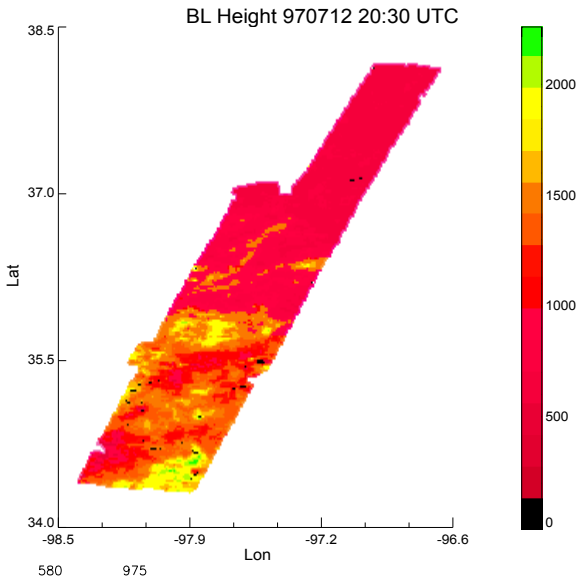
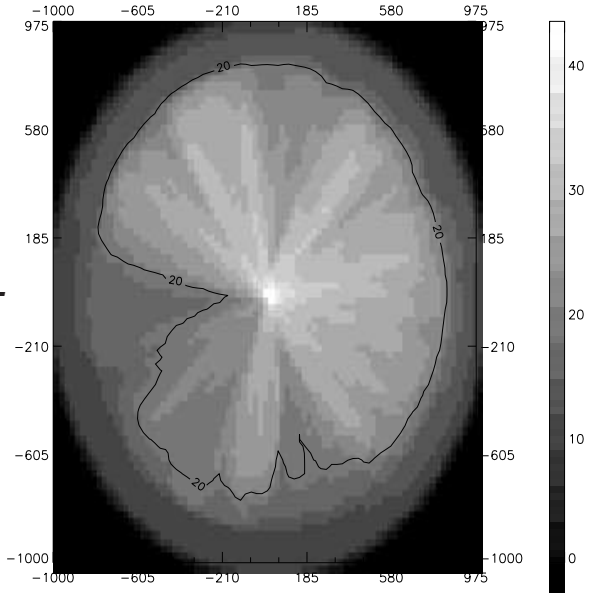
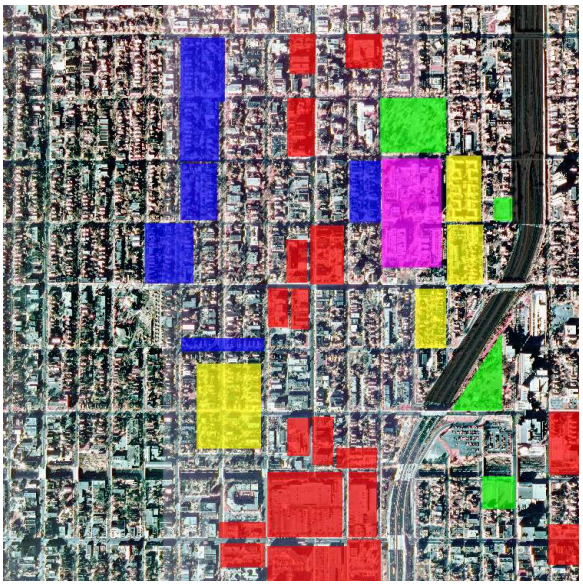
Plan B Papers

Surface energy flux dependence on remotely-sensed soil moisture and vegetation in the Southern Great Plains experiment



Point-source air pollution risk buffers from long-term climate data and Gaussian plume models





Interaction of resolution and classification in high-resolution urban/suburban land cover imagery

Ankur Desai
Deparment of Geography
University of Minnesota
April 2000

Surface energy flux dependence on remotely-sensed soil moisture and vegetation in the Southern Great Plains experiment

Abstract

Remotely sensed surface characteristics are compared to surface energy fluxes. A successful methodology in this realm would provide a method for measuring surface fluxes remotely and characterize surface energy flux heterogeneity across the landscape. Tower measured sensible and latent heat fluxes were compared to AVHRR vegetation indices and airplane-mounted passive microwave soil moisture measurements across the Southern Great Plains Experiment area (SGP-97). Results showed weak correlation between surface fluxes and NDVI. Correlation between remotely sensed soil moisture and surface fluxes was much higher, and better on a dry day than a wet day. This correlation was used to produce hourly SGP-97 flux maps, which were then used to estimate mid-day boundary layer height across the SGP experiment area. Spatial variations in the surface energy balance show some correlation to spatial patterns of boundary layer development. Advection and divergence, not analyzed in this study, are also important in determining boundary layer depth. The modeled boundary layer height was compared to LIDAR measurements of mixed layer depth.

Introduction

Energy balance between the Earth's surface and atmosphere requires that incoming radiation from the sun equals the outgoing energy from the ground, minus the energy that goes into heating the ground. The energy that is absorbed by the surface and subsequently returned to the atmosphere can be partitioned into three components: sensible and latent heat flux, and infrared radiation (Stull, 1995).

This energy balance is measured in terms of flux. Flux is the transfer rate of a quantity per unit area per unit time (Stull, 1995). The net radiation reaching the surface is the sum of downwelling solar radiation transmitted through the air, upwelling solar radiation reflected by the surface (albedo), downwards diffusive IR radiation from the air and upwelling IR emission from the surface. The flux of net radiation should equal the sum of upward sensible and latent heat flux and the energy flux the goes into heating the ground (molecular heat flux and storage of internal energy) at any point in time. Surface sensible heat flux is the rate of heating of the atmosphere by the ground. Surface latent heat flux the rate of moisture vaporization /

condensation between the ground and atmosphere. The way net radiation is partitioned into the component fluxes has a diurnal cycle, with sensible and latent heat fluxes (heating and evaporation) tending to reach a maximum shortly after local noon over land (Wallace and Hobbs, 1977). Atmospheric conditions such as cloud cover, however, can effect how and when this occurs. Night time is typically characterized by a "negative" heat flux (heating of the ground, and condensation of water) (Wallace and Hobbs, 1977).

Less well understood is the spatial variability of these fluxes. Land cover characteristics can greatly affect how the outgoing surface energy is partitioned. For example, a forested or vegetated area may have a lower sensible heat flux and a higher latent heat flux than a desert or bare area. Moister ground exposed to the atmosphere may have a lower sensible and higher latent heat flux than drier ground (Browell et al., 1997).

This variability in surface energy fluxes has been noted to be an important factor in driving and conditioning the atmospheric boundary layer (ABL), the lowest section of the troposphere (the first 10 km of the atmosphere). The ABL is the section of atmosphere that is directly influenced by the presence of the Earth's surface, and responds to surface forcings with a timescale of an hour or less (Stull, 1988). The boundary layer is characterized by terrain and convection generated turbulence leading to highly variable winds and efficient mixing. The height of the boundary layer increases during the day as heat energy from the sun increases the rate of surface-atmosphere energy exchange and related turbulence. This height can vary from near the surface to 3000 meters above ground during the day.

Surface energy flux partitioning has an impact on thunderstorm development, tornado potential, rainfall intensity, cloud cover and atmospheric stability. Spatial variability in the surface energy balance is being recognized as an important determinant in micro and mesoscale

weather evolution. Spatially gridded flux measurements are also useful as inputs to mesoscale atmospheric models such as large eddy simulations (LES). Finally, fine scale maps of surface fluxes can be used to thermodynamically compute atmospheric boundary layer (ABL) inversion height, another important variable in weather and climate. The temporal evolution of the depth of convection in the boundary layer may have a profound influence on local climate and meteorology (Browell et al., 1997).

Knowledge of the relation between land cover variables and surface heat fluxes would provide for a way to spatially map the variability of surface energy fluxes given information about land cover. This information could be derived from remote sensing. Direct remote sensing of surface energy fluxes from satellite is neither currently, nor likely possible, since the signal of flux partitioning is blurred by turbulence and mixing in the atmospheric boundary layer, though advances in Dopper-DIAL Lidar and earth Radar-RASS may change that (Jackson, 1997). In the meantime, a correlation or a physical connection between remotely measurable surface characteristics and surface energy fluxes would provide the key to mapping them. Variables that are related to energy fluxes include direct and diffuse shortwave solar radiation, surface albedo, surface temperature and emissivity, longwave radiation fluxes, cloud cover, soil moisture, air water vapor pressure, surface and aerodynamic resistance to latent heat, leaf area index, atmospheric stability and substrate conductivity (Greenland, 1994). Many of these variables can be measured from remote sensors.

Direct measurement of the surface energy budget from the ground is currently generally limited to point based methods involving tower-mounted equipment (i.e. sonic anemometers). Point based flux measurements are an expensive and difficult way to observe the variation of surface fluxes across the landscape. Numerous researchers have attempted to create and test

methodologies for the measurement of surface fluxes through satellite remote sensing devices (Anderson et al., 1997; Brutsaert et al., 1993; Chehbouni et al., 1997; Pelgrum et al., 1996). These methods can be either universal (relation for all landscape types) or landscape-dependent (i.e. holds only in grasslands), and statistically based (i.e. linear correlation) or physically based (modeled from first principles). No method has been entirely successful, and no universal, physically-based model has been verified. Still, some models have worked well for researchers' local experiments, suggesting that a relationship can be made for different landscapes.

Early relational models often relied on measuring the change in surface temperature and relating that to surface heat flux based on similarity relations of heat flux. Brutsaert et al. (1993) attempted to derive sensible heat flux from satellite measured surface temperature and bulk similarity models of heat flux for forested areas. However, Chehbouni et al (1997) showed that for sparsely vegetated surfaces, this parameterization was not correct, and instead relied on a combination of remotely sensed temperature, leaf area index, and a coupled Soil Vegetation Atmosphere Transfer model to compute sensible heat in southwest Niger. Experiments and models have shown that vegetation cover and water availability appear to be the primary factors in driving flux partitioning, especially in sparsely vegetated areas. Pelgrum et al. (1996) uses spectral band ratio vegetation indices in a surface energy balance algorithm to estimate evaporative fraction, which is related to latent heat flux. The most sophisticated models to date take advantage of multiple sources of data about the land surface. Anderson et al. (1997) argue that determining the spatial variability of fluxes requires both ground and remotely sensed data. They use a coupled soil and vegetation model that relies on multitemporal remote sensing of surface temperature, remotely sensed imagery of vegetation cover, a land cover database of soil and land type, and a morning potential temperature balloon (radiosonde) sounding to solve a set

of equations relating temperature changes and land cover to fluxes. This method was tested successfully at a tallgrass prairie in Kansas.

This paper does not delve into any form of a universal, physically based model, but rather attempts to explore the relationship of two remotely sensed variables to surface energy fluxes for the Southern Great Plains (SGP-97) research area. The purpose of this experiment is to explore the utility of measurements made during SGP-97 and to characterize how surface heterogeneity is expressed as atmospheric variability. Given a positive relation between mapped surface data and tower-measured flux data, surface flux maps can be made, and a time series of flux maps can be used to estimate boundary layer growth over space.

The two remote sensing instruments studied here are Advanced Very High-Frequency Radiometer (AVHRR) and Electronically Scanned Thinned Array Radiometer (ESTAR), and the two respective land cover variables studied are vegetation indices and surface soil moisture. ESTAR is a passive microwave-based remote sensing system that was flown on a P-3 aircraft across the SGP-97 experiment area for several days to validate ESTAR operation and soil moisture retrieval algorithms (Jackson et al., 1999). Along with ESTAR, a differential absorption Light Detection and Ranging (LIDAR) device was also flown to produce high resolution vertical atmospheric scattering ratio and water vapor profiles (Browell et al., 1996). The LIDAR flown is the Laser Atmospheric Sounding Experiment (LASE), a satellite-prototype differential absorption water vapor system developed by NASA Langley.

Theory suggests that both the amount of vegetation and amount of surface moisture should affect energy flux partitioning (Greenland, 1994). For the Southern Great Plains, MacPherson (1998) showed that the ratio of latent to sensible heat flux along several flux aircraft flight tracks was related to vegetation greenness (Figure 1). Anderson et al. (1997) argue that

soil moisture may be a factor in controlling morning surface temperature change, and thus sensible heat flux. Browell et al. (1997) suggest that surface soil water content is the dominant factor in determining how net radiation partitions into latent and sensible heat. Comparing remotely sensed vegetation and soil moisture allows for testing which variable is more important in determining surface flux partitioning for the Southern Great Plains.

For this experiment, ESTAR-derived soil moisture had a stronger relationship to surface energy flux partitioning than AVHRR-measured vegetation indices. This is probably not universally true. A relationship between ESTAR-derived soil moisture and the surface energy flux portioning was used to create boundary layer height maps that were compared to boundary layer height measurements from LIDAR scattering ratio profiles.

Field Experiment Description

The Southern Great Plains (SGP-97) field experiment occurred during late spring and summer of 1997 over northern Oklahoma and southern Kansas (Figure 2). The area can be characterized as subhumid grasslands with flat to moderately rolling terrain and a maximum relief of less than 200m. Rangeland and pasture are the dominant land use types (Jackson et al., 1999).

One of the primary goals of SGP-97 was validation of the soil moisture retrieval algorithms of ESTAR. An additional objective was to examine the effect of soil moisture on the evolution of the atmospheric boundary layer and clouds over the Southern Great Plains (Browell et al., 1997). Research facilities in the area include the United States Department of Agriculture (USDA) Agricultural Research Service (ARS) Little Washita watershed (LW), the USDA ARS Grazinglands Research Lab at El Reno, OK (ER), and the Department of Energy (DOE)

Atmospheric Radiation Measurement (ARM) Cloud and Radiation Testbed (CART) Central Facility (CF) near Lamont, OK. During SGP-97, these agencies recorded large volumes of high density weather data. In this experiment, data from the DOE ARM-CART flux stations and balloon soundings, and the state of Oklahoma Mesonet weather stations are used. A number of researchers at other institutions affiliated with SGP-97 also recorded flux data, including researchers at the University of Wisconsin - Madison, National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL), Georgia Tech, NOAA Atmospheric Turbulence and Diffusion Division (ATDD), NASA Goddard Space Flight Center and the University of Arizona (Hydrology Data Support Team, 2000). Some of that flux data also is used in this experiment. Remotely sensed airplane data from ESTAR and LASE were collected with the assistance of USDA-ARS, NASA Langley Research Center and affiliated researchers.

Data description

The three pieces of data to used to construct the relationship between remotely-sensed surface parameters and the surface energy balance in this experiment are: gridded soil moisture, a remotely sensed vegetation index and a spatially distributed set of measurements of surface energy flux over time. This allows point-based flux measurements to be compared against a grid of soil moisture and vegetation indices. These data come respectively from airplane-mounted ESTAR, NOAA AVHRR, various tower-based flux stations spread across the SGP-97 area. These data are limited to clear, sunny days since otherwise the ESTAR soil moisture relation may not hold, and AVHRR and LASE would be cloud-obscured. For SGP-97, clear, sunny days and the availability of ESTAR, AVHRR, LASE and a full suite of flux data coincide on July 12

and 13. A rainstorm came through on July 10 (MacPherson, 1998). Consequently, the surface is drier on July 13 than July 12, as result of the gradual dry-down from the storm.

Creation of boundary layer height maps requires the use of ARM-CART balloon soundings and are checked against an independent measure of boundary layer height from LASE. Balloon soundings were launched at ARM-CART stations every three hours on both dates. Each suite of measurements are briefly described below:

AVHRR

AVHRR is a suite of National Oceanic and Atmospheric Administration (NOAA) satellites that produce visible and infrared images of the Earth's surface at regular intervals. AVHRR red and near infrared channels are used to create a variety of popular vegetation indices, including the Normalized Difference Vegetation Index (NDVI) (Avery and Berlin, 1992). The output resolution of AVHRR is 1,100m in the green, red and infrared bands. AVHRR data for the SGP-97 region is available in the morning, afternoon and evening on July 12 and the morning of July 13. In this experiment, daytime and afternoon AVHRR data is used from the NOAA-12 and NOAA-14 satellites, respectively. Several vegetation indices were considered, but the most common one in use is the Normalized Difference Vegetation Index (NDVI). NDVI is calculated for each pixel as:

$$NDVI = \frac{NearIR - \mathrm{Re}\, d}{NearIR + \mathrm{Re}\, d} \qquad (1)$$

(Avery and Berlin, 1992).

This value ranges from -1 to 1, with high values related to fully vegetated areas (high amount of red absorption by chlorophyll and relatively low IR radiation release from plants), and low values related to bare or open areas (low amount of red absorption and high amount of

infrared absorption by the ground). Typical NDVI values for vegetated areas tend to be around 0.6, with only marginal increases in NDVI with increasing vegetation cover after that (Gillies et al., 1995). NDVI is often used as a proxy measure for Leaf Area Index (LAI) (Chehbouni, 1997). Figure 3 shows NDVI for the SGP-97 area on July 12 and July 13.

ESTAR

ESTAR was designed as a passive microwave remote sensor to be used in space for global soil moisture mapping at relatively coarse resolutions (10-30 km). Surface (< 2 cm) soil moisture has previously been shown by Jackson (1997) to be related to passive microwave brightness temperature. During SGP-97, USDA ARS and researchers with NASA flew ESTAR aboard a P-3B aircraft from the NASA Wallops Flight Facility at an altitude of 7.5 km. ESTAR was flown 30 times in the months of June and July along a set of four flight lines to test the soil moisture retrieval algorithm.

Jackson et al. (1999) describe the basic method of how ESTAR calculates soil moisture. ESTAR is a synthetic aperture, passive microwave radiometer with a frequency of 1.1413 GHz and bandwidth of 20 MHz. Synthetic aperture radars collect data by using a pair of antennas and computing the complex correlation of their output voltages at different baselines that are calibrated against a black body (water for SGP-97). The inverse Fourier transform from the signal at different baselines provides the microwave brightness temperature response across track. Along track data is provided as the plane moves along its path. The resulting field of view is about 45 degrees on either side of the airplane scanned in 2 degree overlapping steps. This information is then gridded to produce a map of microwave ground reflectance across and along the airplane track.

The soil moisture algorithm used an inverse model equation (Fresnel reflectance equation for horizontal polarization) that relies on microwave brightness temperature, soil temperature, vegetation type, vegetation water content, surface roughness, soil bulk density and soil texture to calculate soil moisture. Soils data from the Oklahoma Mesonet and vegetation information from LandSat Thematic Mapper (TM) were used in the model. The output map was gridded to 800m resolution. This mapped soil moisture was successfully validated against sixteen soil moisture field sites that had 9-14 samples per site (Jackson et al., 1999). Figure 4 shows ESTAR based soil moisture on July 12 and July 13.

Surface fluxes

Flux measurement stations were located across the SGP-97 region as part of the ARM-CART flux measurement program and individual researchers who had set up stations. Surface energy flux is the exchange of heat and moisture energy in time per unit area between the Earth's surface and the atmosphere. The surface energy balance is based on closure of the following equation:

$$R\downarrow = H\uparrow + LH\uparrow + G\downarrow + \Delta s \qquad (2)$$

(Stull, 1988)

where R is net radiation (solar + IR) , H is sensible heat flux, LH is latent heat flux, G is soil heat flux and Δs is the change in storage of heat by the soil due to chemical processes. The latter quantity is known to be negligible in surface energy balance, and so is assumed to be zero. By knowing three of the remaining four values, the fourth can be easily computed. R is measured with a net radiometer. G can be estimated by measuring soil temperature and soil moisture changes over time (Atmospheric Radiation Measurement Program, 1999(2)). Two different

methods were used to calculated H and LH in the SGP-97 region. They are the Bowen ratio method and the Eddy covariance method.

The Bowen ratio method, sometimes called the Priestly-Taylor method (Stull, 1988) attempts to directly measure the Bowen ratio, which is the ratio of vertical sensible heat flux to vertical latent heat flux (H / LH). Unlike the fluxes, Bowen ratio tends to be constant for any particular spot during the daytime. The ratio can be estimated as the ratio of the temperature gradient over the water vapor mixing ratio gradient between two different heights. These two points should be one near the surface and one a few meters above the surface. The basic equation for this is:

$$B = \frac{H}{LH} = \frac{c_p}{L_v} \frac{\partial \bar{\theta} / \partial z}{\partial \bar{q} / \partial z} \tag{3}$$

(Stull, 1988)

where B is Bowen ratio, H is sensible heat flux, LH is latent heat flux, c_p is the specific heat of air at constant pressure, L_v is the latent heat of vaporization of water, θ is potential temperature (temperature normalized for altitude), q is water vapor and z is height about ground. This equation is usually calculated with time averaged (around 30 minute) observations of temperature and water vapor at two points. Once the Bowen ratio is known along with observed R (from a radiometer) and G (from soil temperature/moisture observations), H and LH can be computed as follows:

$$LH = \frac{(R - G)}{(1 + B)} \tag{4}$$

$$H = R - G - L \tag{5}$$

This method is generally intuitive (i.e. since a flux is the vertical change of a quantity over time and space), but the derived values of H and LH from Bowen ratio are not always

correct, especially as LH heads toward zero. Bowen ratio may also be affected by the choice of the measurement heights. A more accurate method is the eddy covariance method. Eddy covariance is based on theories of turbulence in the atmospheric boundary layer and assumes that the vertical component of microscale eddies (turbulent deviations from the mean flow generated by irregularities in the surface and convection) are responsible for the movement of heat energy and moisture in the vertical direction (Wallace and Hobbs, 1977). Knowing the covariance between the change in temperature or moisture and vertical wind velocity yields flux. In this method, H and LH are calculated as follows:

$$H = \rho_{air} c_p \overline{w'\theta'} \quad (6)$$

$$LH = \rho_{air} L_v \overline{w'q'} \quad (7)$$

(Stull, 1988)

where $\overline{w'\theta'}$ is the covariance of turbulence fluctuations (differences from mean) of temperature and vertical velocity and $\overline{w'q'}$ is the same for the water vapor mixing ratio. Eddy covariance requires fast-response measurement of temperature, moisture and wind velocity (on the order of 10 Hz). Temperature and moisture are typically measured with a sonic anemometer and an infrared hygrometer, respectively. Wind velocity is also measured with a sonic anemometer by measuring the transit time of sound between two transducers in three orthogonal directions (Atmospheric Radiation Measurement Program, 2000). The fast-response measurements are averaged over one-half to one hour to get a flux value. Figure 5 shows the locations of Bowen ratio and eddy covariance flux measurement stations across the SGP-97 region.

