H14-209
A METHOD FOR TARGETING CHEMICAL SAMPLERS FOR FACILITY MONITORING IN AN URBAN ENVIRONMENT

Paul E. Bieringer[1], John Hannan[2], Scott Longmore[1], George Bieberbach[1], and Tom Warner[1]

[1]Research Applications Laboratory, National Center for Atmospheric Research, USA
[2]Joint Science and Technology Office,Defense Threat Reduction Agency, USA

Abstract: The deployment of surface-based sensors/samplers is a common practice for emission and air quality monitoring purposes and the proper selection of sites for the measurement equipment is critical to an accurate characterization of the emissions. This is particularly true in urban environments where the availability of suitable sites and the non-intuitive dispersion patterns associated with the wind flow around the buildings and through the urban canyons make site selection difficult. This presentation illustrates a methodology for identifying the optimum locations for air quality monitoring equipment deployed in this complex challenging environment. The method involves using a series of coupled technologies to map the probability of detection (POD) for a given detection threshold. The approach involves the following elements: 1. A high-resolution (40 km horizontal spatial resolution) gridded climatological reanalysis; 2.a multi-dimensional feature extraction and classification technique known as the self organizing maps (SOM) that is used to characterize the weather patterns relevant for atmospheric transport and dispersion and their frequencies of occurrence; 3. the construction of building-aware wind flow fields for the urban environment for each of the SOM weather patterns; 4. interior dispersion modeling that utilizes the wind-loading pressure from the building aware wind model to identify likely material exfiltration paths; 5. and simulations from a Lagrangian particle dispersion model to map the exterior dispersion patterns.The exterior dispersion patterns and associated frequency of occurrence are then combined to estimate the map of the POD for a given detection threshold. The method is flexible and can be tuned to allow the detailed characterization of POD for a given sampler detection threshold and sampling period (e.g. sampling duration, season, time of day). An example of this methodology is illustrated for a single facility in an urban location surrounded by numerous multi-story buildings.

Key words: *Emissions monitoring, sampler locations, transport and dispersion modelling, climatology, air quality monitoring*

INTRODUCTION
Airborne contaminants are a common by-product of modern industrial societies and numerous government and non-government organizations are tasked with monitoring these emissions to assess their impact on the environment. Typically the goal of air quality monitoring is to ensure that industrial emissions are in compliance with local/national/international regulations that have been enacted to ensure public safety. In other instances, law enforcement, homeland security, and national defense organizations are tasked with detecting and identifying the intentional release of chemical and biological (CB) agents into the atmosphere and warning and/or mitigating the effects on the public and/or military forces. A key challenge for all of these organizations is identifying where to place sensing/sampling resources to make an accurate estimation of emissions, or in the case of the CB agents, that the contaminants are detected. This is particularly true in urban environments where the availability of suitable sites to deploy instrumentation and the non-intuitive dispersion patterns associated with the wind flow around the buildings and through the urban canyons make site selection difficult.

A common approach for conducting air quality emissions studies and for assessing where to deploy monitoring instrumentation involves running a series of dispersion simulations across a representative set of meteorological conditionsthat influence the atmospheric transport and dispersion (T&D) for a given area. The representativeness of the meteorological data are influenced by a number of factors: complexity of the site under consideration, proximity of the weather data location to the area of interest, and the period over which meteorological data are available. These factors vary seasonally and by location, and of critical importance is that the meteorological data cover the full range of likely conditions,including outliers that can produce a T&D eventof significance (EPA 2005).Landsberg and Jacobs (1951) examined distributions of atmospheric conditions over different time periods and found that periods in excess of ten years may be required to obtain stability of frequency distributions in weather observations. Burton et al. (1983) also examined the issue of observation record stability through a study of frequency distributions in observational records and their impact on atmospheric dispersion simulations. In this study the authors compared the concentration results from dispersion simulations that used observations of various lengths of time relative to simulation results based on a 17-year record. The goal of their study was to characterize the minimum length of observational meteorological data required to approximate the results produced from a 17-year observational record. The results of this study indicate that for a single location the variability in the dispersion model results were sufficiently reduced if a meteorological record of 5 years or greater were used. Based on the results of Burton et al. (1983), the United States Environmental Protection Agency (EPA) suggest using at least one year of meteorological data in air quality analyses that utilize dispersion simulations and strongly recommend using five years of data (EPA 2005). With the increasing availability of faster, low cost computing, it is now possible to complete a 5-year assessment of material concentrations in a reasonable amount of time with many of the commonly used dispersion models. However, if the assessment requires a more detailed assessment that requires, multiple coupled models or higher fidelity simulations, it is still prohibitively expensive to complete such an analysis with the recommended meteorological data sets.

