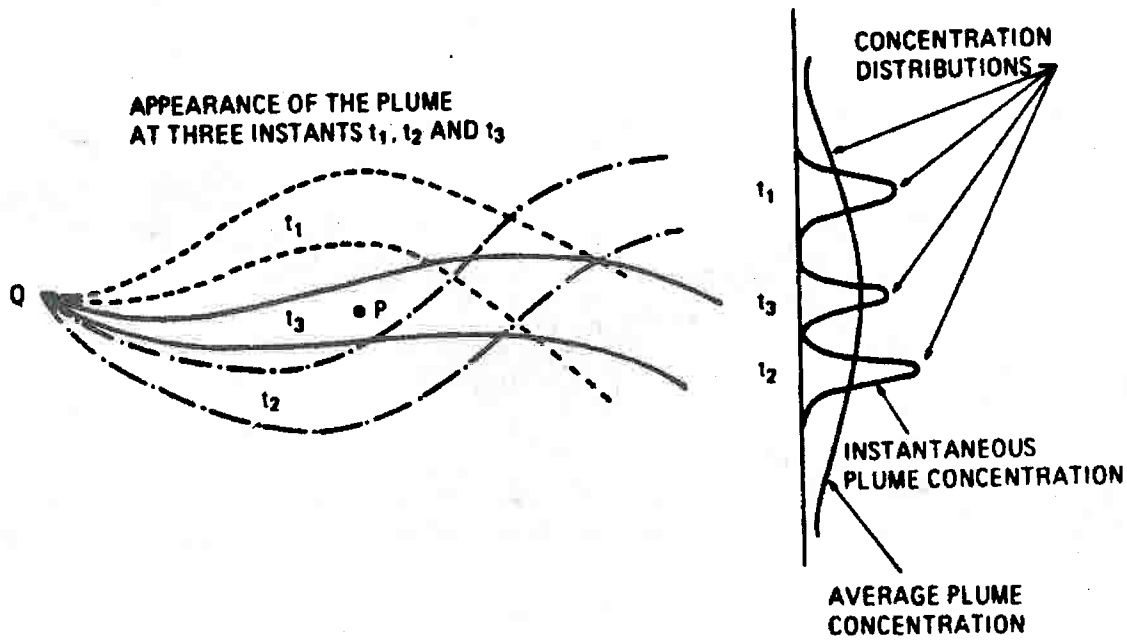


Figure 3. Potential transport trajectories and plume concentrations for various averaging times.