Balloon soundings

Balloon soundings were launched once every three hours from several locations. The weather balloons returned information on temperature, humidity, wind velocity, pressure, location and altitude by radio to ARM-CART facilities (Atmospheric Radiation Measurement Program, 1999(1)). Three soundings were in the AVHRR region, while two of the three were also in the smaller ESTAR region. Potential temperature information was used from these soundings to calculate boundary layer height. Potential temperature is calculated as

$$\theta = T_{dry}\left(\frac{p_0}{p}\right)^{\frac{R_d}{C_p}} \tag{8}$$

(Wallace and Hobbs, 1977)

where θ is potential temperature, R_d is the universal gas constant for dry air, c_p is the specific heat of air at constant pressure, T_{dry} is dry air temperature, p is pressure at balloon altitude and p_0 is surface pressure. The potential temperature is the temperature of air if it adiabatically descended from its current pressure (p) to the reference pressure (p_o). Figure 6 shows several daytime Central Facility potential temperature soundings on July 12. The height of the boundary layer is the area of constant potential temperature, which corresponds to a well-mixed region of the atmosphere.

LIDAR

LIDAR was used in this experiment for boundary layer height validation. LIDAR is an active remote sensing technique that uses lasers to return high resolution information about the atmosphere along the direction of the laser beam. The LIDAR in SGP-97, consisting of silicon avalanche photodiode (Si:APD) lasers is LASE (Laser Atmospheric Sounding Experiment), flown on the same P-3B aircraft with ESTAR. The LIDAR was pointed downwards to provide a

series of vertical profiles of water vapor in the boundary layer along the P-3B flight track (Browell et al., 1997). LASE is a Differential Absorption Lidar (DIAL), which uses two laser wavelengths, one tuned to a wavelength strongly absorbed by the atmospheric gas species to measure, and one tuned off of the absorption wavelength. The returned reflectance from these two lasers at two different ranges can be used to measure average concentration of the gas between these ranges using the equation:

$$n = \frac{1}{2\Delta s(R_2 - R_1)} \frac{P_{on}(R_1)P_{off}(R_2)}{P_{on}(R_2)P_{off}(R_1)} \tag{9}$$

(Browell, 1989)

where P_{on} and P_{off} are the power received "on" and "off" the absorption wavelength at ranges R_1 and R_2, n is the average molecular number density between R_1 and R_2, and Δs is the difference between on-line and off-line absorption. Terms are added to this equation to account for wavelength-dependent atmospheric attenuation and extinction. LIDAR aerosol backscatter profiles (just the returns from the off wavelength) can be used to measure boundary layer height by measuring where the profiles switches from turbulent, well-mixed and higher aerosol concentrations (stirred up at the surface) to less turbulent (or laminar), stratified and lower aerosol concentrations. Finding the upper height of the mixed layer can be calculated using a wavelet analysis that detects sharp, coherent changes in the LIDAR signal of atmospheric scattering ratio (calculated from aerosol backscatter) (Davis et al., in press). Visual analysis of LIDAR images provides a good estimate of the mean height. Figure 7 shows LIDAR images from Flight 21 on July 12. The mixed layer top is the dark band within the red part of the atmospheric scattering ratio image, and range from around 1 to 1.5 km above sea level for most of the image, around 2.5 km above sea level at the southern end.

Methods

The construction of a relation between NDVI or soil moisture to fluxes requires registration of the mapped information to the point information of fluxes. Since fluxes measured from a small tower (around 3 meters) tend to have a "footprint" less than 1 km (i.e. the land surface of which a flux is representative of) (Pelgrum, 1996), the 1.1 km and .8 km resolutions of NDVI and ESTAR respectively appear a good match at least. Flux stations coordinates were projected onto the Universe Transverse Mercator (UTM) projected images of NDVI and ESTAR. Sensible heat, latent heat and Bowen ratio were compared to NDVI and soil moisture. Registration in time was another problem. NDVI values were available for morning and afternoon, though the two were not very different. ESTAR soil moisture was based on the time of the overflight, typically late morning. Flux values were available for every half hour. Flux values compared to ESTAR and AVHRR were morning and afternoon (when sensible heat was maximum) as well as various other times. The best correlations for both NDVI and ESTAR were then analyzed for their fit. Additionally, a visual correlation was attempted for NDVI by draping a 3D surface of interpolated (krigged) surface fluxes over a color corrected NDVI image. This image allowed for additional understanding of any scale dependence nature of the NDVI-flux relationship.

The best fit from the ESTAR-flux relationship was used to create a boundary layer height map. Individual inverse linear fits were made between hourly Bowen ratio and ESTAR soil moisture. This information produced a set of hourly flux maps across the ESTAR region. The calculation of boundary layer height can then be computed thermodynamically using the follow equation:

$$c_p \rho_{air} \int_{z_0}^{z_1} \left(\theta_{(t_1,z)} - \theta_{(t_0,z)} \right) dz = \int_{t_0}^{t_1} H dt \qquad (10)$$

(Stull, 1988; Wallace and Hobbs, 1977)

where c_p is the specific heat of air, ρ_{air} is the density of air averaged from z_0 to z_1, $\theta_{(t,x)}$ is potential temperature at time t and height x, z_0 is surface elevation, z_1 is predicted boundary layer height at time t1, and t_0 is the time of the sounding (early morning). This relationship basically states that the height of the convectively-mixed boundary layer is a function of the initial state of the atmosphere (from the early morning sounding) and the sum of the surface heat flux added over time. The equation is solved for z_1 (the point where $\theta_{(t1,z)}$ and $\theta_{(t1,z)}$ intersect) by iteratively increasing the value of $\theta_{(t1,z)}$ until the equation is satisfied. $\theta_{(t1,z)}$ is assumed to be constant with z. The equation is essentially a way of calculating how much energy is needed to turn an early morning stable sounding (potential temperature increase with height) to a well-mixed isothermal sounding (potential temperature constant with height under the boundary layer). Figure 8 shows a visual depiction of the equation and its solution for soundings and flux stations in the SGP region. This equation assumes that boundary layer growth is entirely a function of heating from the surface. Entrainment flux (transfer of heat energy to the free troposphere), advection (motion of heat energy in the horizontal direction) and mean vertical winds caused by synoptic-scale convergence or divergence are ignored. Potential temperature is assumed to be constant with height in a well-mixed unstable boundary layer. Despite these assumptions, the equation provides a relatively good estimate of boundary layer height. The use of this equation was validated at ARM-CART sounding and flux locations by comparing computed boundary layer height using equation 10 to actually height observed in late morning and afternoon soundings.

For this experiment, the height equation is solved for every pixel in the suite of hourly flux maps made from the ESTAR-flux relationship. To reduce computing time, pixel sums are classed in 10 w/m^2 intervals and the equation is solved only on unique integral values. The

boundary layer height values under the P-3 flight track were extracted and compared visually to LIDAR images of boundary layer structure. The visual comparison first required converting the boundary layer heights from meters above ground level to meters above sea level. This was accomplished by extracting elevation values from United State Geological Survey (USGS) 1:250,000 30 arc-second Digital Elevation Model (DEM) grids interpolated to points along the P-3 flight track. Boundary layer height was calculated for points along the P-3 flight track from 14:30 UTC to 18:30 UTC at every 3 seconds using the ESTAR boundary layer height model and interpolating the output heights over time and space to the P-3 flight track. This height was added to the interpolated DEM elevation values. The graph was then overlaid with an image of LIDAR atmospheric scattering ratio.

Results

The results from the experiment can be divided into two parts: the results of the linear regression of soil moisture and NDVI against fluxes, and the creation and validation of boundary layer height maps. Figure 9 shows the relationship between daytime NDVI and daytime sensible heat flux, latent heat flux and Bowen ratio for July 12 and 13. Figure 10 shows the same for July 12 afternoon NDVI and flux quantities taken at the time when sensible heat flux is maximum. Both relationships are only weakly correlated. Other combinations and multiple linear and non-linear regressions of AVHRR measurements and fluxes were compared and similar weak correlations held. This might suggest that for a grasslands area like the Southern Great Plains, vegetation cover does not greatly affect the variability of surface energy flux across the landscape, at least relative late in the growing season, when much of the winter wheat is already harvested, and the remaining vegetation may have senesced as a result of hot, dry conditions.

The small variability in NDVI between the measured flux points suggests that vegetation variability is too small to be the primary influence on flux variability. However, the lack of correlation may also be a function of sensor spectral and spatial resolution. MacPherson (1998) noted that in comparing the Twin Otter Bowen Ratio data to greenness, the relationship between the ratio and vegetation was not as strong in some areas (El Reno) as other flight tracks (Figure 1). MacPherson hypothesizes that different agricultural practices under the different tracks may have an impact on this relationship. The lack of a relation between NDVI and the flux towers compared to the good relation between Twin Otter fluxes and greenness may be because the aircraft flux footprint is much larger than the tower.

There is evidence for a scale dependence in NDVI. Figure 11 shows the visual gridded and krigged 3D surface of fluxes against NDVI colors. These graphs suggest that a visual correlation exists between broad patterns of flux heterogeneity and vegetation, contrary to the poor point-by-point correlation. Perhaps, vegetation interacts with flux variability at a different scale then the 1.1 km scale of AVHRR or the 1 km scale of the tower flux measurements.

Figure 12 shows the relationship of ESTAR soil moisture to midday surface sensible heat flux, latent heat flux and Bowen ratio for July 12 and 13. The correlation is much stronger, especially to Bowen ratio. This correlation suggests that surface soil moisture, a measure of water vapor availability, has a strong causal relationship with Bowen ratio, a measure of surface energy partitioning. The correlation to Bowen ratio is stronger on July 13 than July 12. This may be because the typical soil moisture values are lower on July 13 than July 12, and on both days, the linear regression appears to fit better with those points that have drier soil moisture values than those that are wetter. The wet soil moisture values are quite high, where greater than

45% soil moisture is essentially a saturated soil. This suggests that the ESTAR soil moisture-Bowen ratio relationship breaks down or flattens out with very high soil moistures.

A similar fit between ESTAR and midday fluxes was done for ESTAR soil moisture and fluxes at regular intervals across the day, with an additional inverse parameter added to flatten out Bowen ratio modeled values in response to high soil moistures and to avoid negative Bowen ratios with very low soil moisture. The fits were all generally good, though some hours had worse fits than others, and there appears to be a regular pattern of change in the slopes and correlation coefficients with time. Table 1 shows the parameters of the fit for each hour. This table of fits may warrant future exploration of patterns of Bowen ratio dependence on soil moisture over time.

The fits were used to produce maps of Bowen ratio across the ESTAR region. To be able to map latent and sensible heat flux, maps of net radiation (R) and soil heat flux (G) were also needed. On a clear day, R is relatively constant across the Southern Great Plains region, with major exceptions being variations in upwelling solar and IR radiation. Net radiation values from eddy covariance and Bowen ratio systems were gridded using krigging across the ESTAR region. G is relatively small compared to R, H, and LH, and was krigged and gridded across the region in a similar method to R. Overall results of the fit between predicted and actual Bowen ratios were only marginally poorer by using constant values for R or G, or zero for G. H and LH for each pixel at each hour can be known given R, G, and the Bowen ratio for each pixel using equations 4 and 5. Figures 13 shows how well these fluxes relate to measured surface energy fluxes for July 12 and 13. Note the better fit on July 13 than 12. It is also interesting note that on both days, the fit to latent heat flux is better than sensible heat flux.

Boundary layer height is calculated for each pixel using the method described by equation 10, the sensible heat flux derived from ESTAR and an early morning sounding from the central facility, which is near the center of the ESTAR image. Figure 14 shows boundary layer height derived from this method for the ESTAR region for July 12 and 13. The July 12 ESTAR derived growth-model boundary layer height image has a strong south to north gradient with higher heights in the south. Predicted boundary layer heights on July 13 are higher overall than July 12 (since it was drier), but there is greater variability in heights, especially in the south. Observed July 13 depths are more shallow than on the July 12. It is important to note that a strong southerly wind component of about 10 m/s would have induced a good deal of advection into the scene, and so the boundary layer height may be more "blended" than the figures suggest. The figures are more a map of "potential" boundary layer heights that assume no entrainment, no advection and only direct surface effects on mixed layer height.

Figure 15 shows the overlay of July 12 boundary layer height against LIDAR scattering ratio from 14:30 to 18:30. Interestingly, ESTAR-derived growth-model predicted boundary layer height does not seem to mimic the same overall structure as what LIDAR shows. Although the general trend (high south, low north) is seen in the larger ESTAR boundary layer height maps, along the P-3B track boundary layer height appears to be relatively constant, though with quite a bit of microvariability. Overall, the average boundary layer height from the ESTAR-derived growth-model matches well with the average boundary layer height from LIDAR. Also, both the growth-model and LIDAR show similar intervals of microvariability in boundary layer height structure. Further research is needed to know if this variability is the same in LIDAR and ESTAR.

Discussion

Remotely sensed soil moisture and vegetation cover were compared to points of known surface energy flux. The best fit was found between ESTAR 800m resolution passive-microwave derived surface soil moisture and mid-day Bowen ratio. This fit was recomputed for hourly Bowen ratios using an inverse fit to account for very high and very low soil moistures. These sets of fits were then used to map Bowen ratio across the entire image. Bowen ratio, along with interpolated net incoming radiation and soil heat flux, were used to compute surface sensible heat flux and latent heat flux over space and time. These fluxes were then used to estimate boundary layer height using a simple thermodynamic growth model. When compared to LIDAR estimated boundary layer heights, the calculated heights tended to differ especially to the south, though the average height is the same. Agreement between LIDAR and the growth model for July 13 is worse than for July 12, despite a between fit between the fluxes and soil moisture on July 13. This recap highlights several questions: 1.) despite the theory and prior evidence, why didn't NDVI correlate to fluxes?, 2.) what are the possible causes for the soil moisture to Bowen ratio relation, and 3.) how can the boundary layer height model be improved?

As previously mentioned, the lack of NDVI correlation with Bowen ratio may be related to issues surrounding spectral and spatial resolution of AVHRR NDVI and the flux footprint of the tower. NDVI varies very little across the Southern Great Plains and so one can conclude that either 1.) NDVI variability is too small to have an effect on flux variability in an homogenous area like the Southern Great Plains or 2.) NDVI does not accurately measure vegetation variability in the Southern Great Plains. However, these may be refuted by noticing that the images of NDVI in figure 3 do show a general gradient from west to east. Western areas are less vegetated than the eastern part of the image. The visual 3D flux surface maps of figure 11 show

the relation of higher Bowen ratios in areas of lower NDVI. The lack of a direct correlation with the flux stations instead points to a relatively low number of flux stations that were all located in areas of similar NDVI.

Another argument that is worth looking further into is the idea that NDVI interacts with fluxes on a different scale than 1 km. The 3D surfaces do show that despite the general trend (green east, brown west yields low Bowen ratio east, high Bowen ratio west), there is quite a bit of variability in NDVI at individual locations. The specific NDVI of a pixel collocated with a flux tower may not represent the overall vegetation cover of the area around the tower. Indeed, larger agglomerations of pixels (3x3) were considered, but the results were essentially the same as using one pixel. A further study might try using different radii of NDVI pixels and seeing if the fit between fluxes and NDVI changes with resolution. Friedl (1997) showed that a relationship between NDVI and modeled Leaf Area Index (LAI) and fraction of absorbed photosynthetically active radiation (FPAR) in the Kansas tallgrass prairie had different behaviors depending on the resolution scale of NDVI. It may be the case the alternate case that the 1100 m resolution of AVHRR is too coarse to reflect the true field-to-field variability of the tower flux measurements. However, getting infrared and red spectral information at higher resolution decreases the revisit time. For the Southern Great Plains experiment, the next higher resolution satellite image does not occur until July 25, 1997 when LandSat 6 Thematic Mapper flew over the area. By that time, vegetation cover may have changed due to agricultural plowing of winter wheat occurring in the area at the time. As previously described, MacPherson (1997) notes that plowing and variability of precipitation may have affected the Twin Otter Bowen ratio-greenness relation.

Given the poor results of the vegetation correlation, what makes the soil moisture correlation so strong? It can be argued that soil moisture is a more direct measure of atmospheric moisture available, and thus related to the amount of partitioning of fluxes between latent and sensible heat (Browell et al., 1997). This may especially be the case in a hot and dry month like July 1997. Our results suggest that evaporation may dominate in July, given the strong correlation between surface soil moisture and Bowen ratio. A reasonable test would be to see if flux tower Bowen ratio and surface soil moisture correlates over longer time periods.

The ESTAR derived surface soil moisture is based on a model. The model is based on data from point-based soil measurements from Oklahoma Mesonet and LandSat land cover maps. The output soil moistures do correlate very well with the soil moistures test sites, but that does not necessarily mean that soil moisture across the rest of the region is as correct. The mapping relies on data from GIS databases of soil type along with LandSat derived vegetation cover from July 25, by which time vegetation may be different from July 12. The original output product from the passive microwave sensor is surface brightness temperature footprints (Jackson et al., 1999). It may be more the case that brightness temperature is related to Bowen ratio, however this experiment cannot answer that. There is good enough cause to think that soil moisture and Bowen ratio are connected. It is interesting, as previously pointed out, to note the different correlation on July 12 and 13, and the better correlation with lower soil moistures than high soil moistures. This suggests that soil moisture affects Bowen ratio only to a point. Once the ground becomes extremely wet, other factors must be involved to induce variability in fluxes among those areas.

Another interesting point is the difference in where the main gradient of variability lies. NDVI shows a strong west to east gradient as is expected with the change in vegetation with

elevation. ESTAR, however, shows a primarily north-south soil moisture gradient (wet north, dry south). This is partly because the east-west field of view of ESTAR is 1/3 the size of the NDVI field of view, but it is also a function of the location of the rain storm that passed on July 10. Consequently, the Bowen ratio map from ESTAR shows high Bowen ratios in the north and low in the south, leading to boundary layer heights higher in the south. Any similar east-west gradient is too small to discern from the ESTAR scene. The north-south gradient is consistent with July 12 and July 13 LASE observations of boundary layer height.

Finally, the boundary layer height map needs to model advection, entrainment and possibly vertical divergence. Advection, however, is difficult to model without some information about upwind conditions. At a 10 m/s southerly wind, within one hour air located in one area would have moved 36 km north. During this time, its temperature and moisture characteristics would have been affected by a number of different surfaces. In this experiment, modeling advection would require a parcel of air located over a specific pixel to be subject to the surface flux conditions of 45 pixels in the downwind direction. The net effect of advection would be homogenization of surface flux variability in the direction of the wind. Thus, instead of using this downwind pixel propagation method which effectively limits the number of pixels for which one could get boundary layer height information from, it may be possible to estimate the effect of advection by averaging the flux response to a 45x45 pixel grid and calculating boundary layer height from that. However, the results from LIDAR show that the variability of boundary layer height is at a much finer scale than 36 km. Still, this may be the result of convective turbulence, which will exist even over a perfectly homogenous surface (Deardorff et al., 1980). A larger path or field of view would allow for modeling advection in the pixel propagation method and still get a large enough scene from which to discern boundary layer

height variability. Modeling entrainment is easier and requires some knowledge of the amount that boundary layer growth is affected by the mixing of boundary layer air with free troposphere air. Since this amount tends to be relatively small compared to surface effects, a scene average may be used, and it may help produce more reliable boundary layer height numbers. The high July 13 mean boundary layer height may be a function of synoptic subsidence (average vertical wind velocity < 0) from a high pressure system. This value could be estimated with model reanalyses of wind profiles from the region using balloons and radar profilers.

Conclusion

One of the goals of the Southern Great Plains experiment was to validate soil moisture retrieval algorithms of ESTAR for possible future use of passive microwave remote sensing in space. Recent Earth observing satellites such as TERRA have not included a passive microwave sensor system, though plans are in the works for a yet-to-be launched passive microwave system. Spatial resolution in space at ESTAR frequencies would be relatively coarse, on the order of 10-30 km (Jackson et al., 1999). While this may still be useful for creating global soil moisture datasets, it would complicate the global remote sensing of mesoscale surface energy flux variability.

This experiment showed that for the Southern Great Plains experiment, on a clear, cloudless day in July, soil moisture derived from a model that relies on ESTAR passive-microwave brightness temperature and soil temperature has a relatively good correlation to surface energy flux partitioning, i.e. the Bowen ratio. This correlation is greater than a similar experiment done with vegetation cover as measured with a standard spectral vegetation index from AVHRR. This suggests that surface energy flux variability in the Southern Great Plains is

driven more by moisture availability than vegetation cover and type. This experiment did not set out to construct a universal relationship between fluxes and remotely sensed vegetation or soil moisture for use on all land types. The ESTAR-surface energy flux correlation may hold only for grasslands, or perhaps only for the Southern Great Plains. It may be the case that in other landscapes with more vegetation and higher vegetation variability, or at different scales or different days, vegetation indices are better than soil moisture at estimating surface energy fluxes (Brutsaert et al., 1993).