The approach described in this paper was designed to make it logistically feasible to conduct a complex dispersion modelling assessment that utilizes meteorological conditions that are representative of the full range of conditions.Furthermore, the approachalso addresses the limitation of spatial representativeness that is typically a challenge for air quality assessments that cover larger areas. Here the approach is illustrated in the context of a methodology for identifying the optimum locations for air quality and CB agent detection monitoring equipment deployed in an urban locale. This application requires a high fidelity solution that utilizes a building-aware urban wind-flow and T&D coupled with an interior building aware dispersion model that incorporates wind-loading pressures into the interior contaminant flow and exfiltration and exterior dispersion estimates.The modeling results are then usedto determine the locations to deploy surface based airborne chemical sampling equipment. The complexity of this solution makes it prohibitively expensive to use the recommended 5-year climatological to map the probability of detection (POD) for a given detection threshold making the proposed methodology necessary.

Two key elements of this method area 21-year, climate reanalysis initially developed to support atmospheric T&D simulations and a multi-dimensional pattern recognition technique that can characterize the weather patterns relevant for T&D and their frequencies of occurrence. These atmospheric data were then used as input to urban wind-flow and dispersion models to diagnose the material dispersion patterns and detection probabilities. The approach is very flexible and can be tuned to provide a map of POD tuned specifically for the location of interest, material being emitted, and the detection hardware used. The technique is demonstrated for a scenario where contaminants are being emitted from a small building in a commercial zone where the buildings range from one four stories in height.