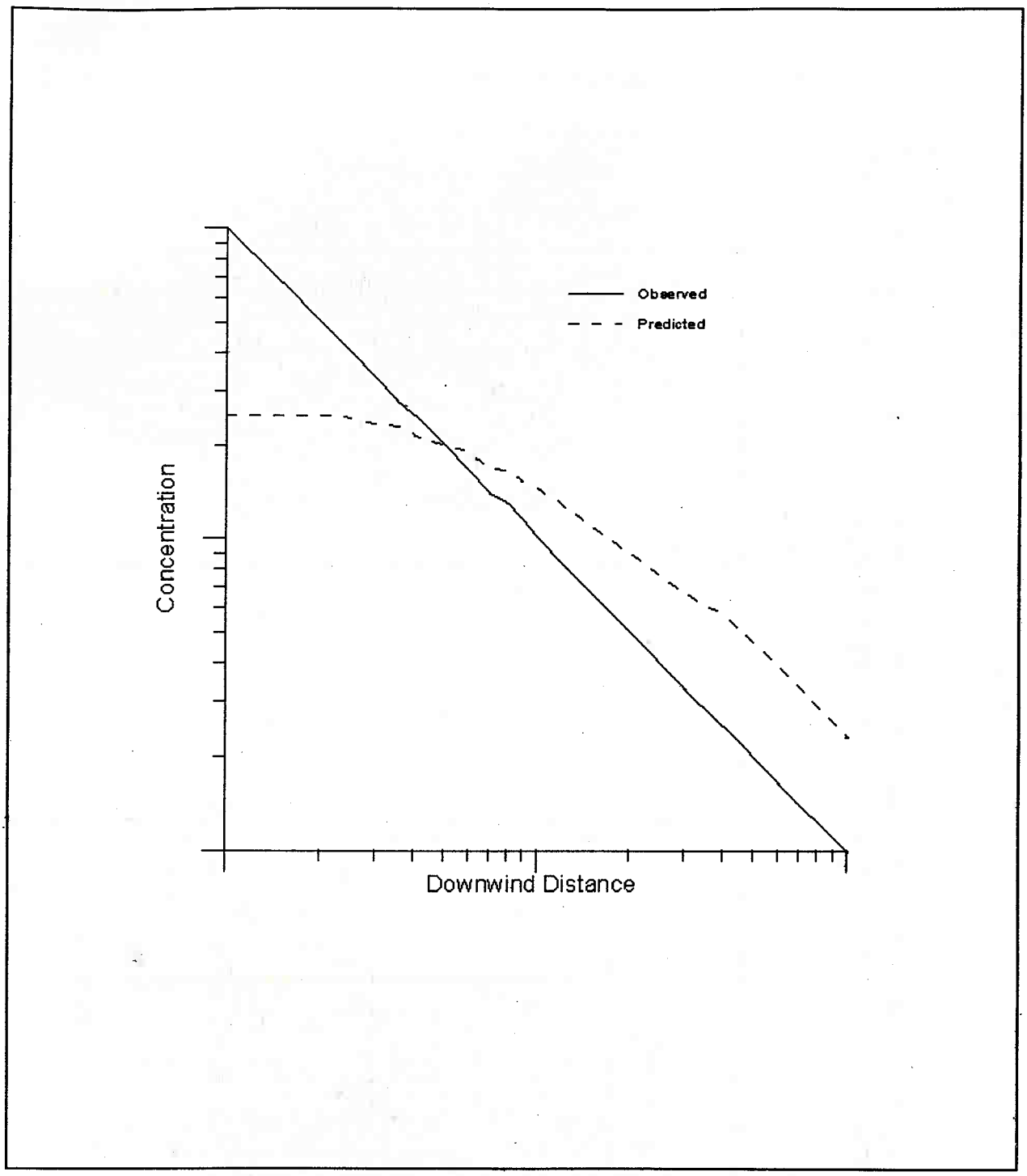
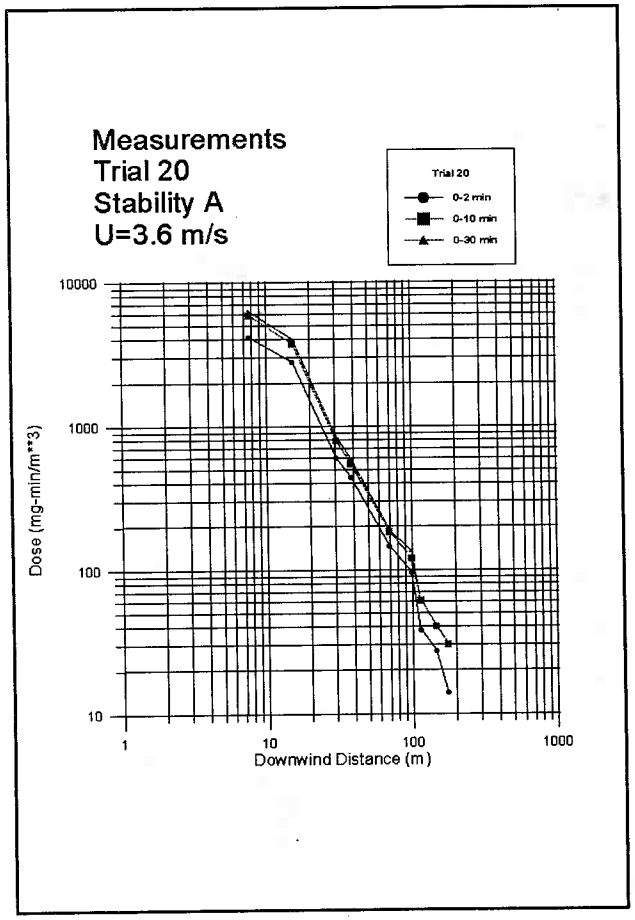
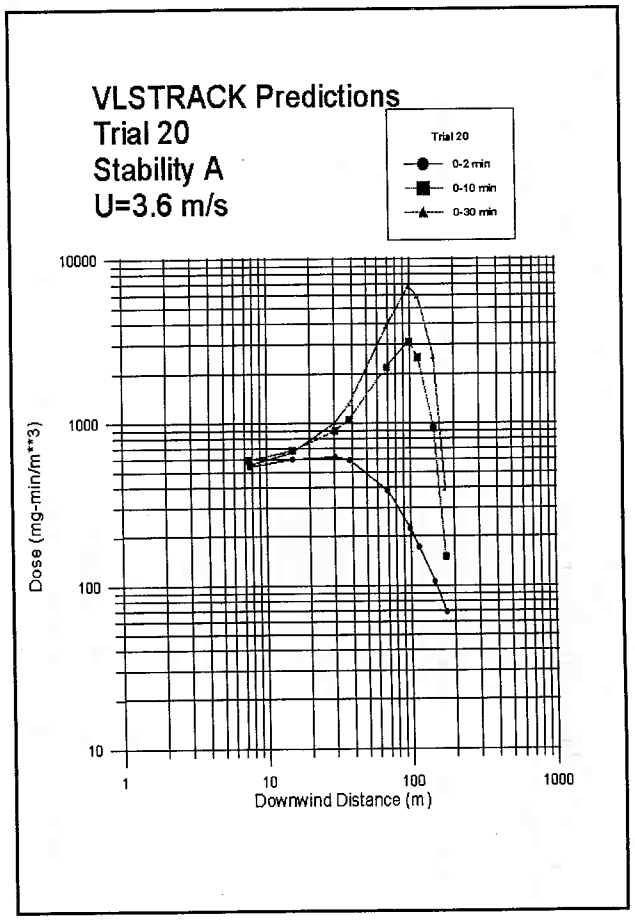


Figure 4. Illustration of problem associated with sample size and mean difference factor. By selecting more sample points near the source (or beyond the intersection point of the two curves) weights the mean difference factor towards either under- or overprediction.