Sensible heat mapped from the soil moisture relationship was used to drive a thermodynamic boundary layer height model. The model, while not matching boundary layer height at all points, possibly due to the lack of advection and entrainment modeling, did match over a larger scale the height gradient of the boundary layer. The broad, though not specific, similarity between LIDAR and ESTAR-derived growth model strengthens the hypothesis of soil moisture control of surface energy flux variability. Further research might 1.) compare in greater depth the difference between the east-west vegetation gradient and the north-south soil moisture gradient, 2.) test the scale dependence hypothesis on the NDVI-flux relationship, and 3.) perform the modeling of advection in the boundary layer height model.

Acknowledgements

This research would have been impossible without the ideas, support and encouragement from Dr. Ken Davis in the Department of Soil, Water and Climate at the University of Minnesota, Twin Cities, MN. This experiment was also aided with the help of Dr. Ron Dobosy at the NOAA Atmospheric Turbulence and Diffusion Division, Dr. Christoph Senff at CIRES, University of Colorado, and the staff of the LASE project at NASA Langley Research Center. This work was funded in part by a grant from NASA's Land Surface Hydrology Program, grant number NASA/NAG-1-193.

Works Cited

Anderson, C., J.M. Norman, G.R. Diak, W.P. Kustas, and J.R. Mecikalski. 1997. A Two-Source Time-Integrated Model for Estimating Surface Fluxes Using Thermal Infrared Remote Sensing. *Remote Sensing of the Environment*, v. 60: 195-216.

Atmospheric Radiation Measurement Program. 1999 (1). *Balloon-Borne Sounding System (BBSS)*. Department of Energy, http://www.arm.gov/docs/instruments/static/bbss.html

Atmospheric Radiation Measurement Program. 1999 (2). *Energy Balance Bowen Ratio (EBBR)*. Department of Energy, http://www.arm.gov/docs/instruments/ebbr.html

Atmospheric Radiation Measurement Program. 2000. *Eddy Correlation Flux Measurement System (ECOR)*. Department of Energy, http://www.arm.gov.instruments/static/ecor.html

Avery, Thomas E. and Graydon L. Berlin. 1992. *Fundamentals of Remote Sensing and Airphoto Interpretation.* Upper Saddle River, NJ: Prentice-Hall, Inc.

Browell, Edward V. 1989. Differential Absorption Lidar Sensing of Ozone. *Proceedings of the IEEE*, v. 77, n. 3, 419-432.

Browell, E. V., S. Ismail, W. M. Hall, A. S. Moore, S. A. Kooi, V. G. Brackett, M. B. Clayton, J. D. W. Barrick, F. J. Schmidlin, N. S. Higdon, S. H. Melfi, and D. Whiteman, 1996. LASE: validation experiment and atmospheric case studies. *18th International LASE Radar Conference Abstract Book, July 22-26*. Berlin, Germany, 32.

Browell, Edward V., Syed Ismail, Kenneth J. Davis, and Donald H. Lenschow. 1997. *Investigation of Mesoscale Variability in Convective Boundary Layer Development Using LASE*. NASA Langley Research Center.

Brutsaert, Wilfried, A. Y. Hsu, and Thomas J. Schmugge. 1993. Parameterization of Surface Heat Fluxes above Forest with Satellite Thermal Sensing and Boundary-Layer Soundings. *Journal of Applied Meteorology,* v. 32: 909-917.

Chehbouni, A., D. Lo Seen, E.G. Njoku, J.P. Lhomme, B. Monteny, and Y.H. Kerr. 1997. Estimation of sensible heat flux over sparsely vegetated surfaces. *Journal of Hydrology*, v. 188-189: 855-868.

Davis, Kenneth J., N. Gamage, C. R. Hagelberg, D. H. Lenschow, P. P. Sullivan and C. Kiemle. An objective method for determining atmospheric structure from airborne LIDAR observations. *Journal of Oceanic and Atmospheric Technology*, in press.

Deardorff, J. W., G. E. Willis, and D. K. Lilly. 1980. Laboratory studies of the entrainment zone of a convectively mixed layer. *Journal of Fluid Mechanics*, v. 100, part 1, 41-64.

Friedl, Mark A. 1997. Examining the Effects of Sensor Resolution and Sub-Pixel Heterogeneity on Spectral Vegetation Indices: Implications for Biophysical Modeling, in Dale A. Quattrochi and Michael F. Goodchild, eds. *Scale in Remote Sensing and GIS*. Boca Raton, FL: Lewis Publishers.

Gillies, Robert R. and Toby N. Carlson. 1995. Thermal Remote Sensing of Surface Soil Water Content with Partial Vegetation Cover for Incorporation into Climate Models. *Journal of Applied Meteorology*, v. 34: 745-756.

Greenland, David. 1994. Use of satellite-based sensing in land surface climatology. *Progress in Physical Geography*, v. 18, n. 1: 1-15.

Hydrology Data Support Team. 2000. *Surface Flux Data From The Southern Great Plains Hydrology Experiment*. NASA Goddard, http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/SGP97/srf_flx.html

Jackson, Thomas. 1997. *Southern Great Plains 1997 (SGP97) Hydrology Experiment Plan*. USDA Agricultural Research Service. http://hydrolab.arsusda.gov/sgp97/explan/

Jackson, Thomas J., David M. Le Vine, Ann Y. Hsu, Anna Oldak, Patrick J. Starks, Calvin T. Swift, John D. Isham, and Michael Haken. 1999. Soil Moisture Mapping at Regional Scales Using Microwave Radiometry: The Southern Great Plains Hydrology Experiment. *IEEE Transactions on Geoscience and Remote Sensing*, v. 37, n. 5: 2136-2151.

MacPherson, J.L. 1998. *NRC Twin Otter Operations in the 1997 Southern Great Plains Experiment*. National Research Council Canada, Institute for Aerospace Research. LTR-FR-146.

Pelgrum, H. and W.G.M. Bastiaanssen. 1996. An intercomparison of techniques to determine the area-averaged latent heat flux from individual in situ observations: A remote sensing approach using the European Field Experiment in a Desertification-Threatened Area data. *Water Resources Research*, v. 32, n. 9: 2775-2786.

Stull, Roland B. 1988. *An Introduction to Boundary Layer Meteorology*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Stull, Roland B. 1995. *Meteorology Today for Scientists and Engineers*. St. Paul, MN: West Publishing Company.

Wallace, John M. and Peter V. Hobbs. 1977. *Atmospheric Science: An Introductory Survey*. London: Academic Press, Ltd.

List of Figures and Tables

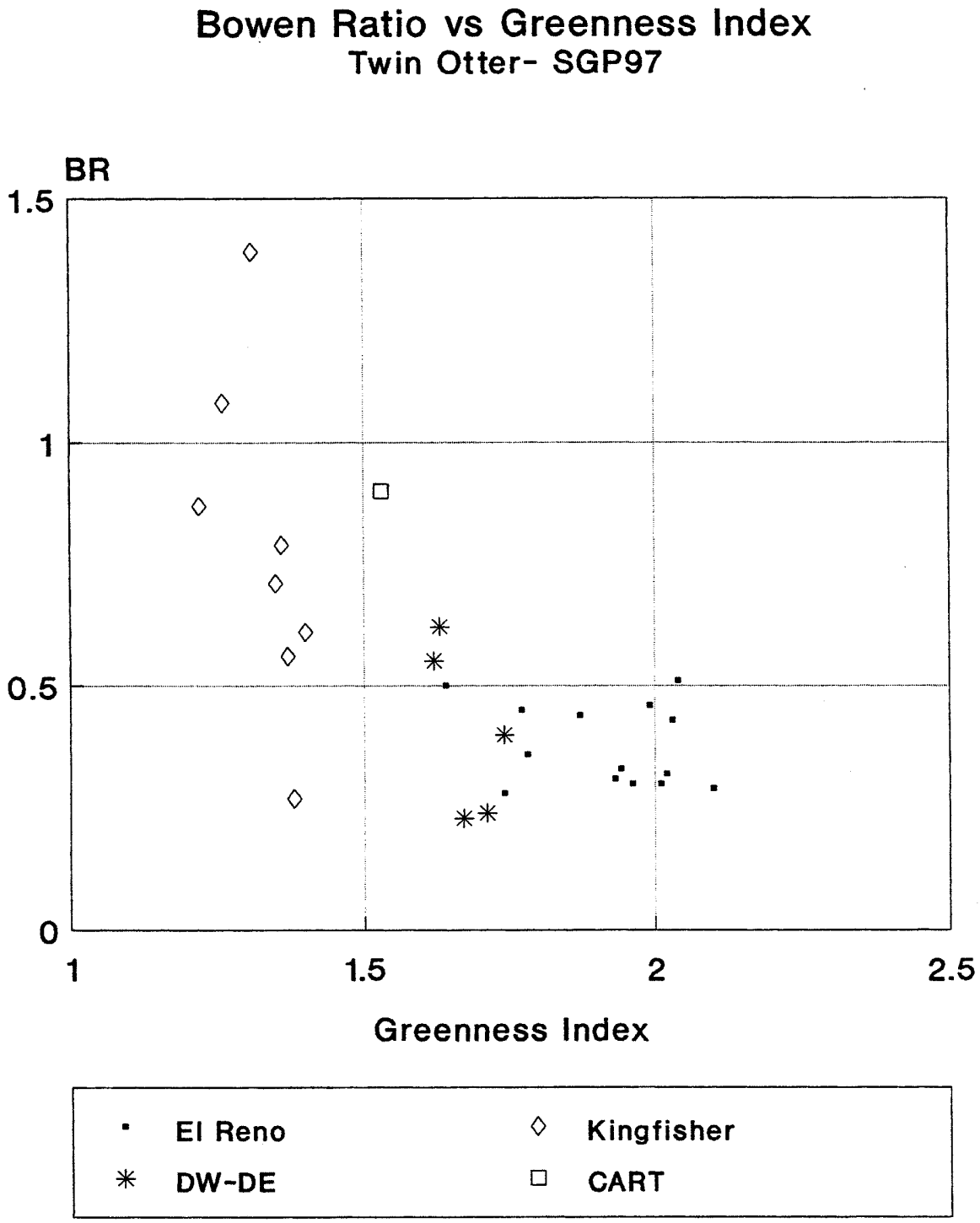


Figure 1. Bowen ratio vs. Greenness Index as measured by Twin Otter aircraft (from MacPherson, 1998)

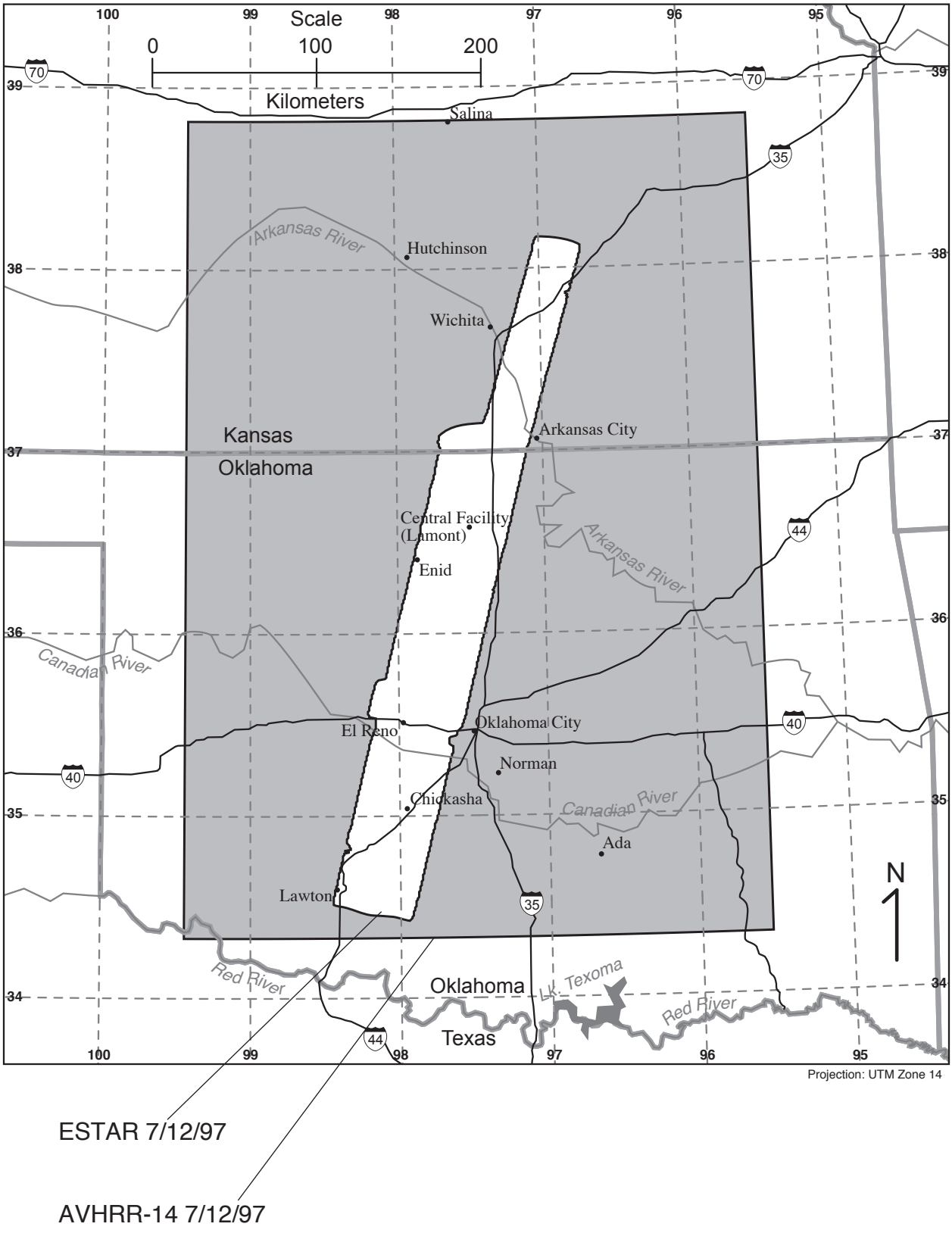
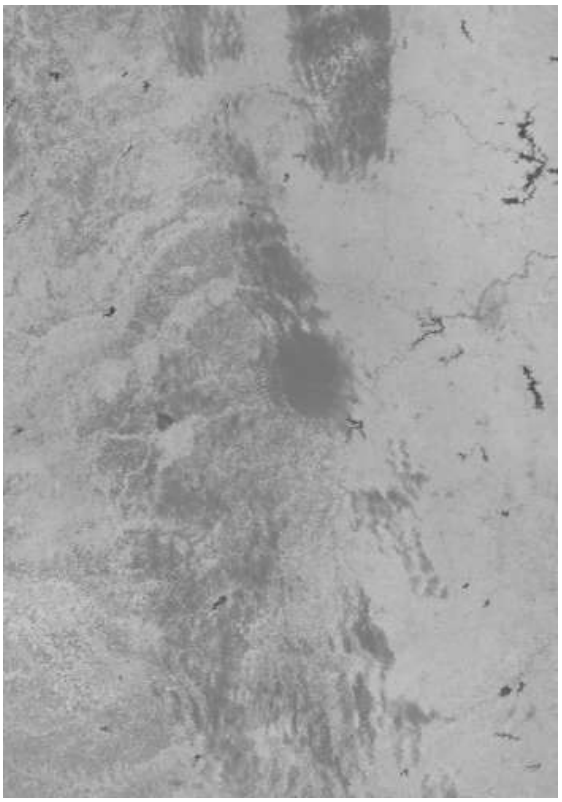
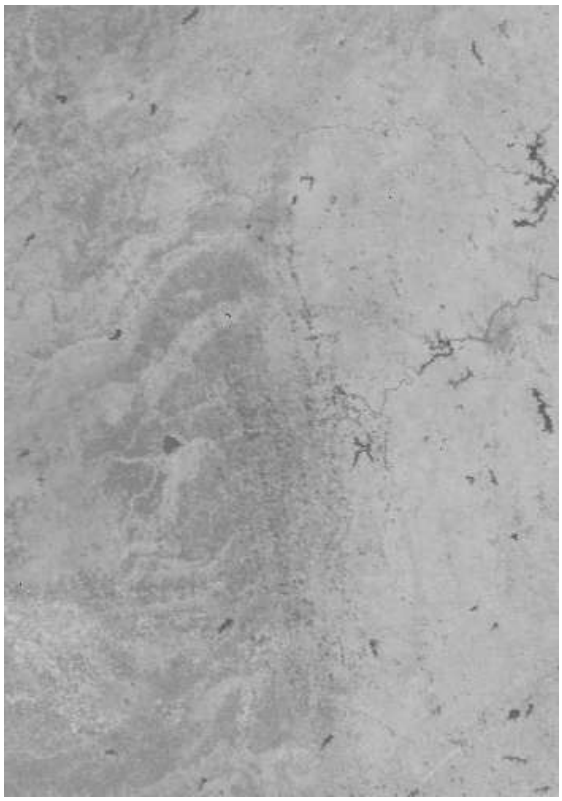


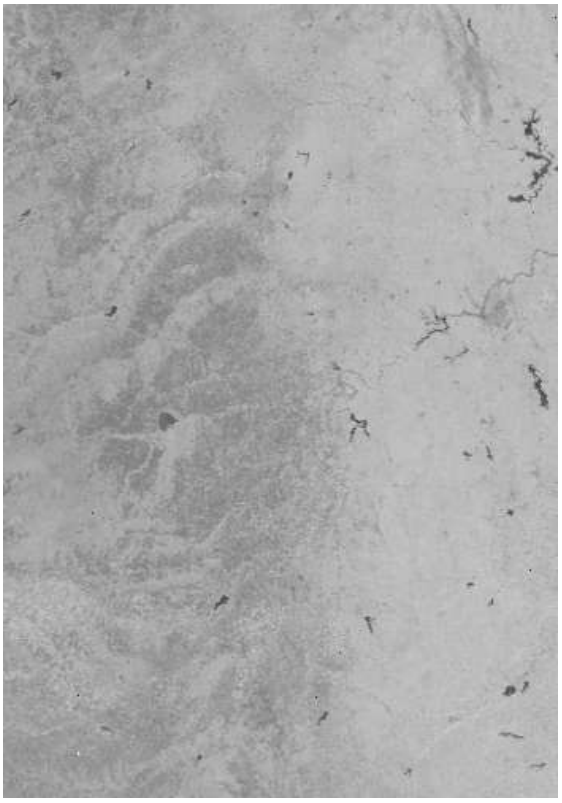
Figure 2. Map of SGP-97 experiment area and NDVI and ESTAR coverage.



(a)



(b)

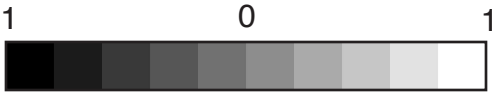


(c)

a) NOAA 12: 7/12/97 13:00 UTC

b) NOAA 14 7/12/97 20:00 UTC

c) NOAA 12 7/13/97 13:00 UTC



NDVI Scale

Figure 3. NOAA AVHRR satellite Normalized Difference Vegetation Index (NDVI) for 7/12 and 7/13 over the SGP-97 area (see figure 2 for location). Local time = UTC - 5

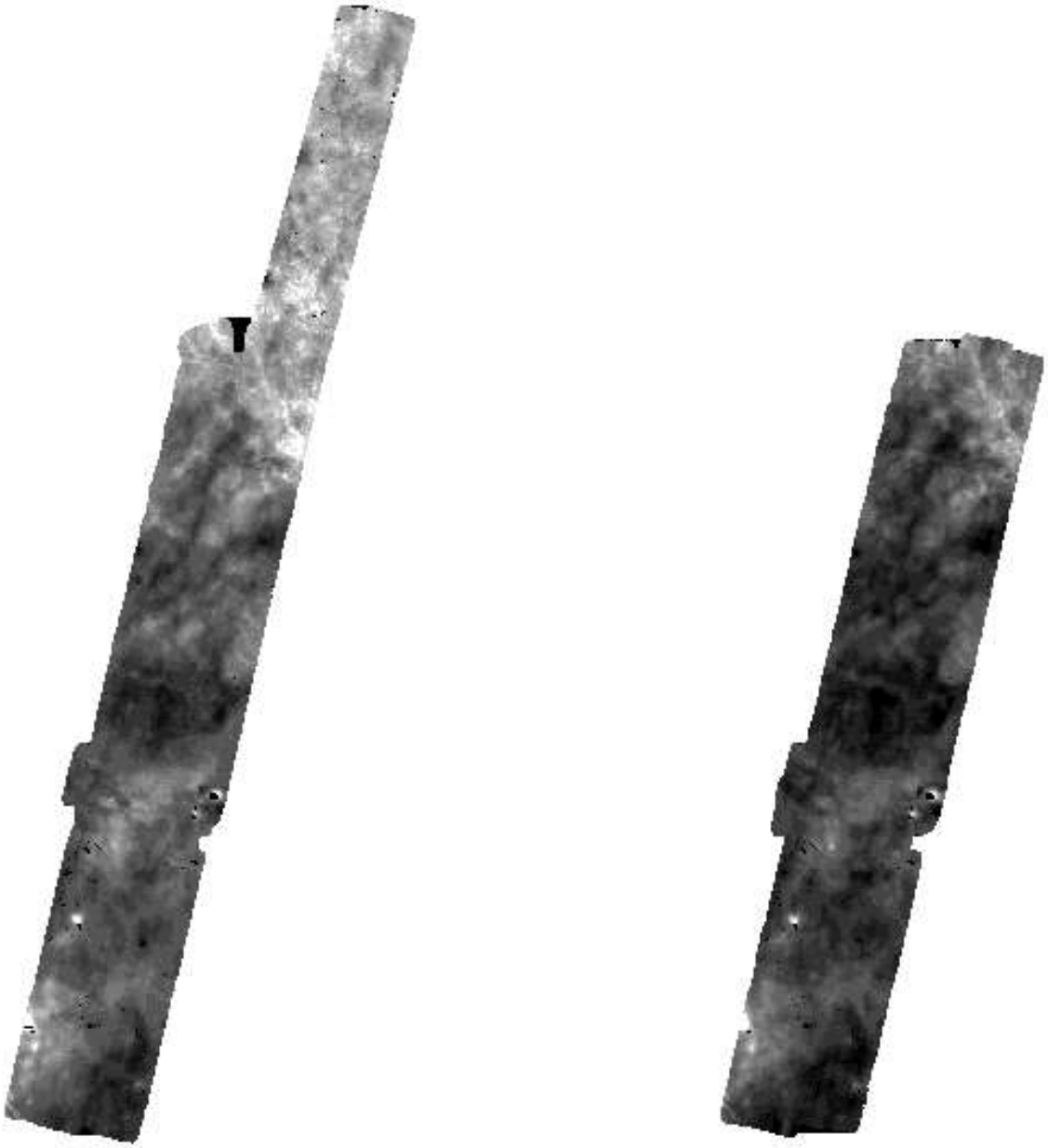


Figure 4. ESTAR derived soil moisture for SGP-97 area. See figure 2 for location.