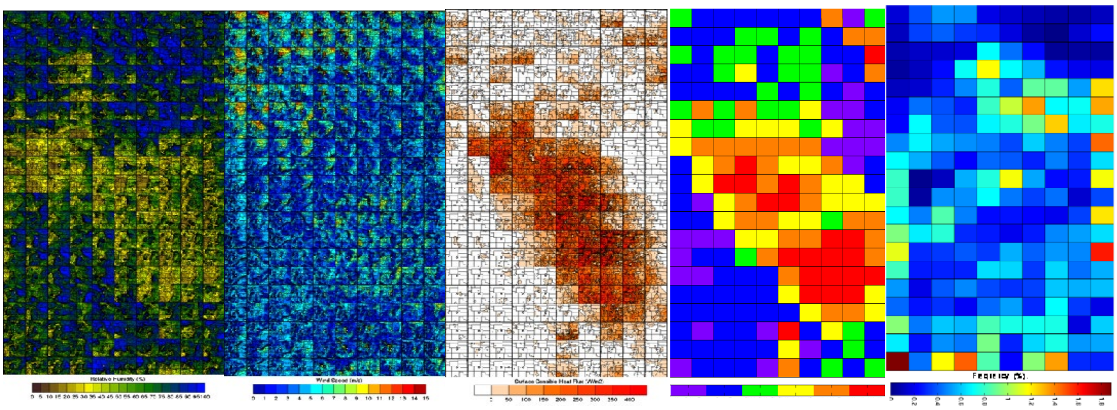
METHODOLOGY

Thesampler placement analysis is conducted in the central commercial district of Boulder CO, USA and encompasses an area of approximately 12 city blocks. The climatological database, used for this study, was developed by the National Center for Atmospheric Research (NCAR) to support aerosol transport modeling for the United States (US) Department of Defense (DoD) Joint Effects Model (JEM) and the Defense Threat Reduction Agency (DTRA) Hazard Prediction and Assessment Capability (HPAC). The database contains hourly, three-dimensional analyses of all standard meteorological variables over a 21-year period (1985 – 2005) on a global 40-km horizontal grid that extends from the surface up to a height of approximately 60,000 feet (Rife et al., 2010).The ability of this dataset to represent the mean and statistical behavior of the atmosphere, particularly within the planetary boundary layer, has been validated, utilizing a variety of observation datasets, which were not used to derive the original database. Of particular interest to this study was the validation of this dataset'sability to recreate the observed characteristics of the Great Plains Nocturnal Low Level Jet (NLLJ) using observations from the National Oceanic and Atmospheric Administration (NOAA) Profiler Network (NPN), located in the central United States (Rife et al., 2010). Additionally, the dataset's ability to replicate the observed spatial patterns of rainfall (particularly the diurnal cycle of rainfall) has also been verified by Monaghan et al., (2010). Both of these processes (boundary layer winds and moisture) are important drivers of the meteorological conditions that effectchemical and aerosol transport/dispersion.

To accurately account for the effects of all possible meteorological conditions, ideally T&D simulations would be performed for every hour contained in the full 21-year database, or approximately 183,960 simulations (21-years x ~365-days/year x 24-hours/day) per release scenario and CB threat. Although comprehensive, this technique is computationally expensive, as noted by the large number of simulations. Additionally, when one realizes that meteorological conditions or patterns tend to repeat themselves over the course of the 21-year database record this technique is wasteful. Therefore, to reduce the computational burden and associated waste, the original 21-year hourly database is decomposed into a smaller set of representative meteorological conditions using a feature extraction and classification technique known as the Self Organizing Map (SOM) (Kohonen, 1982, 1998). The SOM is an artificial neural network technique, which uses an unsupervised iterative learning technique to group large amounts of data into smaller set of clusters or "nodes", with similar characteristics. These nodes are then organized into the final map or SOM. The SOM technique has been widely used in a variety of scientific disciplines and more recently, has gained wider use within the synoptic climatology community (Hewitson and Crane 2002, Cassano et al. 2006, Schuenemann et al. 2009). In particular, SOMs were utilized by Cassano et al (2006) to analyze arctic circulation patterns, from an ensemble of global circulation model predictions, while Schuenemann et al. (2009) used the SOM technique to identify sea level pressure (SLP) patterns over Greenland from the 40-year European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40 database). Liu et al. (2006) performed an extensive evaluation of the SOM technique's ability to extract features from known patterns and provided a list of recommended SOM configurations, when developing a SOM.

For this analysis each SOM node was initialized randomly from a sample in the training data, which consists of 21 years of 24-hour periods of output for a 1000 x 1000 km domain centered over Boulder, CO.The choice of domain size was based on the need to resolve the synoptic scale patterns that play a dominant role in determining wind direction, speed, and boundary layer depth characteristics. The data vectors used to initialize and train the SOM were composed of 2-dimensional arrays of meteorological variables that most strongly impact the aerosol T&D. The variables chosen include boundary layer winds at two levels (U_{10m}, V_{10m}, U_{850mb}, V_{850mb}), relative humidity at a 2-meter elevation (RH_{2m}), and surface sensible heat flux ($SH_{surface}$).A 10x20 hexagonal lattice of nodes was used to construct the NBAF SSRA SOM. A hexagonal lattice was chosen over a rectangular lattice structure for several reasons. 1) It provides more neighbors within a given radius of a node, which can potentially strengthen the relationship between adjoining nodes;2) the overall rectangular shape of the mapping, 10 in the x direction 20 in the y direction, helps to ensure the stability of the learning process. The map size of two hundred was chosen because it is the nexus between a manageable number of meteorological scenarios, in terms of aerosol transport simulation computational expense, and a sufficient oversampling of meteorological patterns contained in the historical record.In order to maintain equal weighting during the training of the SOM, the meteorological variables, above, were each normalized to have zero mean and unit standard deviation. The SOM was trained by taking each 24-hour period of data from the training data set and finding which SOM node it best matches. Each input vector was presented to the SOM and the best matching unit (BMU) was determined by finding the smallest Euclidian distance between the input vector and each map node. The BMU and its neighboring nodes were adjusted by the weighted difference between the node and the input vector. For this study, a "bubble" neighborhood function was used, which weights all neighboring nodes, within the neighborhood radius, equally. The training process was repeated for each input vector in the data set. After cycling through the data set, the neighborhood radius and training weight were recomputed and the process repeated. After the SOM was successfully trained, the original meteorological input data were presented to the map a final time to determine node membership for each meteorological input vector. This was used to determine the frequency of occurrence of each SOM node. The best match between the training samples and the SOM patterns was then used to define a typical day (24-hour period) for each pattern, which was defined as the input vector with the closest Euclidian distance to the SOM node vector. It was this typical 24-hour period that was used to drive the resulting aerosol transport simulations for each SOM pattern.