(a)



(b)

Figure 5. Comparison between observed and predicted doses for chemical Trial 20. VLSTRACK typically underpredicts dose level near the source and overpredicts with distance. As shown in Figure 5b, VLSTRACK seems to overestimate contribution from secondary sources.

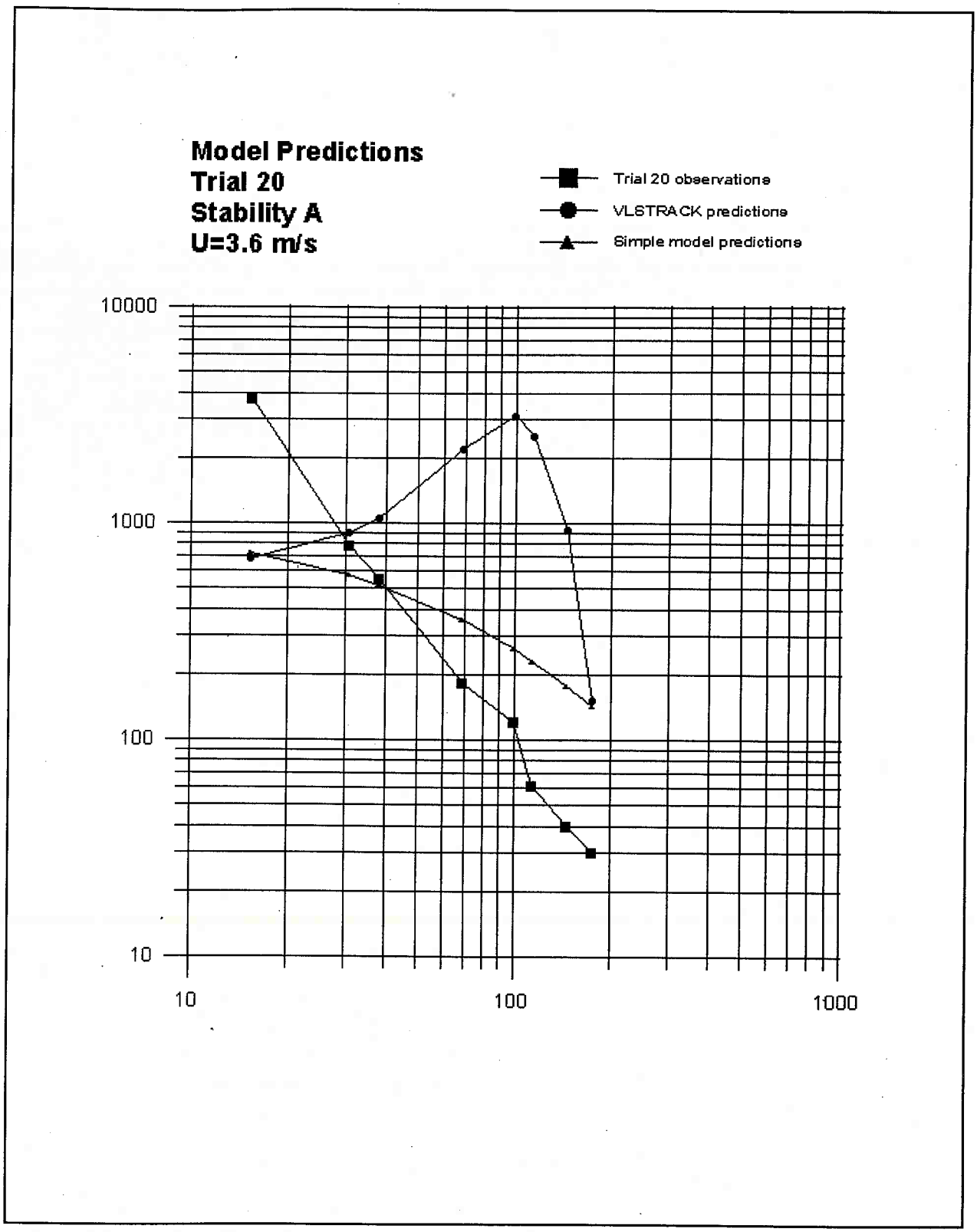


Figure 6. Observed and predicted doses as a function of downwind distance for Trial 20. Both VLSTRACK and simple model predictions are plotted.