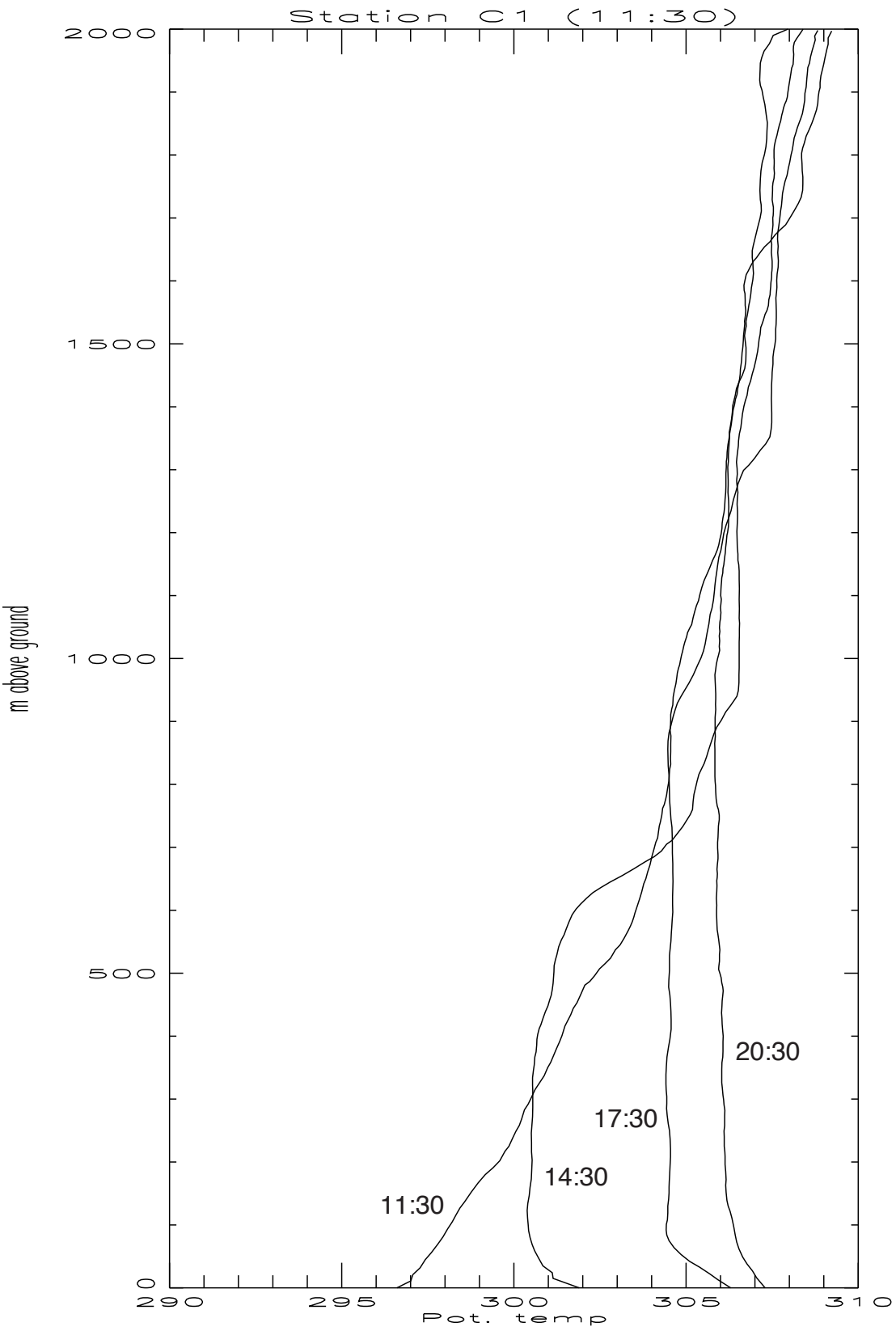
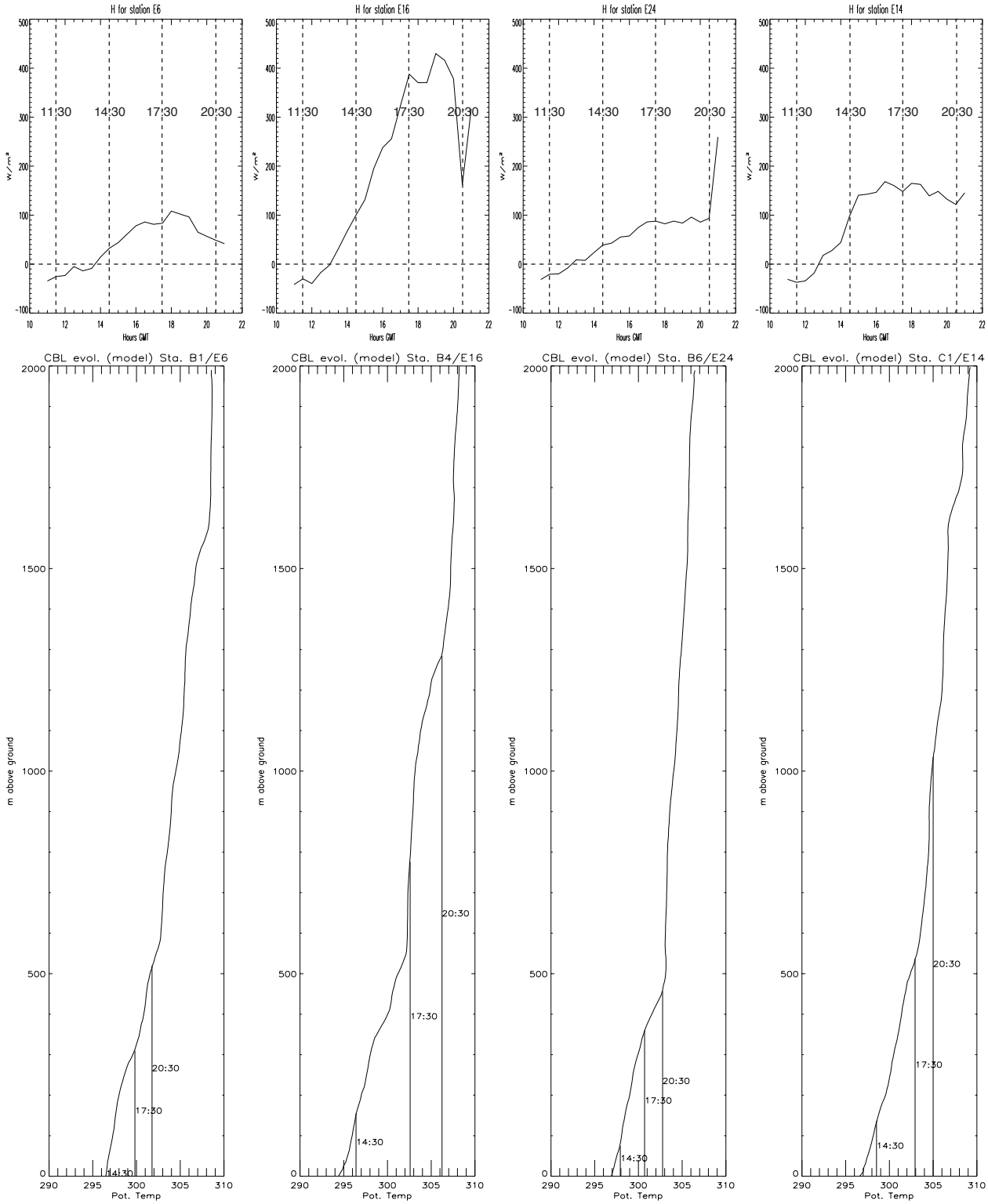


Figure 6. Central facility sounding on 7/12/97 from 11:30 to 20:30 UTC (local time = UTC - 5). Note growth in area of constant potential temperature with height over time.



$$c_p \rho_{air} \int_{z_0}^{z_1} \left(\theta_{(t_1,z)} - \theta_{(t_0,z)} \right) dz = \int_{t_0}^{t_1} H dt$$

Figure 8. Thermodynamic boundary layer height model, uses sensible heating over time (top graph: four flux stations), early morning sounding (bottom graph: 11:30 UTC, four nearest stations) and iterative solution to equation shown above.

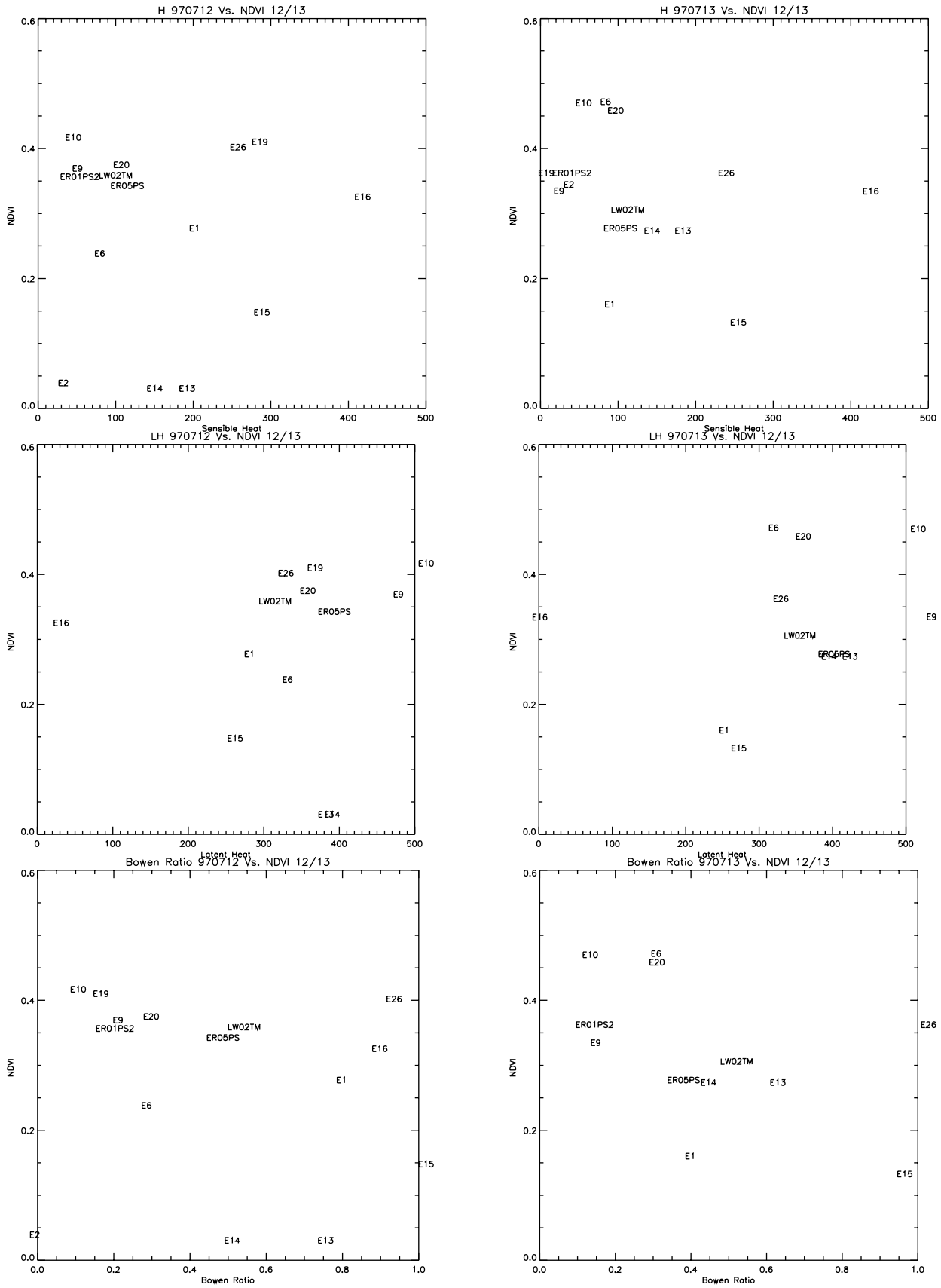


Figure 9. Morning (13:00 UTC) NDVI vs. morning flux station measured sensible heat flux (top), latent heat flux (middle) and Bowen ratio (bottom) for 7/12/97 (left) and 7/13/97 (right). See figure 5 for flux station locations.

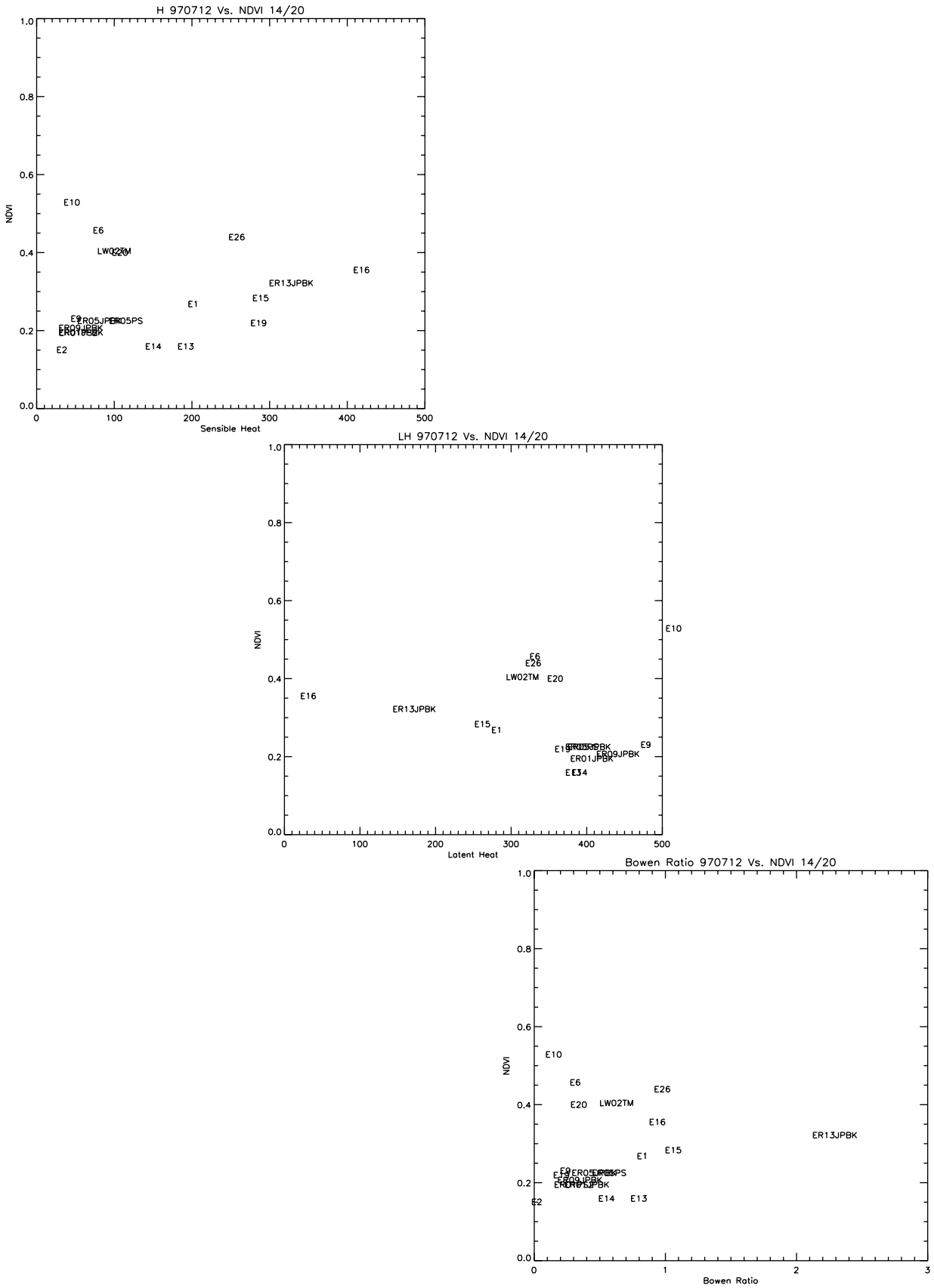


Figure 10. Afternoon (20:00 UTC) NDVI for 7/12/97 and flux station sensible heat (top), latent heat (middle) and Bowen ratio (bottom). See figure 5 for flux station locations.

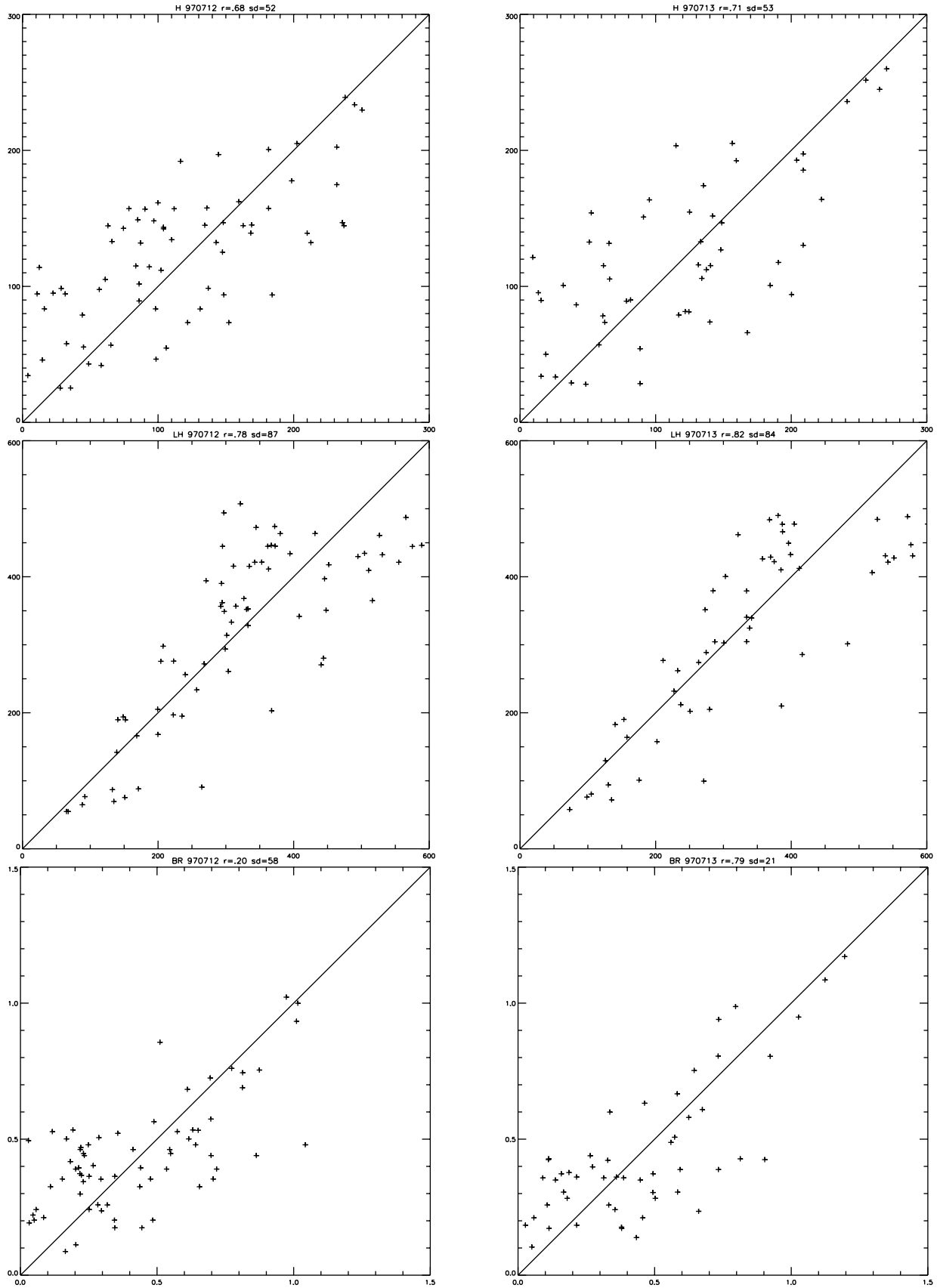


Figure 13. Comparison of ESTAR derived fluxes (y-axis) to station measured fluxes (x-axis) at all times during the day, with 1:1 line. Top is sensible heat flux, middle is latent heat flux, bottom is Bowen ratio. Left is 7/12/97 and right is 7/13/97.