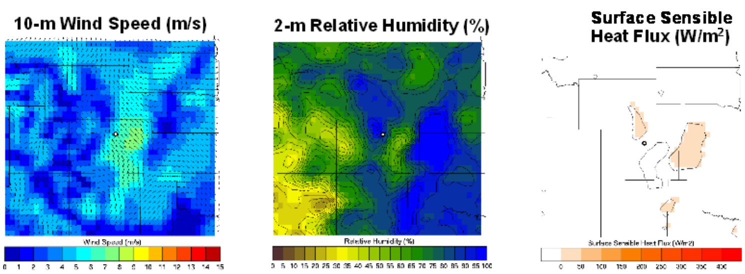
Figure 1 illustrates the spatial patterns in the relative humidity (RH) (1a), wind speed (1b), and surface heat flux (1c) identified by the SOM analysis. For the purposes of illustration, the figures represent the relative humidity, wind speed, stability classification, and the surface heat flux at the initial time of the representative SOM day. Furthermore, it is also possible to estimate the Pasquill stability classification (1d) for each of the nodes based on the 10-meter wind velocity and the incoming solar radiation. While the details in the lattice map images are difficult to see for any individual node, the images illustrate the wide range of patterns associated with each variable, and how similar patterns tend to localize in various areas of the map and are physically consistent with the other variables. For example, the RH results depicted in figure 1a show areas of higher values near the top and bottom of the lattice which correspond to low surface sensible heat flux that occurs at night. The SOM technique also provides the frequency of occurrence (1e) for each of the weather patterns. Each of the nodes represent a distribution of similar atmospheric conditions for a given weather pattern. Figure 2 illustrates the weather conditions from themost representative day at the initial hour for node 190 in the lattice (e.g. lower left corner of the image).The weather pattern associated with node 190 occurred ~1.9% of the time and for this location was the most frequently occurring pattern. From each node, the meteorological conditions relevant to atmospheric dispersion analysis (e.g. wind speed, direction, and atmospheric stability) based on the dates determined by the SOM analysis are then extracted from the climate reanalysis for the location desired, which in this case was Boulder CO and is denoted by the dot in figure 2 (a,b,c). These meteorological conditions then serve as the input to the interior and exterior dispersion simulations. The methodology provides both a physically consistent set of conditions required by the dispersion model and a corresponding frequency of occurrence that can be used to weight the dispersion model output relative to other weather patterns.



a b c d e

Figure 18. The SOM lattice maps for relative humidity(a), wind speed(b), surface sensible heat flux (c), Pasquill stability categories (d). The frequency at which each node is present in the climatological record is depicted in figure e.