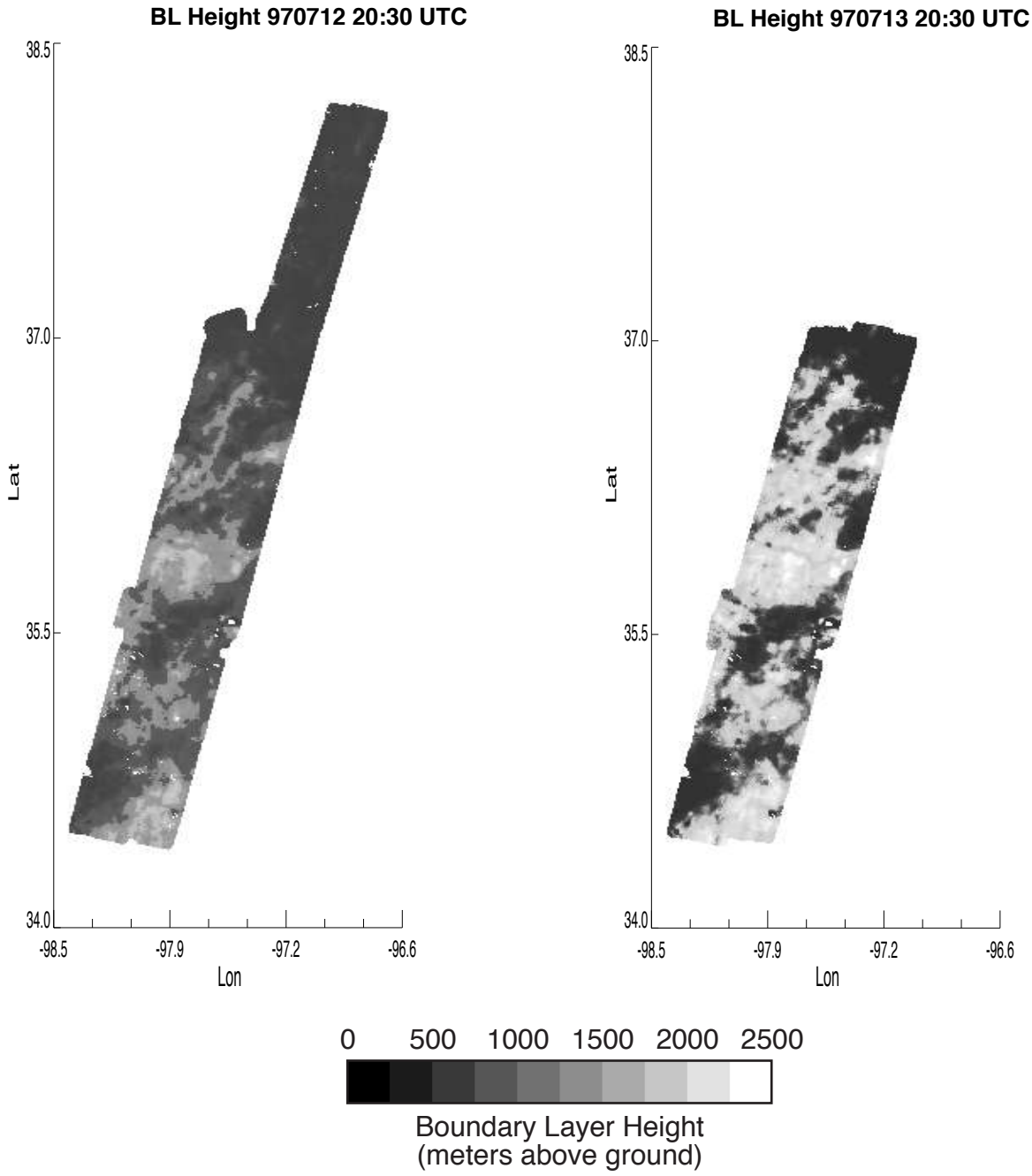


Figure 14. 20:30 UTC boundary layer height based on 14:30 UTC Central Facility sounding, ESTAR derived sensible heat flux and thermodynamic model. Left is 7/12/97, right is 7/13/97 height.

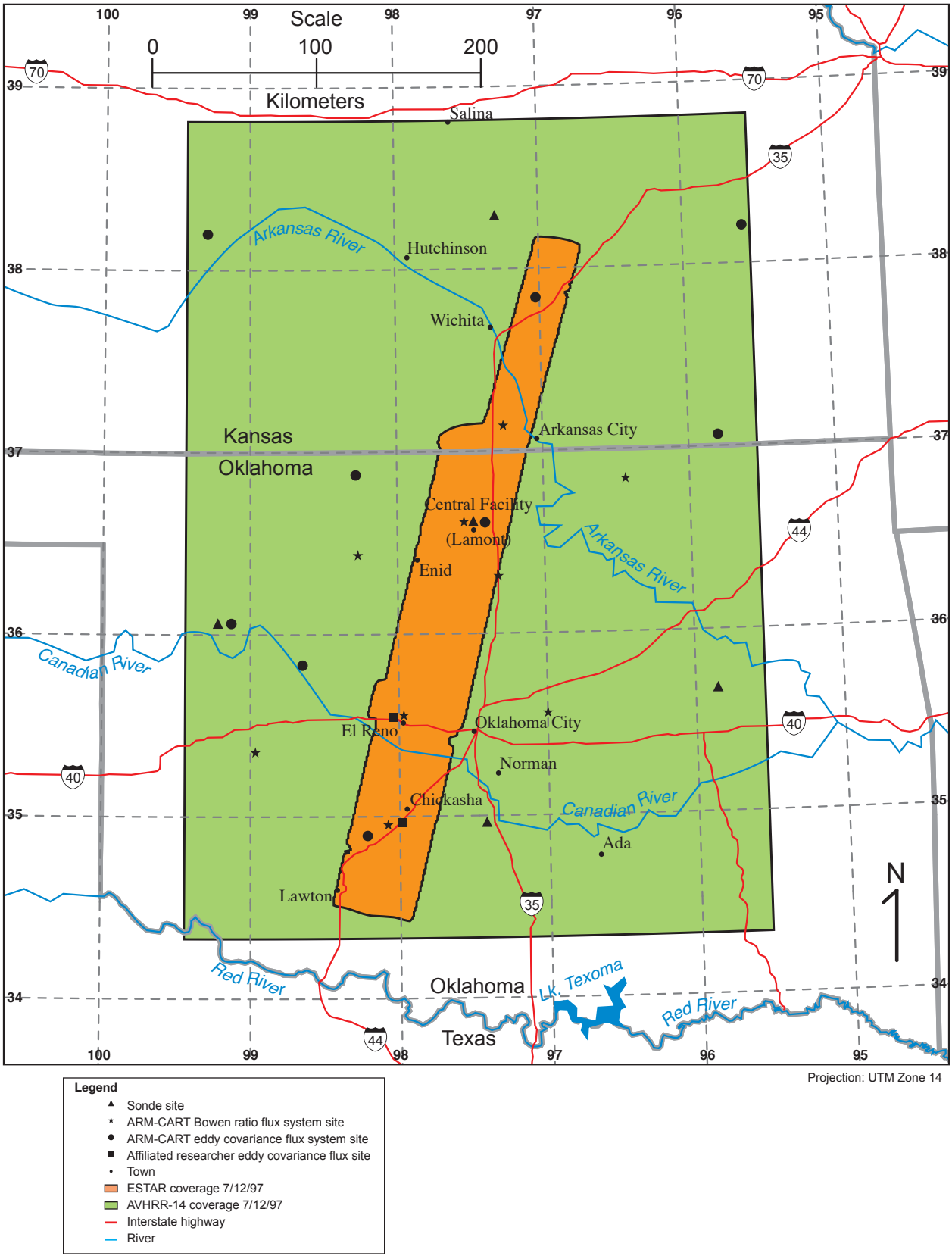


Figure 5. Location of flux towers and sounding sites.

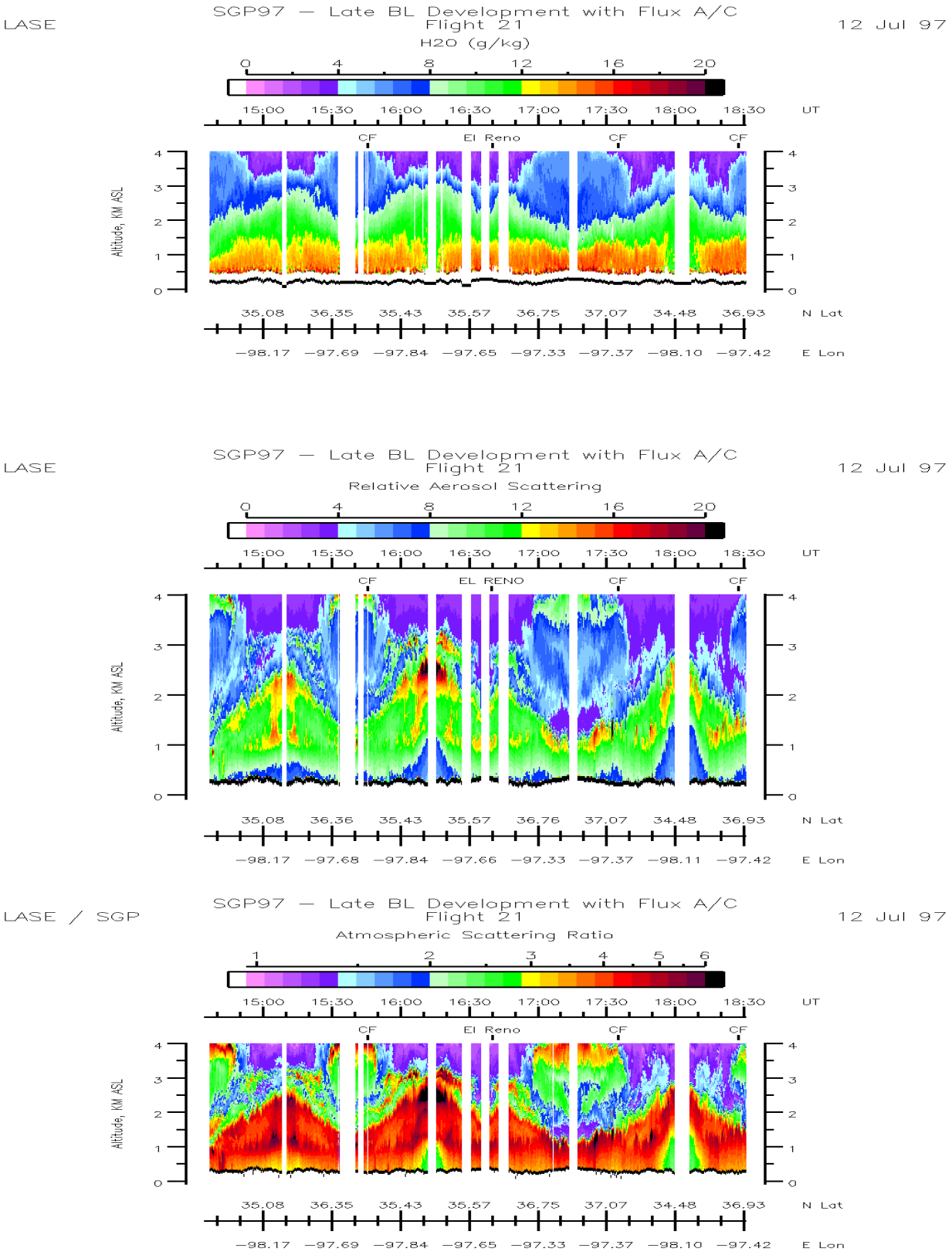


Figure 7. Images from 7/12/97 LIDAR runs. Top shows water vapor, middle is relative aerosol backscatter, and bottom is atmospheric scattering ratio. Dark line within red portion of bottom graph is boundary layer height.

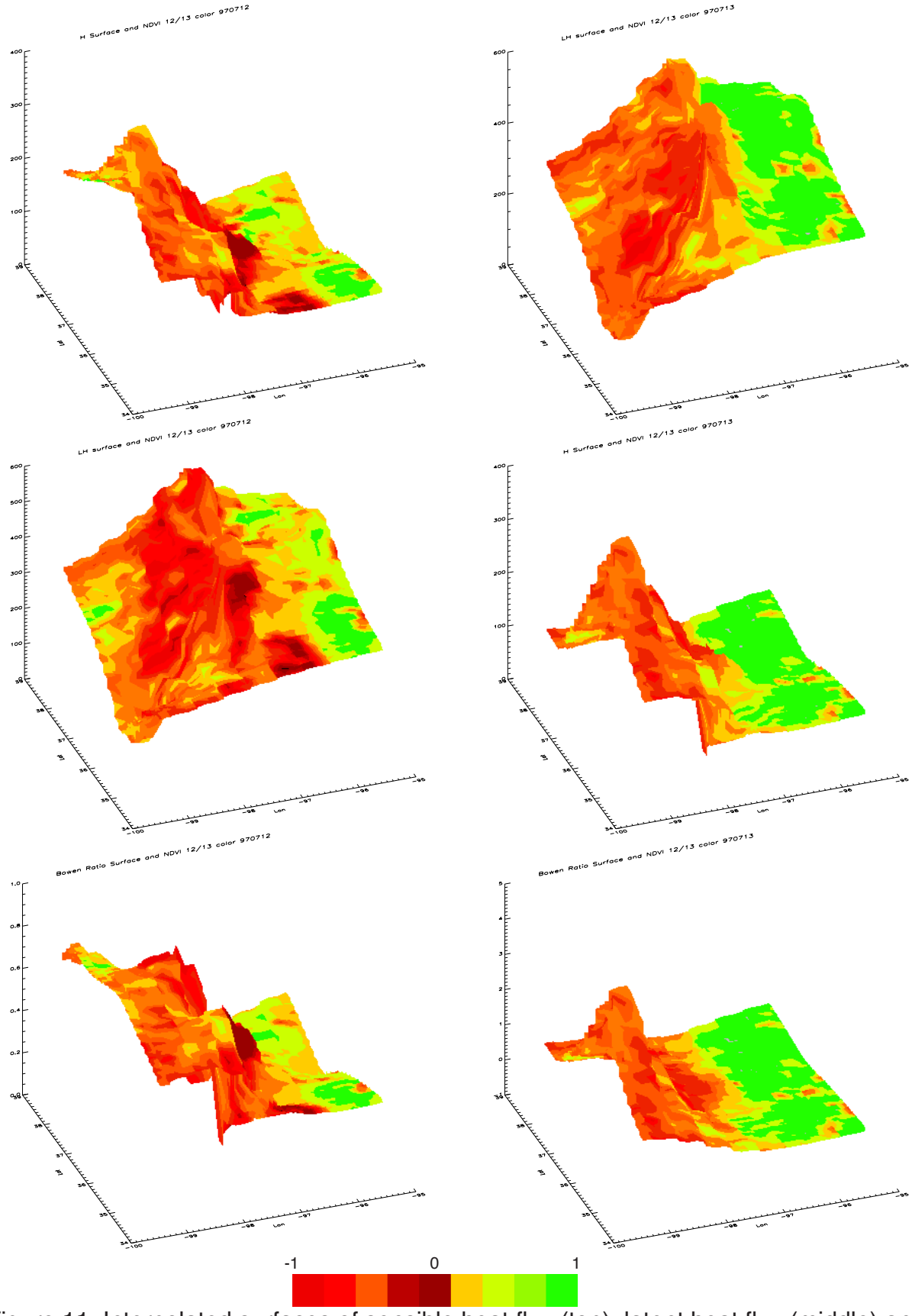


Figure 11. Interpolated surfaces of sensible heat flux (top), latent heat flux (middle) and Bowen ratio (bottom) overlaid with 7/12 (left) and 7/13 (right) NDVI colors - red (0) - green (1).

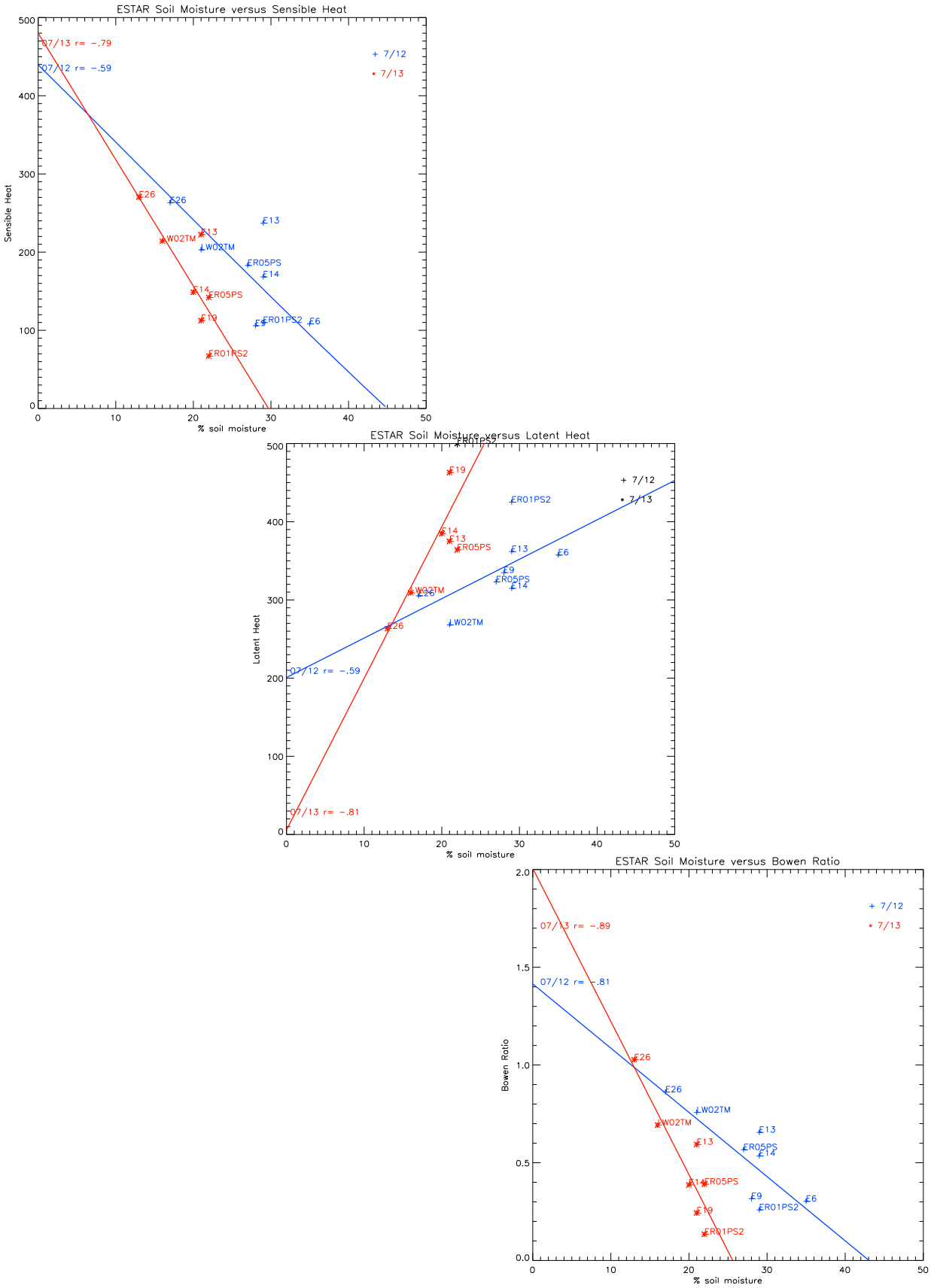


Figure 12. Relationship of ESTAR derived soil moisture and flux station sensible heat flux (top), latent heat flux (middle) and Bowen ratio (bottom)

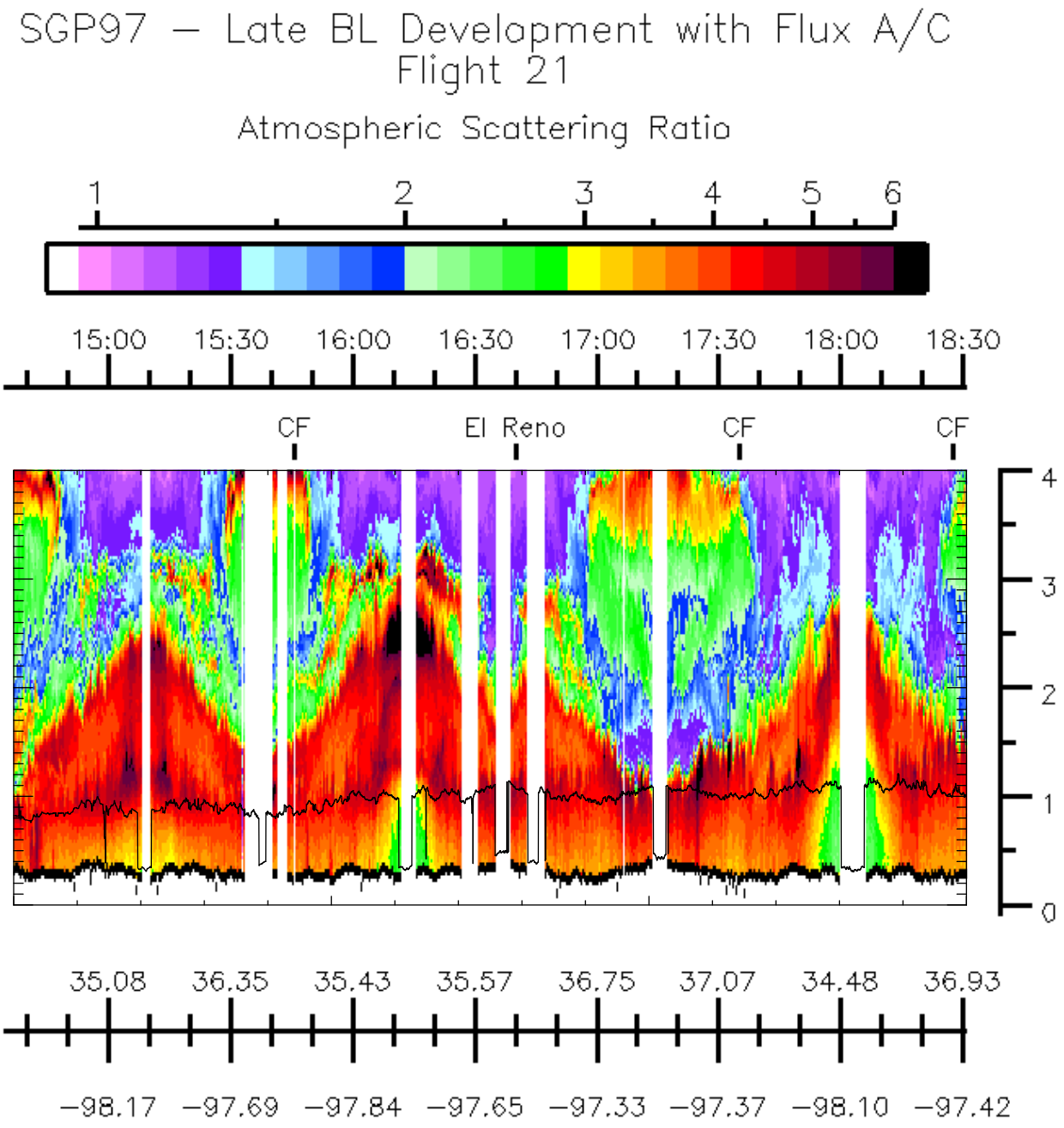


Figure 15. LIDAR atmospheric scattering ratio on 7/12/97 from 14:30-18:30 overlaid with ESTAR derived boundary layer height. LIDAR boundary layer height is dark section within red area. X-axis is position (lat,lon,time). Y-axis is altitude in kilometers.