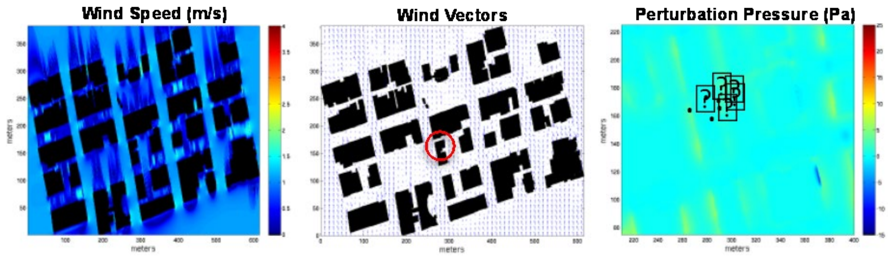
The next step in this analysis is to run a “building-aware” model to compute the wind flow through the urban terrain, the pressure loading on the buildings, and the dispersion of the contaminants around the buildings. This study used the Röckle (1990) based Quick Urban Industrial Complex (QUIC) Dispersion Modeling Systemdeveloped at the Los Alamos National Laboratory (LANL). QUIC is a fast-response, urban dispersion modeling system capable of meeting the requirements described above. The system is comprised of a wind model (QUIC-URB), a building wind loading pressure model (QUIC-PRES), a Lagrangian dispersion model (QUIC-PLUME), and a graphical user interface (GUI), QUIC-GUI (LANL, 2007). The QUIC URB/PRES/PLUMEmodeling system was chosen because its performance is representative of theRöckle class building-aware models and has been used in other relevant urban T&D applications like those developed for the Pentagon and surrounding facilities (Warner et al. 2007). Figure 3 illustrates an example of the output from the QUIC models.



a b c

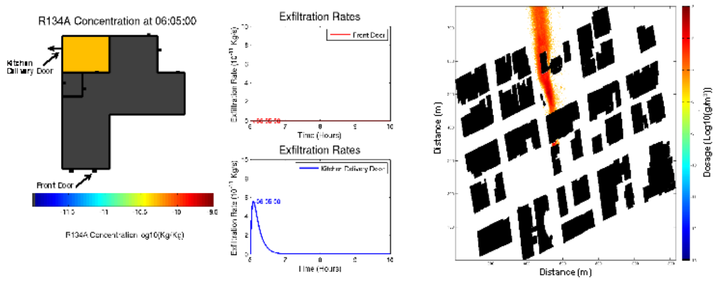
Figure 19. The SOM node 190 maps forwind speed (a), relative humidity (b), surface sensible heat flux (c). Node 190 is located at the lower left corner of the lattice map and had the highest frequency

The exterior atmospheric conditions can influence interior dispersion and for some buildings may dictate the exfiltration points on the buildings where the contaminants move from the interior to the exterior environment. To account for these effects this study used the CONTAM interior dispersion model. CONTAM is a multi-zone model capable of representing the movement of air and contaminants within structures - through open windows and doors, and through leakage around closed windows and doors (Walton and Dols, 2008). The perturbation pressure computations from the QUIC modeling system, an example of which was illustrated in figure 3c, are used as input to CONTAM. The dots represent the infiltration/exfiltration points used. CONTAM utilizes information on the room configurations, and the differential pressures between the rooms and the exterior of the building to compute the transport of the contaminants throughout the building and in this case out of the exfiltration points. The scenario used in this study involved a five-minute release of a contaminant in the kitchen of building identified by the red circle in figure 3. Figure 4 illustrates an example of the output from the CONTAM model and depicts the interior contaminant concentration (denoted by the orange color) and the flow of contaminant material (denoted by the arrow). For node 190, the wind loading pressure is on the front of the building (bottom of the image 4a). Due to the differential pressures on the building provided by QUIC-PRES, the contaminant material flows out the back kitchen deliverydoor (Figure 4b). The exfiltration estimates produced by CONTAM are then used as a source term for the exterior "building-aware" dispersion simulations. The exterior dispersion solution from QUIC-PLUME for node 190 is illustrated in Figure 4c.



a b c

Figure 20.The output from the QUIC modelling system for SOM node 190 forwind speed (a), wind direction vectors (b), and perturbation pressure (c). The dots on figure c correspond with infiltration/exfiltration points that correspond with the configuration of the CONTAM interior dispersion model.