	7/12			7/13		
Time	R	B	M	R	B	M
14.0000	0.784844	-0.417540	22.2378	0.603085	-0.173028	13.1567
14.5000	0.655432	-0.384856	23.9171	0.792620	-0.780177	25.3653
15.0000	0.619566	-0.375909	24.9035	0.863047	-1.04745	30.9029
15.5000	0.585101	-0.257723	21.3735	0.835209	-1.22468	34.6374
16.0000	0.648093	-0.383904	24.1991	0.756249	-0.548426	19.8731
16.5000	0.726773	-0.379169	22.3088	0.687698	-1.14700	29.0189
17.0000	0.758970	-0.484653	25.1826	0.809468	-0.772745	23.2097
17.5000	0.635140	-0.199847	16.0474	0.846154	-0.739904	21.9519
18.0000	0.570425	-0.155022	13.7246	0.731358	-0.180126	9.79395
18.5000	0.615800	-0.190542	14.9538	0.584454	-0.287861	14.1995
19.0000	0.667615	-0.358034	18.0618	0.771858	-0.543649	16.6416
19.5000	0.632011	-0.324288	15.2689	0.698477	-0.388507	12.5876
20.0000	0.443935	-0.282240	12.4644	0.747364	-0.336419	10.8568
20.5000	0.619443	-0.333457	14.7195	0.702183	-0.626494	16.0597
21.0000	0.386955	-0.216666	11.0252	0.681655	-0.658194	16.2305
21.5000	0.465176	-0.335277	14.0313	0.625516	-0.664543	15.8191
Max H	0.806483	-0.195027	18.6820	0.897687	-0.734517	22.9181

$$BowenRatio_{time} = B_{time} + \frac{M_{time}}{SoilMoisture}$$

Table 1. List of coefficients for converting soil moisture to Bowen ratio. Equation is changed from linear to inverse equation to allow for trailing off at high soil moisture levels. Below low soil mositure amount (around 15%), Bowen ratio is linearly interpreted to a fixed value. R is the correlation coeffient of this equation against measured Bowen ratio at given time. Max H is the fit when sensible heat flux is maximum.

Point-source air pollution risk buffers from long-term climate data and Gaussian plume models

Abstract

A Gaussian plume model developed by the U.S. EPA and NOAA is used to predict the long-term pollution impact of a hypothetical industrial smokestack on its immediate surrounding area. The model is used to predict hazardous ground-level pollutant concentration from an one-hour daytime accidental release. In conjunction with research on environmental justice issues in the Minneapolis metro area, this experiment considers the use of 20 years of hourly surface weather data from the Minneapolis-St.Paul Airport in the creation of risk buffers for GIS / environmental-equity analysis. The basis of Gaussian plume models is reviewed. A clustering methodology is developed to reduce the number of climate data points needed to model the plumes. Additionally, an overlay method is developed to create an annual probability buffer that shows the chance that any given point surrounding a stack would be subject to hazard in the event of an accidental release. The results are compared to traditional GIS environmental equity analysis methods. They show that while a simple buffer may be appropriate in many cases, a detailed cluster analysis based plume can pick out finer details of risk surrounding industrial emissions.

Introduction

Environmental geographic information systems often need to buffer environmental risks. Point-based risks are often modeled simply by location to population units or with arbitrarily sized buffers (McMaster et al., 1997). For example, a 500 meter buffer might be constructed to look at the impact of an industrial stack on its surrounding population and landscape. It may make more sense, however, to look at a model of air pollution dispersion and determine where precisely are the most affected areas. Such a model might consider the impact a stack would have in the case of accidental releases at any random time using given weather information. The buffer model developed in this paper considers the use of a relatively simple air pollution dispersion model together with long-term hourly meteorology.

Constructed buffers of these sorts may be used in conjunction with demographic data to understand the nature of distributed costs and benefits of environmental siting (Sheppard et al., 1999). This can raise concerns of whether environmental injustice is occurring in the siting of a

stack. The work in this paper is done in conjunction with research in progress by McMaster, Leitner and Sheppard (1997) on environmental justice issues in the Minneapolis / St.Paul metropolitan area.

McMaster et al. (1997) review some of the history of environmental equity analyses. Political challenges to the siting of industrial sites, the publication of reports on the relation between race or class and industrial siting (i.e. the United Church of Christ's Commission on Environmental Justice Report), the passage of federal acts that require environment equity analyses (SARA Title III) and the creation of the Office of Environmental Justice in the U.S. Environmental Protection Agency (EPA) have all catalyzed the environmental justice movement. However, most studies that have attempted to prove environmental bias have been limited by crude statistical methods. In many studies, characteristics of population in arbitrarily-sized enumeration units are statistically compared to the specific locations or arbitrarily-sized circular buffers of industrial sites. Proximity measures, however, do not rigorously account for exposure (Sheppard et al., 1999). Existence and location are not good surrogates for risk (McMaster et al., 1997). Moreover, statistical results from proximity measures are greatly influenced by the choice of population enumeration unit resolution (e.g., block group, tract, county) and of geodemographic variables used. Finally, proving inequity (statistical bias) does not necessarily prove injustice (intentional or institutional bias against specific subsets of people). A social analysis of all the actors involved in the siting of a stack along with measurements of inequity are required to prove environmental injustice.

The goal of this paper is to look at ways to incorporate information of chemical toxicity and local climatological information into environmental equity analyses. This experiment looks specifically at a potential way to model the risk associated with air pollution from a smokestack.

A number of assumptions about meteorology and pollutant dispersal go into this model, and they will be discussed below. I intend to show that despite these assumptions, the risk buffers created can be used to produce a clearer picture of the potential risks associated with a stack in a neighborhood. This paper reviews Gaussian plume models and develops the risk buffer methodology. Further research is needed to understand how to apply this model to geodemographic analysis at multiple scales.

The overall methodology for risk buffer creation involves 1.) the selection of relevant meteorological and pollutant data, 2.) the modeling of various dispersal scenarios with these data and 3.) a method for visually depicting the sum of the model runs and creating a "probability of risk" buffer from this view. The experiment described here uses a clustering methodology on 20 years of weather data along with a Gaussian plume model developed by the U.S. EPA and the National Oceanic and Atmospheric Administration (NOAA). The pollutant modeled is a hypothetical accidental one-hour release of 5000 kilograms of hydrogen chloride gas (not the same as hydrochloric acid), but could in general apply to the release of any neutrally buoyant gas that does not boil upon injection into the atmosphere (NOAA Office of Response and Restoration, 1999).

Site description

The Minneapolis / St. Paul metropolitan area is a low to medium density urban area situated within flat to gently rolling terrain. The overall climatology of the region is best described as northern continental with both strong winters and summers. Wind direction is highly variable, but is often westerly or south westerly. The winter is marked by a larger component of northerly and northeasterly winds. The high variability of wind, temperature and

atmospheric stability make predictions of pollutant dispersal difficult. Nevertheless, the generally flat terrain and low building heights / density allow for relatively simple pollutant models to be used in conjunction with known data about the variability of surface climatology.

Environmental justice has only become a recent issue in the metropolitan area as immigration influx and suburban outgrowth have lead to increased race and class stratification across the region. This has led to an unequal distribution of the costs and benefits of industrial plant location. Sheppard et al. (1997) has shown in an analysis of toxic releases and population demographics that this "inequity" is difficult to measure and highly dependent of the scale of analysis. For example, there is a greater correlation between minority population and industrial sites at the county level than at the scale of the city of Minneapolis.

This experiment looks specifically how one might be able to discern the very large cartographic scale impact of a smokestack in a neighborhood. Although it may not necessarily be the case that industries willfully site smokestacks "upwind" of the politically less powerful, it is quite possible that the geographic distribution health and pollution risk associated with a stack at the large scale may not be known at the time of siting. The methodology described below would help determine whether a new stack does harm a population in a politically-defined "unfair" way. Several pollution models do incorporate measures of chemical hazard as a surrogate for risk in the form of a toxicity index (McMaster et al., 1997; Pratt et al., 1993), but often do not incorporate the spatiality of risk due to variability in weather. The model described here hopes to alleviate that.

Data description

Data for this experiment comes from the U.S. Environmental Protection Agency (EPA) and the National Weather Service (NWS). Information about pollutants in the Minneapolis area comes from the Toxic Release Inventory (TRI) produced by the U.S. EPA Office of Pollution, Prevention and Toxics. All major industrial emitters are required to report the nature and quantity of their air emissions to the U.S. EPA. EPA compiles yearly statistics for each site in the TRI report. Although this particular experiment only explores the impact of a hypothetical smokestack, a full understanding of pollution risks and benefits would require the modeling of all major fugitive and regular releases from TRI sites across the region. The hypothetical release considered in this experiment was chosen as a reasonable release based on the TRI report for Minneapolis.

Risk levels for chemical toxicity were based on the Immediately Dangerous to Life or Health (IDLH) levels of concern sets by the National Institute for Occupation Safety and Health (NIOSH). IDLH levels were originally defined for the use of respirators in workplaces, but they are also considered standard levels of acceptable pollutant risk by many agencies (NOAA Office of Response and Restoration, 1999). The IDLH is defined as the maximum concentration a worker can be exposed to a chemical for thirty minutes without suffering permanent or escape-impairing health effects. For the general public, this level is often reduced one order of magnitude (1/10 IDLH). The 1/10 IDLH has been adopted by the U.S. EPA, U.S. Department of Transportation (DOT) and the Federal Emergency Management Agency (FEMA) for modeling conservative scenarios of pollutant risk to the general public, since different people respond in wide-ranging ways to toxic chemicals.

Climatology data came from National Weather Service (NWS) weather data for the Minneapolis / St. Paul Airport. These data were obtained from the National Climate Data Center (NCDC) through the assistance of the Minnesota State Climatology Office. Data only after 1969 were used since prior to then, wind directions were reported in multiples of 20 degrees, whereas after 1969, directions were reported in multiples of 10 degrees. Thus, hourly weather data from 1969 to 1989 were considered. This seemed to be a reasonable average of long-term hourly weather variability for the region. Six surface weather variables were considered for the plume model: temperature, wind direction, wind speed, relative humidity, percent cloud cover, and atmospheric pressure. These variables are used to approximate mean wind and turbulence or stability characteristics of the atmosphere, and also represent the variables needed by the plume model.

The plume model used was developed jointly by the U.S. EPA Chemical Emergency Preparedness and Prevention Office and the NOAA Office of Response and Restoration. The model, called Areal Locations Of Hazardous Atmospheres (ALOHA), was designed for use by emergency response planners in the case of chemical spills. It is designed to model a wide variety of spills given a relatively small amount of meteorological and chemical information. The resulting plumes are mapped onto GIS databases of the users' choosing (i.e. street map, building footprints, etc...) (NOAA Office of Response and Restoration, 1999).

Atmospheric Turbulence and Pollution

Pollution occurs at a number of scales. The local industrial smokestack produces pollution that affects the surrounding area. A conglomeration of stacks affects the overall pollution hazard for an urban area. The flow of pollutants over large distances from an industrial

area affects entire regions. The sum of theses regions affects pollutant flow in the continental and finally global scale (Stern et al., 1984).

The scale of the analysis in this paper is the industrial smokestack on the neighborhood. Pollutant modeling at this scale is typically done in a source-receptor framework (Stull, 1995). The pollutant is emitted at a known rate from a point source and is followed downwind as it is dispersed and diluted by turbulence and diffusion until it reaches a point of interest, the receptor. To create the plume of the pollutant, a set of receptors are selected for the source point, and a contour line is drawn around a given concentration level of concern (ex. 1/10 IDLH) from the "control points" or receptors.

The kind of pollutants that are the easiest to model and what are modeled in this paper are known as passive, conservative tracers (Stull, 1995). These pollutants are typically gases or fine particles that ride in the wind "passively." They do not decay, react or fall out during transport in the atmosphere. A single particle of this kind pollutant disperses entirely as a function of atmospheric turbulence. For this to be the case, the pollutant has to have about the same density as the atmosphere, and be neutrally buoyant (does not rise on its own due to high exit temperature or other factors.) With these constraints, the major factors that affect pollutant flow are: horizontal wind speed generated by the geostrophic wind altered by terrain frictional forces and local winds, atmospheric stability and turbulence characteristics, temperature profiles, synoptic meteorology, terrain, pollutant emission rate and stack height (Zannetti, 1990). Pollutant plumes from passive tracers are sometimes known as momentum-dominated plumes, where the exhaust temperature is not much higher than ambient air (Economopoulos, 1992)

Pollutants from low industrial stacks generally disperse into the lowest layer of the atmosphere, known as the atmospheric boundary layer (ABL) or planetary boundary layer

(PBL). The boundary layer is the region of atmosphere that is directly affected by the surface via exchanges between the atmosphere and surface of mass and energy, and the generation of turbulence due to variation in land cover and terrain (Stull, 1988). This is analogous to the combination of the no-slip condition of fluid mechanics and energy transfer laws of thermodynamics. The boundary layer can range in height from near the surface to several kilometers above the ground, and occasionally extends beyond the troposphere during thunderstorms. Above the boundary layer, atmospheric flow is generally laminar and aligned with the geostrophic (pressure gradient) wind flow. Mean wind flow in the boundary layer is altered by turbulence creating eddies (deviations from mean flow) of various sizes. Turbulence leads to a greater amount of mixing of atmospheric constituents. Turbulent eddies are involved both in the transport and diffusion of pollutants. Eddies smaller than the pollutant plume tend to have the effect of diffusing the pollutant, while eddies larger than the plume tend to transport the pollutant (Oke, 1987).

The boundary layer can be subdivided into a mixed layer where turbulent motion is generated by terrain roughness and convection, and a thin surface sublayer where turbulence is not fully developed. The height of this sublayer is sometimes called the roughness length, and can be approximated as a fraction of the average height of surface protrusions or estimated empirically from a log-log profile of wind with height (Stull, 1988). The roughness length is an important factor in determining the strength and scales of turbulence that occur in the surface layer (Snyder, 1981). Typical roughness lengths vary from .01-.1 cm for sand, 1-10 cm for grass, 90-100 cm for forests, and 35-100 cm for towns and cities.

Scales of turbulence in the surface or mixed layer are also affected by thermodynamic heating of the Earth's surface by the sun. Heat energy radiated from the surface drives

convection, which creates larger eddies and rolls in the boundary layer. The height to which mixing due to turbulence and convection occurs in sometimes known as the boundary layer height (Stull, 1988). This height is often similar to the inversion height, the height at which temperature, which usually decreases with height adiabatically in the daytime boundary layer, begins to decrease at a lower rate or even increase. If this height is low, pollutants released by a stack may bounce off the top of the boundary layer, in addition to reflections occurring off the surface (Oke, 1987).

Another measure of levels of turbulence in the boundary layer is atmospheric stability. Stability is generally measured from how temperature changes with height. Unstable conditions occur when temperature decreases with height at a rate higher than the adiabatic lapse rate. Clear, daytime conditions with a positive surface heat flux are typically unstable, and the boundary layer is well-mixed due to moderate to strong turbulence. Neutral conditions are characterized by an adiabatic temperature decrease with height. Neutral stability is often found during evening nighttime transition, overcast skies or strong winds. Stable conditions occur when temperature decrease is less than the adiabatic lapse rate, or when temperature increases with height (i.e. temperature inversion). Stable air is typically found over land on clear nights with weak winds. Heat flux is typically downward (negative) and turbulence is weak, leading to stratification of the boundary layer into multiple stable layers (Zannetti, 1990). Stable layers are the most difficult of the three stability conditions to parameterize for dispersion modeling, but also can produce some of the most dangerous pollutant concentrations, due to a lack of efficient diffusion (NOAA Office of Response and Restoration, 1999). Figure 1 shows some idealized plumes in different stability conditions (Oke, 1987).

Usually, in meteorological plume modeling, information about temperature change with height is not known. Instead, stability characteristics are estimated from empirical methods that relate surface weather characteristics to stability. A common method is a set of dispersion classes originally created by F. Pasquill and later modified by F. A. Gifford, known as the Pasquill-Gifford dispersion classes (NOAA Office of Response and Restoration, 1999). Table 1 shows the classes and how they related to measured weather values. The classes are estimated from wind speed, solar insolation (a function of sun angle) and cloud coverage (Zannetti, 1990). Other stability typing schemes include the Turner class (similar to Pasquill-Gifford) and the Golder nomogram (based on Monin-Obukhov Length and roughness length) (Kunkel, 1985; Zannetti, 1990).

Pollutant Dispersal Modeling

Knowledge of atmospheric diffusion parameters is critical to the modeling of pollutant plumes. Models of pollutant dispersal can be either statistical or deterministic in method (Zannetti, 1990). The statistical methods involve creating semiempirical relationships between available data and measurements made of actual pollutant plumes (Zib, 1977). Proportion models assume that the concentration at a receptor is some proportion of concentration oat the source. Regression models attempt to find a relationship between meteorological parameters and concentration at the receptor. Thrane (1988) attempted to find a regression relationship between receptor measurements of fluoride and polycyclic aromatic hydrocarbons (PAH) and the locations of aluminum plants over two years of meteorological data, and found that the regression worked better with PAH than fluoride. The regression is only as good as the choice and quality of data used.

Deterministic models rely more on fundamental descriptions of atmospheric processes. These models involve combining the basics of atmospheric fluid dynamics with atmospheric chemistry and parameterizing these equations for the atmospheric boundary layer. The most general application of this would include the fully defined 3D Navier-Stokes equations of motion, along with equations for pressure-height relations (i.e. hydrostatic pressure), 1st and 2nd law of thermodynamics, gas state behavior, plume rise, and wet and dry deposition behavior of the pollutant (Zannetti, 1990). Modeling of these equations in the atmosphere requires a statement of simplifying assumptions, boundary conditions, modeling time / space steps, scaling, and similarity and closure analyses for characterizing scales of turbulence. This model is highly dependent on boundary conditions. However, rarely in emergency planning or in long-term climate data is full information about atmospheric boundary conditions and pollutant release available. Computational time is also a factor, especially when modeling complex terrain or chemical interactions. Sometimes, scale models in wind or water tunnels may be faster and more accurate (Snyder, 1981).

The Gaussian plume model is a statistical simplification of a deterministic model and can be derived from either Eulerian, Langrangian or empirical modes of deterministic modeling. Eulerian methods of analysis take place in a fixed reference frame (often centered at the pollutant source) and look at the "big" picture of pollutant flow by modeling changes in "state" of the surrounding atmosphere over a grid of space and time. The analytical solution to the Eulerian equations of motion would show the concentration of pollutant at any point downwind of the source. The simplest model Eulerian model, the box model, consists of only grid points at the edges of a box of interest and essentially gives general information of overall pollutant concentration (Zib, 1977). The most advanced Eulerian models to date involve the use of large

eddy simulations (LES), where the smallest scales of turbulence are empirically estimated and the rest are modeled on a fine grid scale.. The basic Eulerian equation is:

$$\frac{\partial c}{\partial t} = -V \cdot \nabla c + D\nabla^2 c + S \qquad (1)$$

(Zannetti, 1990)

where $\delta c/\delta t$ is the change in pollutant concentration over time, V is wind velocity, ∇c is gradient of concentration over space, D is molecular diffusivity, and ∇^2 is the LaPlacian operator. The first term is an advection term for the motion of the pollutant over space, the second is a term for molecular diffusion and S is a source/sink term for deposition.

Lagrangian models look at pollutant dispersal differently. Instead of modeling the state of the atmosphere at one time, Lagrangian models follow "particles" released from the stack over time. The reference frame moves with these particles. Particles are not actual molecules of pollutant species, but are generally some discretized length / time scales that define the particles to follow. Interactions between particles must be explicitly modeled (Zannetti, 1990).

Gaussian plume models rely on statistical interpretations and simplifications of Lagrangian models (Gaussian puff models) and Eulerian models (steady-state assumptions), as well as empirical, statistical experiments on plume shape. The Gaussian plume model is designed to calculate the concentration of a pollutant over space and time given limited boundary conditions, namely stationary meteorological and emission conditions (Economopoulos, 1992). It is the most common air pollution model. The model assumes that the pollutant is transmitted in the direction of the mean wind from the source as a function of average emission rate and wind speed. Turbulence is assumed to be stationary and homogenous around the mean wind, though vertical and horizontal turbulence can be different. Other assumptions include constant

eddy diffusivity, complete reflection at the top of the mixed layer and the ground, and generally flat topography (Economopoulos, 1992; Stull, 1995).

With these assumptions, it can be shown that pollutant concentration in crosswind and vertical deviations from the mean wind decay has a Gaussian or a normal bell curve distribution. The basic Gaussian plume model equation is:

$$c(r_{x,y,z,t}) = \frac{Q}{2\pi\sigma_y\sigma_z|\bar{u}|} \cdot e\left[-\frac{1}{2}\left(\frac{y}{\sigma_y}\right)^2\right] \cdot \left\{e\left[-\frac{1}{2}\left(\frac{z-z_{CL}}{\sigma_z}\right)^2\right] + e\left[-\frac{1}{2}\left(\frac{z+z_{CL}}{\sigma_z}\right)^2\right]\right\} \qquad (2)$$

(Stull, 1995; Zannetti, 1990)

where c is the concentration at r = (x,y,z), Q is the emission rate, σ_y and σ_z are the standard horizontal and vertical deviations of plume concentration, u is average wind velocity at emission height and z_{CL} is the height of plume centerline above ground.

This equation assumes neutrally buoyant gases (about the same density as air) and homogenous and stationary meteorology conditions between source and receptor. The model does not work as wind speed approaches 0. Additionally, the model must be run for at least a 10 minute period, and concentration values are only valid in an approximate range of 100 meters to 10 km downwind (Oke, 1987). The Gaussian equation assumes steady diffusion along the mean downwind direction and one standard deviation of variance in concentration around this mean wind. Figure 2 shows different views of an idealized Gaussian plume concentration surface (Oke, 1987).

The key to creating a good Gaussian plume model is the selection of good values for σ_y and σ_z, since they are the main measures of turbulent diffusion and transport in the equation. Most Gaussian models rely on empirical formulations of these values, typically based on Pasquill-Gifford stability class and downwind (x) distance to source. Zannetti (1990) reviews

some of the basic schemes for calculating σ. The earliest formulations were based on an empirical plume modeling project from 1956 called Project Prairie Grass (NOAA Office of Response and Restoration, 1999). During the project, which occurred over flat grasslands in Nebraska, plumes of sulfur dioxide were released and plume concentration was measured at numerous points. Turner and later Green turned the results from this experiment into coefficients that could be used in an equation. Power law functions from another experiment in rougher terrain were used to create Brookhaven sigmas. Briggs created a typing scheme (table 2) that separated values for σ for urban and rural areas based on the results of Pasquill-Gifford, Brookhaven and urban dispersion experiments by McElroy and Pooler. Because of the variety of empirical formulations for σ, choice of σ is subject to subjective bias, which could cause controversy in decision making (Hanna, 1982). Briggs' model appears to be the most accepted in Gaussian plume models.

While the Gaussian model is good at getting quick views of averaged concentration over space and time, it does have some limitations. Plume rise from buoyancy effects are not modeled. They have to be explicitly added with additional terms to the equation (Zannetti, 1990). For overall concentration values, however, buoyancy may not have that large an effect (Economopoulos, 1992). Also, recent research suggests that plume shapes may not have so much a normal distribution as a clipped-gamma distribution (Yee, 1997).

It is important to note that the Gaussian shape is an average concentration over some time period. The dosage received at any point is not smoothly received. Turbulence causes concentration to vary all over the place and the chemical arrives in short and necessarily intense bursts (Jones, 1996). This is especially the case in areas with obstacles, such as urban areas. Obstacles will create local eddies, reflection and variations in concentration and diffusions in

complex ways that are partly a function of plume size. An experiment with urban dispersion using an 2 m^3 obstacle array suggests that small obstacles may not affect overall plume concentration, but do change plume structure due to increased upward air motion (Davidson et al., 1995). Larger obstacles will increase the deviations of concentration from the Gaussian model.

Newer models have attempted to treat Gaussian models in a Lagrangian manner using "puff" models that apply Gaussian equations to each "particle" in the flow (Zib, 1977; Zannetti, 1990). Increased computation speed has allowed for easier modeling of Eulerian grids, involving numerical integration of the differential equations for mass, momentum and energy conservation and treating plumes as a combination of both gas and vapor (Gangoiti et al., 1997). Gaussian plumes can also be embedded into a larger grid model (Seigneur et al., 1983). These developments will enhance the usefulness of the Gaussian plume dispersion model.

The Gaussian plume model used in this experiment, ALOHA, uses the Pasquill-Gifford stability classes (table 1) and the Briggs equations (table 2) for calculating σ. ALOHA was used in this experiment since it provided an easy way to map plume levels for a Geographic Information System (GIS) and had a built-in chemical library of IDLH values. ALOHA also has modules for dealing with heavy gas dispersion (useful since hydrogen chloride is actually a bit heavier than air), and wet and dry depositions. It provides warnings when chemicals will most likely fail to hold Gaussian assumptions. While terrain is assumed to be obstacle free, ALOHA does allow the user to set roughness lengths for modeling differences in wind speed with height.

Probability Risk Buffer Methodology

The overall goal of this experiment is to understand risk probabilities around a stack emission. This is accomplished by using ALOHA to model plumes for all possible weather conditions for a point, overlaying these "potential plumes", calculating a hazard probability density function, and mapping this information. These steps require two basic methods: a method for selecting "all possible weather conditions", and a method for overlaying and calculating statistics on the overlaying plumes.

Hourly weather data from Minneapolis / St. Paul International Airport were used to calculate average weather conditions. A variability analysis shows that simple monthly averages would be inaccurate for modeling plumes. Figure 3 shows the direction and magnitude of wind in vector form for every hour for the month of January 1981. From this is can be seen that neither monthly nor daily averages of wind velocity reflect the variability of wind within any given time frame. Figure 4 shows a 3D histogram of the wind vector at all hours for Januarys of 1969-1989. Figure 5 shows the same for all months of the year. These two figures show that while an average monthly velocity may be inaccurate for modeling purposes, there are certainly clusters of wind velocities that exist in various places. Further analyses showed these clusters of winds to be related to other variables including temperature and humidity. Thus, a minimum-distance-to-means ISODATA clustering algorithm for the six variables ALOHA needs (temperature, pressure, humidity, wind speed, wind direction, cloud cover) was chosen. Additional clustering weights (2x) were given to wind speed and wind direction (normalized to an orthogonal {u,v} wind vector), since they are the primary controlling factor on plume direction and shape. These weights force the clusters to have different wind directions. The clustering method iteratively computes clusters of each six-dimensional data point until a

measure of divergence (ratio of in-class to between-class variability) reaches a threshold value. The initial clusters are randomized to data values that are spread across the mean value (Jensen, 1996). Ten clusters along with their frequency of occurrence were computed for each month, producing 120 clusters. Figure 6 shows how the same winds from figure 3 clustered against all the data. Other clustering methods that could have been used include agglomerative hierarchical clustering (Weber et al., 1995) designed for interpolating wind fields between stations, or a principal components analysis of variability (Green et al., 1993). For this experiment, it was decided that the ISODATA algorithm provided reasonable clusters.

A secondary clustering method was also done. This time, eight clusters were computed for each month. Each cluster consisted of hourly readings whose wind directions were closest to one of the eight cardinal or ordinal directions of a compass rose. The motivation behind this method was to create GIS buffers by connecting the centerline of plumes in each of the eight directions. It also served as a comparison against the ISODATA clustering.

Ten clusters for twelve months yields 120 weather conditions to run on ALOHA. (The second clustering method produces 96 plumes). These are far fewer conditions than having every data point or some frequency of every data point. On the other hand, information about time of day and day of year is lost, which is used in the stability calculations to determine sun angle. Because of this, nighttime weather data was not used. A further refinement of this clustering analysis might include clustering the time of day along with the other variables. For this experiment, a constant time (summer afternoon) is used for all plumes. Overall, a variability analysis showed that the plume shapes do not change all too much with this limitation.

The clusters were run against a typical accidental release in which it was assumed that the chemical was neutrally buoyant and would not boil over or freeze. The release was of 5000

kilograms of hydrogen chloride gas over one hour in the city of Minneapolis (83.3 kg/minute). An average urban roughness length of 100 cm and typical stack height of 10 m were selected for the model. ALOHA was asked to map the extent over which the plume concentration of hydrogen chloride exceeded its IDLH of 50 ppm during one hour. The output from ALOHA is a set of geographic coordinates defining both the contour line of the plume where concentration exceeds the IDLH and the plume confidence interval, defined as one standard deviation from the edges of the contour line. The confidence interval is a conservative measure of variability of wind direction through the hour, which may vary from 30-45 degrees within any one hour depending on atmospheric stability (Oke, 1987).

Each set of ten plume confidence intervals for a month were rasterized into one square meter grid cells extending over the entire domain of the plumes (about 1 kilometer radius from the center). The value of each grid cell for each plume (created for each weather cluster) was equal to the number of hours that the related cluster of weather observations occur in that month. Each plume' grid cells were then overlaid and added together. The value of each grid cell was divided by the total number of hours in the month to produce a monthly probability density map showing the percentage probability that any one point would be exposed to dangerous levels of an accidental release. These images were smoothed to 25 square meter grid cells to smooth out the image as a form of spatial generalization. Ideally, in addition to plume frequency, actual or average concentration values inside the plume should have been used. But since that information is not directly available from ALOHA and since the overall goal was to create a generalized probability density buffer, that information was not used. The output monthly probability maps were then overlaid with the other months to produce both seasonal and annual composite probability density buffers

Another plume overlay method using the second clustering method involved connecting the centerline of each plume cluster for the month and overlaying the months. The buffers produced this way were hard to analyze. Frequency information was added, and instead of connecting the centerlines, a plume "pie" chart was created showing the probability that any one direction from a plume would be subject to health hazards.

To validate these clustered overlaid plumes, another risk buffer was created using every hour from January 1981. This required running ALOHA approximately 720 times. These plumes were then rasterized and overlaid in the first overlay method described above. This risk buffer was compared to risk buffers created with clusters for all Januarys.

Results

Figure 7 shows the plume risk buffers for each month based on the first clustering method. Contour lines of either the 20% or 40% risk probability are shown. The results show average risk plumes extending out to 1 km, though at the 20% risk level, the health hazard extends only about 500 m for the pollutant modeled. Overall, the plumes for each month track closely with the wind velocity histograms for that month. This suggests that wind direction and velocity are the primary determining factors in plume extent. Temperature, humidity, sky cover and pressure primarily affect the stability factor in the plume model. It is harder to discern how those variables affects plume extent from the figures. They have more an effect on plume shape. It is also possible that warmer, southerly winds are more stable than cold, northerly winds, and this will change the width of the plumes of different directions in different ways. However, the final output product masks these results. Overall, the effect is probably small compared to the importance of wind direction and velocity. Zannetti (1990) argues that using Gaussian plume

models in the long term way suggested here may actually provide better overall results since individual errors cancel out.

Figure 8 shows the plume "pie chart" buffers based on the second clustering and overlay method. The buffers are similar to those seen in Figure 7 though the details of features get smoothed out in the process. Both figures 7 and 8 tend to show that areas south or east of point source emissions typically are at lowest risk from hazardous exposure depending on the month. Figure 9 shows the plume buffers overlaid by season with the first method. The seasonal buffers show greater risk to the southwest in the winter, northwest in the spring, northeast in summer, and north in fall. Figure 10 shows the annual plume buffer. The annual plume buffer is essentially the spatial risk buffer around this point-source emission. For this chemical in Minneapolis, it shows that generally areas to the northwest of the point have a higher probability of hazardous risk than areas to the east. Areas to the south fall in between the two. Figure 11 shows the annual plume buffer from the second clustering method. Both clustering methods produced probability maps that compare reasonably well to each other, especially in the annual buffer. The biggest difference is that individual plumes are narrower in the first clustering method, leading to higher risk probabilities for smaller areas. The overall direction of risk, however, is similar between the two methods.

Figure 12 shows the validation risk buffer based on hourly overlaid plumes for January 1981. This risk buffer is a bit different from the January image in Figure 7. This is partly because the validation buffer is only for 1981 while the January buffer is for all Januarys. The shape of the space of risk is similar, though, suggesting that the clustering method is a reasonable approximation of risk around a point source emission.

Further verification of these plumes might require empirical / historical studies of pollutant release and risk assessment. Since long term data collection on weather is limited to ground based measurements at the hourly level, the Gaussian plume model is the most straightforward way to model pollutant plumes. With additional information, it would make sense to use a more robust model that rigorously incorporates variability of the atmosphere over space and time.

Discussion

The results show that the cluster/overlay method is validated against an hourly plume model for January 1981. The composite annual risk buffer image shows a general assessment of risk across an area. It is not to be taken as ground truth, i.e. the risk at an individual grid cell has a large margin of error due to the limitations of Gaussian plume models. The big picture is what is more important for incorporating the plumes into a GIS for equity analysis.

The largest source of error is the use of ground level hourly weather measurements from the airport. Urban canyon effects, microvariability of weather and chemicals over space and time, and estimation methods for modeling boundary layer conditions can all lead to problems with the Gaussian plume model (Davidson et al., 1995; Zannetti, 1990). Photochemistry, wet/dry deposition and particulate matter chemistry of the pollutant can also affect hazard plume shape and size. The chemical chosen, hydrogen chloride, is occasionally modeled better as a heavy gas. As such, the plumes from this experiment can be considered generalized for any large Gaussian spill as opposed to specifically representing hydrogen chloride, except for the IDLH value and release amount.

Another source of error is the overlay method of risk buffer creation. Each individual plume is treated as just an area of human health hazard, not the actual concentration value which varies from the centerline out to the edge. The ALOHA IDLH plume is a measure of the area at ground level that would experience concentrations of the chemical that are hazardous to human health with over 30 minutes of exposure. A better measure of risk would use the actual average ppm concentration over any hour and get an average ppm risk from all the plumes that overlay any one point. ALOHA, however, does not provide a direct way to access plume centerline concentration other than graphs. Future work should look into ways of getting better estimates the concentration values.

Still, with the data given, the Gaussian plume model and the clustering/overlay methodology appear to be a good model for making broad, general statements about the geography of risk to human health from accidental release of an industrial air pollutant released from a stack. To use this model in a real case, one might need to further refine the choice of weather data and model multiple release scenarios. Historical information about a plant's operation may provide data on when industrial pollutants are released, if not all the time. This information can be used to mask out non-operating hours' climate conditions. Toxic Release Inventory data may give insight into the size of typical accidental spills from that stack.

The larger question in this experiment is how to use this information about probability risk buffers in equity analysis. Once one does create a probability risk buffer, how does one go about determining its relation to sociodemographic variables in the surrounding neighborhoods? Traditional buffers which assume equal risk in all directions would have looked for a correlation between stack location and poverty or minority status (McMaster et al., 1997). A risk buffer could be treated simply as a more detailed buffer by simply selecting an arbitrary probability risk

level and using the contour of that percent probability level as a buffer. The results from this experiment tend to show that depending on the probability level one chooses, the overall result in analysis may not be all that different than simple circular buffers . The main difference is a reduction of risk to the average "upwind" side of a stack. What one does get out of the plume model is a sense of what an appropriately sized buffer should be. Each pollutant emitter in an area can have scaled buffers in relation to their plume sizes and risks with this model.

An even more nuanced version of demographic analysis might use additional probability information from the risk buffer. This information could be used to find a regression model between probability of risk at sets of grid cells and some demographic variable. A model of this sort assumes that the resolution of the demographic data is at least as good as the plumes'. Otherwise, it probably makes sense to take the average plume risk for each demographic unit (i.e. census block group).

These kinds of demographic analyses, while an important start in understanding the intersection of industrial pollution and neighborhood composition, do not form a complete environmental justice analysis. They can only show correlation in space. Without data about change in demography and pollution over time, and without an understanding of the large number of other factors at play in industrial siting (i.e., access to transportation, political resistance, cost of land and development, etc...), an environmental justice analysis is lacking critical information. The Gaussian probability risk buffer model described here is simply a way to develop better buffers around point source air pollutants.

Conclusion

The Gaussian method of modeling plumes from a point source was used to model the

concentration of a pollutant around a stack. ALOHA, a plume model developed by the U.S. EPA. and NOAA is designed to create plumes around areas where ground level concentration exceeds a user-defined toxicity level. The IDLH toxicity level index was used to create plumes for an hypothetical accidental release of hydrogen chloride in the city of Minneapolis. Hourly ground-level weather data from the airport were used to create clusters of "possible" weather conditions which were then used to model the possible plumes from an accidental release. These hazard plumes and their confidence intervals were then overlaid and aggregated based on their occurrence frequency by month, season and year. Spatial aggregation created 25 square meter grid cells measuring probability of exposure to hazardous pollution from an accidental release that occurred over one hour. The results were validated against a similar model using all hourly measurements for one month. These results reasonably agreed suggesting the clustering method is not oversimplifying the data too much.

The risk buffers could be used in conjunction with toxic release and demographic data across a city or region of interest to understand how risk varies over space and how demography and pollution intersect. Additional information on site history is critical to formulate a strong thesis on environmental justice. While better plume models could be used to get a more accurate picture of pollutant hazard levels across space and time, most of the time one is limited by the data available: hourly surface weather observation and a list of yearly and fugitive stack emissions. Little information is usually available on chemical mixing, urban obstacles, boundary layer structure or exit temperature / velocity. Thus, with the information available, a Gaussian plume model run on clusters of long-term climate data in conjunction with an overlay method for calculating overall risk provides a good base from which to analyze the spatial variability of hazardous exposure.

Acknowledgements

This project was aided by the generous support of Dr. Bob McMaster in the Department of Geography at the University of Minnesota. I am also indebted by the help of the rest of the research colleagues in the environmental justice project, including Dr. Helga Leitner, Dr. Eric Sheppard, Jeff Matson and Hongguo Tian. I would also like to thank U.S. EPA and NOAA for providing the ALOHA program, and to Greg Spoden at the Minnesota State Climatology Office for providing assistance in retrieving the hourly climate data.

Works Cited

Davidson, M.J., K.R. Mylne, C.D. Jones, J.C. Philips, R.J. Perkins, J.C.H. Fung, and J.C.R. Hunt. 1995. Plume Dispersion Through Large Groups of Obstacles - A Field Investigation. *Atmospheric Environment*, v. 29, n. 22: 3245-3256.

Economopoulos, Alexander P. 1992. A Model for the Maximum Credible Hourly Impact on Any Ground Receptor from Point Sources with Momentum-Dominated Plume Rise. *Environmental Monitoring and Assessment*, v. 21: 99-131.

Hanna, S. R. 1982. Applications in Air Pollution Modeling, in F.T.M. Nieuwstadt and H. van Dop, *Atmospheric Turbulence and Air Pollution Modeling*. Dordrecht, Netherlands: D. Reidel Publishing Company.

Gangoiti, G., J. Sancho, G. Ibarra, L. Alonso, J.A. Garcia, M. Navazo, N. Durana, and J.L. Ilardia. 1997. Rise of Moist Plumes From Tall Stacks in Turbulent and Stratified Atmospheres. *Atmospheric Environment*, v. 31, n. 2: 253-269.

Green, Mark C., Robert G. Flocchini and Leonard O. Myrup. 1993. Use of Temporal Principal Components Analysis to Determine Seasonal Periods. *Journal of Applied Meteorology,* v. 32: 986-995.

Jensen, John R. 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Upper Saddle River, NJ: Simon & Schuster, Inc.

Jones, Chris. 1996. Something in the air. *The Chemical Engineer*, 28 November: 21-26.

Kunkel, Bruce A. 1985. *Development of an Atmospheric Diffusion Model for Toxic Chemical Releases.* Hanscom AFB, MA: Air Force Geophysics Laboratory, Environmental Research Papers, n. 939, AFGL-TR-85-0338.

McMaster, Robert B., Helga Leitner and Eric Sheppard. 1997. GIS-based Environmental Equity and Risk Assessment: Methodological Problems and Prospects. *Cartography and Geographic Information Systems*, v. 24, n. 3: 172-189.

NOAA Office of Response and Restoration. 1999. *About ALOHA*. NOAA. http://response.restoration.noaa.gov/cameo/aloha.html.

Oke, T. R. 1987. *Boundary Layer Climates.* New York: Routledge.

Pratt, Gregory C., Paul E. Gerbec, Sherryl K. Livingston, Fardin Oliaei, George L. Bollweg, Sally Paterson, and Donald Mackey. 1993. An Indexing System for Comparing Toxic Air Pollutants Based Upon Their Potential Environmental Impacts. *Chemosphere*, v. 27, n. 8: 1359-1379.

Seigneur, Christian, T. W. Tesche, Philip M. Roth, and Mei-Kao Liu. 1983. On the Treatment of Point Source Emissions in Urban Air Quality modeling. *Atmospheric Environment*, v. 17, n. 9: 1655-1676.