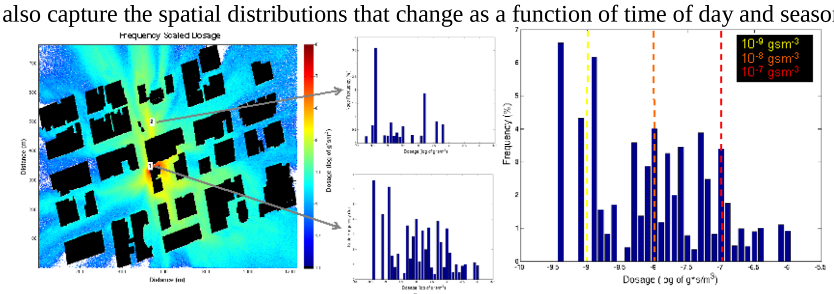
a b c

Figure 21.The output from the CONTAM interior and QUIC exterior "building-aware" dispersion models. The images depict the interior contaminant concentration and direction of material flow (a), the exfiltration rates through the front and kitchen delivery doors (b), and the exterior contaminant dispersion (c) for the weather conditions from node 190.

RESULTS AND CONCLUSIONS

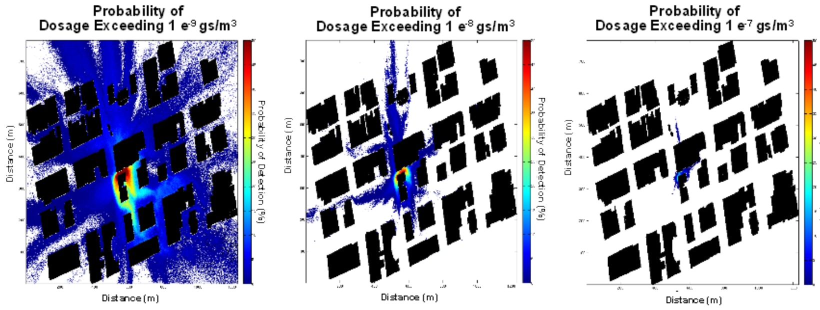
Coupled interior and exterior dispersion simulations, like those shown in figure 4c, were produced for all 200 weather patterns identified by the SOM analysis. The results were compiled in a number of ways to provide guidance for chemical sampler placement. In most cases, time-integrated samples are used for emissions monitoring. For the purpose of illustration in this study the following examples assume a 1-hour integration time and are referred to as a dosage estimate. This sample period could however be adjusted to correspond to match any sampling strategy desired.An indication of the most common dispersion modes within the urban environment is often useful for instrument placement. An example of this is shown in figure 5a that depicts a map of the 1-hour integrated dosage scaled by the frequency of the weather pattern. The 1-hour dosage from all of the SOM weather patterns can also be combined with their corresponding frequency of occurrence to produce a distribution of the frequency that a given dosage will occur. This distribution is available for every location in the domain. Two example distributions are shown in figure 5b. Assuming that the minimum detectable signal for the sampling equipment is known, it can then be applied as a threshold to this distribution to estimate the POD for the sampler at a given location. This is determined simply by summing the frequencies for all of the dosages higher than the threshold. Since this probability of detection is known at every location, it can then be plotted on the map to illustrate the spatial distribution of POD for a sampler with a given detection threshold.Figure 6 depicts the POD maps for 1-hour integrated samples at the three detection thresholds given the assumption that it is equally likely that the material release could occur at any time during the year. This map along with logistical information can then be used to make sampler placement decisions. Furthermore, by adjusting the SOM analysis to only account for weather patterns in a given season or time of day, the POD maps can be further tailored to meet the specific sampling objectives of any emission survey.

Identifying where to place sensing/sampling resources to ensure an accurate estimation of contaminant emissions continues to be a critical unresolved challenge for the air quality, law enforcement, and defense communities. This is particularly true in urban environments where the availability of suitable sites to deploy instrumentation and the non-intuitive dispersion patterns associated with the wind flow around the buildings and through the urban canyons make site selection difficult. A common approach for conducting air quality emissions studies and for assessing where to deploy monitoring instrumentation involves running a series of dispersion simulations across a range of meteorological conditions considered to be representative of the weather influencing the atmospheric transport and dispersion for a given area. In spite of significant advances in computational resources used for air quality models, it is still prohibitively expensive to run dispersion models to make equipment placementdecisions in complex urban situations. The method demonstrated here leverages weather pattern recognition techniques to address this limitation and enables the construction of detailed POD maps for a given sampler detection threshold and sampling period. The approach is flexible and can be designed to not only address sampler characteristics, but also capture the spatial distributions that change as a function of time of day and seasonal weather effects.



a b c

Figure 22.Exterior contaminant dosage results. The images depict the 1 hour integrated dosage scaled by the SOM frequency (a), the distribution of dosage frequency across all 200 nodes for two locations (b), and an example of the setting dosage thresholds to estimate the probability of exceeding a given dosage threshold (c).



a b c

Figure 23.Probability of detection maps. The images depict the 1 hour integrated dosage POD for a dosage exceeding 1 e^{-9}gs/m^3 (a), 1 e-^{8}gs/m^3 (b), 1 e^{-7}gs/m^3 (c).

REFERENCES

Burton, C.S., T.E. Stoeckenius and J.P. Nordin, 1983: The Temporal Representativeness of Short-Term Meteorological Data Sets: Implications for Air Quality Impact Assessments. Systems Applications, Inc., San Rafael, CA. (Docket No. A–80–46, II-G–11).

Cassano, J.J., Uotila, P., and Lynch, A., "Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, Part 1: Arctic," *Int. J. Climatol.*, Vol. 26, 2006, pp. 1027-1049.

Environmental Protection Agency, 2005: Revision to the Guideline on Air Quality Models: Adoption of a Preferred General Purpose (Flat and Complex Terrain) Dispersion Model and Other Revisions; Final Rule. United States Federal Register, Rules and Regulations, Vol. 70 No. 216.

Hewitson, B.C., and Crane, R.G., "Self-organizing maps: applications to synoptic climatology," *Clim. Res.*, Vol. 22, 2002, pp. 13-26.

Kohonen, T., "Self-organized information of topologically correct feature maps," *Biol. Cyber.*, Vol. 43, 1982, pp 59-69.

Kohonen, T., "The self-organizing map," *Neurocomputing*, **21**, 1998, pp. 1-6.

Landsberg, H.E. and W.C. Jacobs, 1951: Compendium of Meteorology. American Meteorological Society, Boston, MA; pp. 976–992.

Liu, Y., Weisberg, R.H., and Mooers, C.N.K., "Performance evaluation of the self-organizing map for feature extraction," *J. Geophys. Res.*, Vol. 111, 2006, pp. 1-14.

Los Alamos National Laboratory, cited 2007: QUIC-URB & QUIC-PLUME: Fast Response Urban Dispersion Modeling System [Available online at [http://www.lanl.gov/orgs/d/d4/atmosphere/docs/LA_UR_02_2559.pdf.]

Monaghan, Andrew J., Daran L. Rife, James O. Pinto, Christopher A. Davis, John R. Hannan, 2010: Global precipitation extremes associated with diurnally varying low-level jets. *J. Climate*, **23**, 5065–5084.

NCDC (National Climate Data Center), 2011: "Quality Controlled Local Climatological Data," Rocky Mountain Regional Airport Hourly ASOS Observations [online database], http://www.ncdc.noaa.gov/oa/ncdc.html [retrieved 8/2011].

Rife, Daran L., James O. Pinto, Andrew J. Monaghan, Christopher A. Davis, John R. Hannan, 2010: Global distribution and characteristics of diurnally varying low-level jets. *J. Climate*, **23**, 5041–5064.

Röckle R., 1990: Bestimmung der Stromungsverhaltnisse i m Bereich komplexer Bebauungsstrukturen, Ph. D. dissertation, Vom Fachbereich Mechanik, der Technischen Hochschule Darmstadt, Germany.

Schuenemann, K.C., Cassano, J.J., and Finnis, J., "Synoptic Forcing of Precipitation over Greenland: Climatology for 1961-99," *Journal of Hydrometeorology*, Vol. 10, 2009, pp. 60-78.

Walton, G.N. and W.S. Dols. CONTAM 2.4c User Guide and Program Documentation.*NISTIR7251*. National Institute of Standards and Technology (October 2008).

Warner T. and Coauthors, 2007: The Pentagon Shield field program: Toward critical infrastructure protection, *Bull. Amer. Meteorol. Soc.*, **88**, 167.