Sheppard, Eric, Helga Leitner, Robert B. McMaster and Hongguo Tian. 1999. GIS-based measures of environmental equity: Exploring their sensitivity and significance. *Journal of Exposure Analysis and Environmental Epidemiology*, v. 9, n. 1: 18-28.

Snyder, William H. 1981. *Guidelines for Fluid Modeling of Atmospheric Diffusion.* Research Triangle Park, NC: Environmental Sciences Research Laboratory, U.S. Environmental Protection Agency, EPA-600/8-81-009.

Stern, Arthur C., Richard W. Boubel, D. Bruce Turner, Donald L. Fox. 1984. *Fundamentals of Air Pollution*. London: Academic Press, Inc.

Stull, Roland B. 1995. *Meteorology Today for Scientists and Engineers*. St. Paul, MN: West Publishing Company.

Stull, Roland B. 1988. *An Introduction to Boundary Layer Meteorology*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Thrane, K. E. 1988. Application of Air Pollution Models: A Comparison of Different Techniques for Estimating Ambient Air Pollution Levels and Source Contributions. *Atmospheric Environment*, v. 22, n. 3: 587-594.

Weber, Rudolf O. and Pirmin Kaufmann. 1995. Automated Classification Scheme for Wind Fields. *Journal of Applied Meteorology*, v. 34: 1133-1141.

Yee, Eugene and R. Chan. 1997. A Simple Model for the Probability Density Function of Concentration Fluctuations in Atmospheric Plumes. *Atmospheric Environment*, v. 31, n. 7: 991-1002.

Zannetti, Paolo. 1990. *Air Pollution Modeling: Theories, Computational Methods and Available Software*. New York: Van Nostranf Reinhold, Inc.

Zib, P. 1977. *Urban Air Pollution Dispersion Models: A Critical Survey*. Department of Geography and Environmental Studies, University of the Witwatersrand, Johannesburg, Occasional Paper No. 16.

List of Figures and Tables

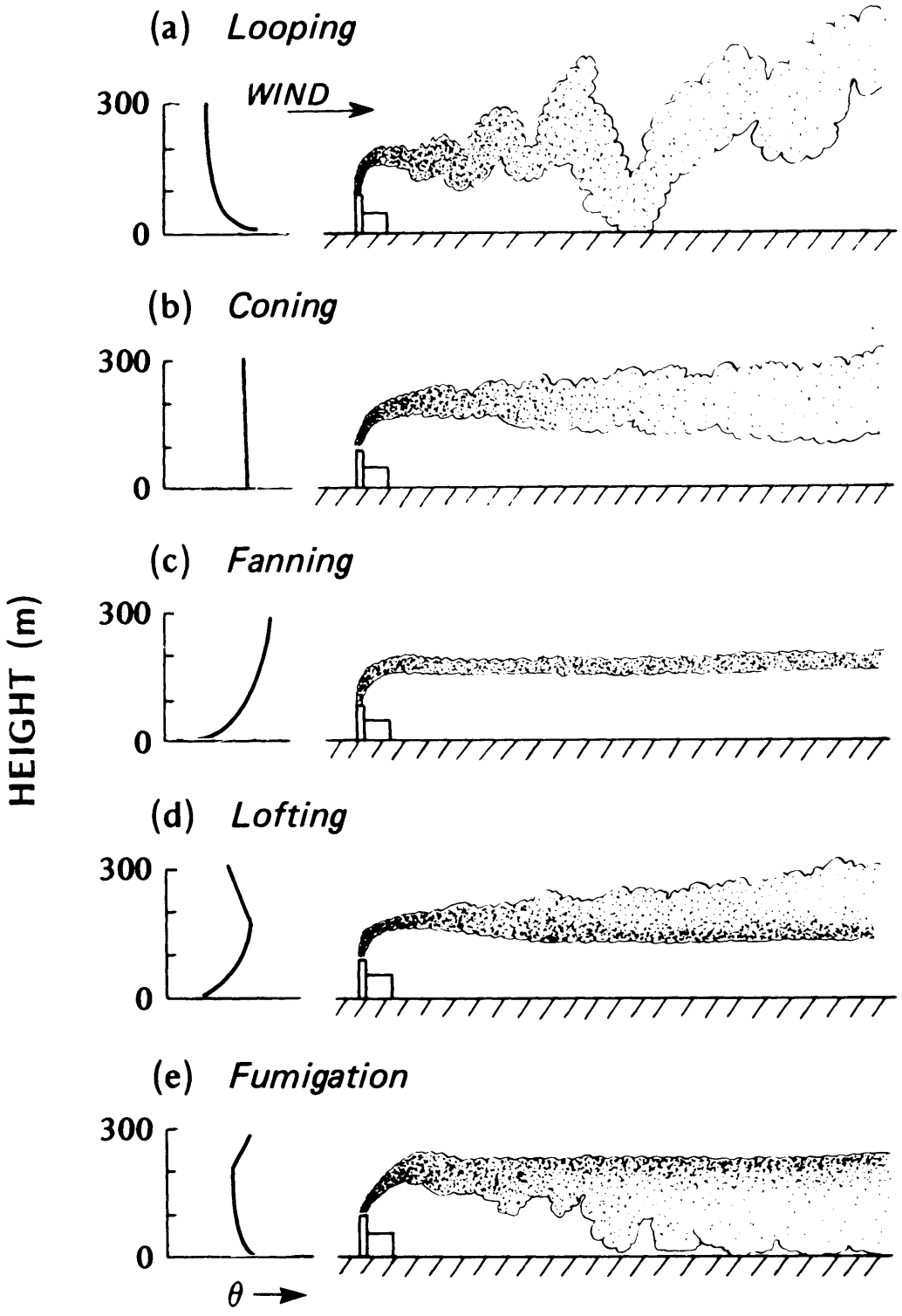


Figure 1. Schematic from Oke (1987) showing smoke plume shape under different stability conditions. Right half of graph is potential temperature vs. height.

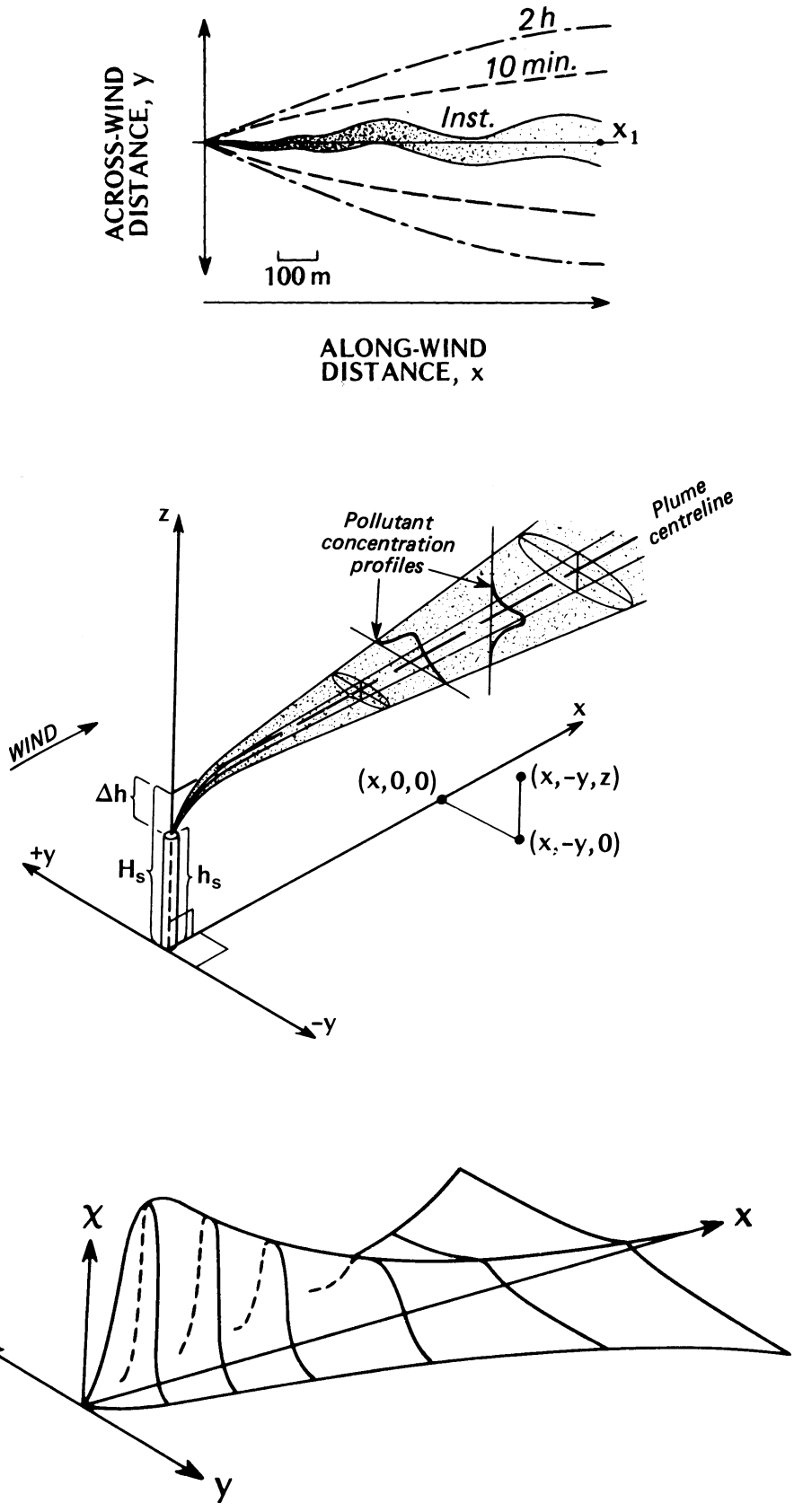


Figure 2. Schematics from Oke (1987) showing different views of a Guassian plume. The top graph shows how an instantaneous plumes average out to a Guassian shape over time. Central graph shows the reference frame for a plume. Bottom graph shows idealized Guassian distribution of pollutant concentration in the centerline direction.

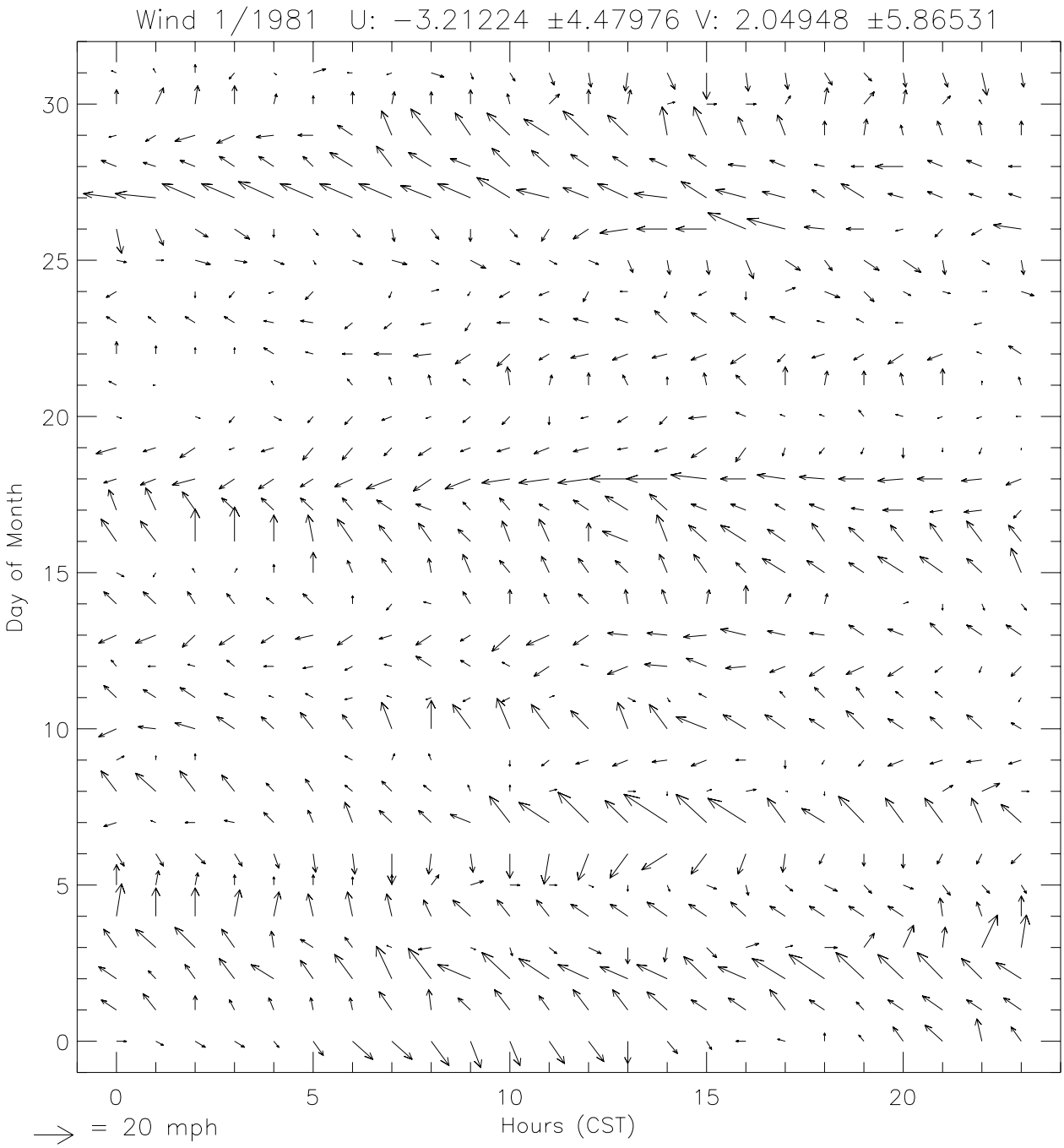


Figure 3. Minneapolis/St.Paul Airport wind data in vector form for January 1981. Arrow direction is wind direction, magnitude is speed.

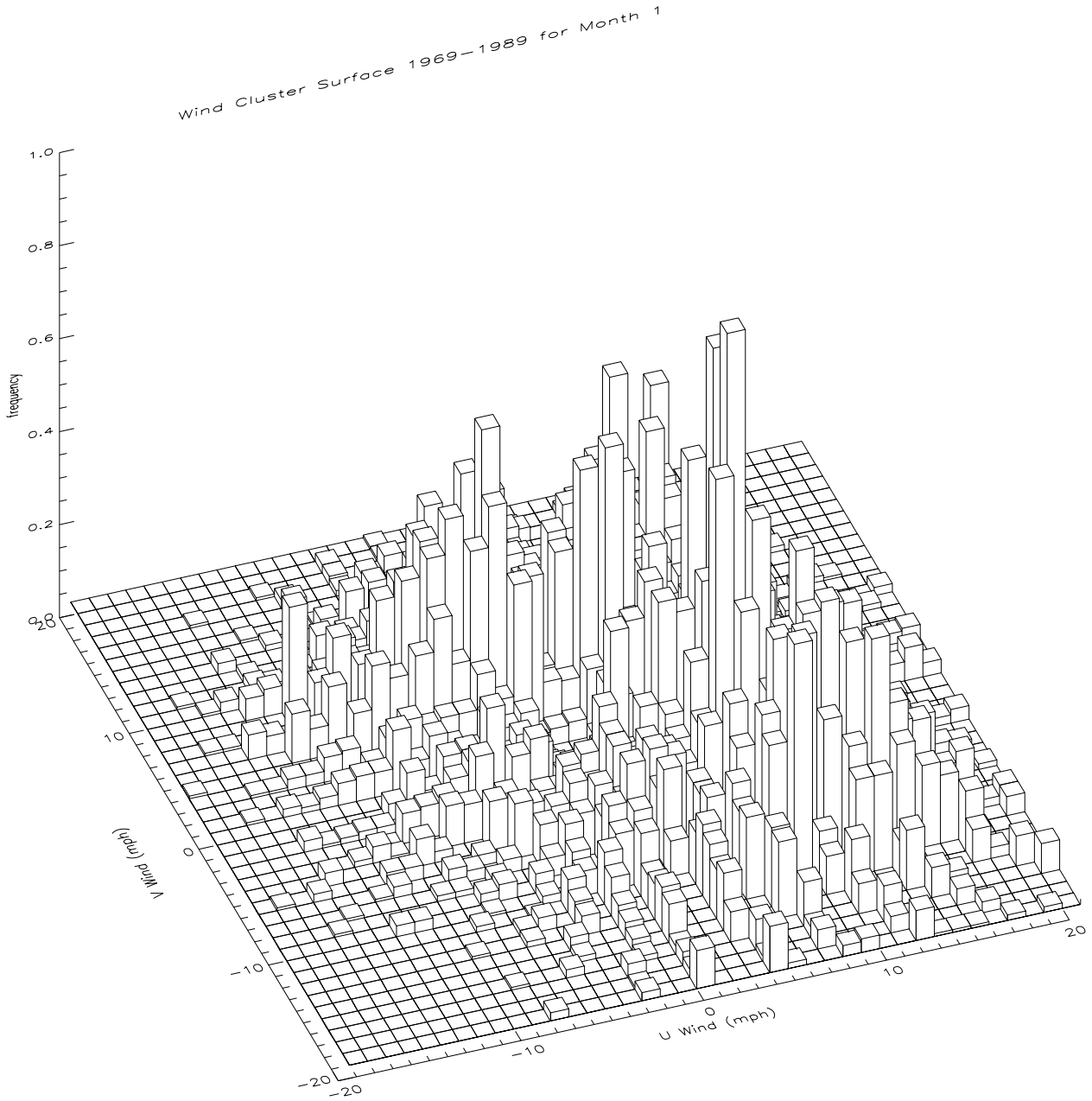


Figure 4. Two-dimensional histogram of all winds in January. X-axis is easterly component of wind, y-axis is northerly component of wind, z-axis is hours of occurence scaled to 1.0 for most common wind velocity.

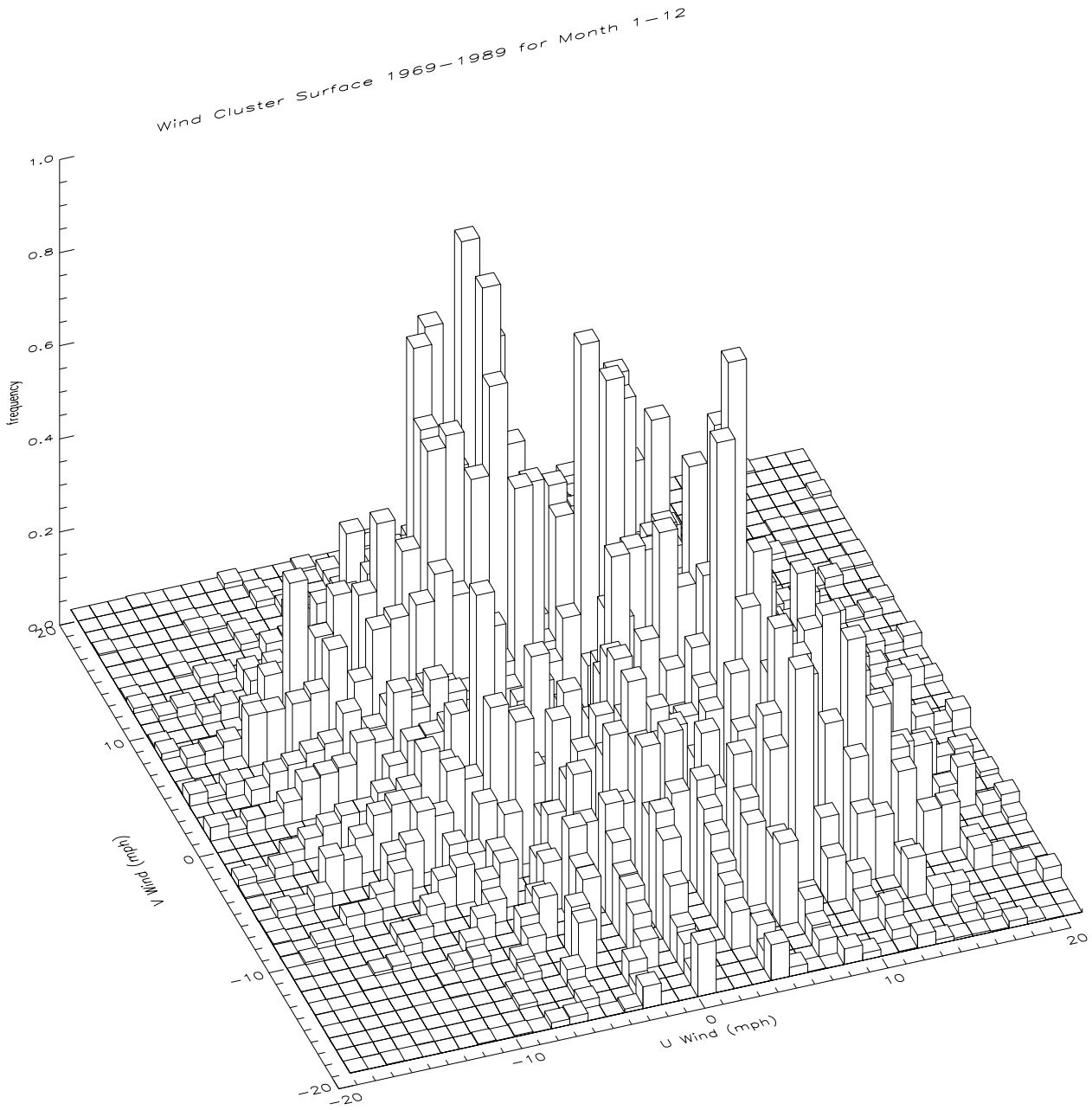


Figure 5. Same wind histogram as Figure 4 except for all months of the year.

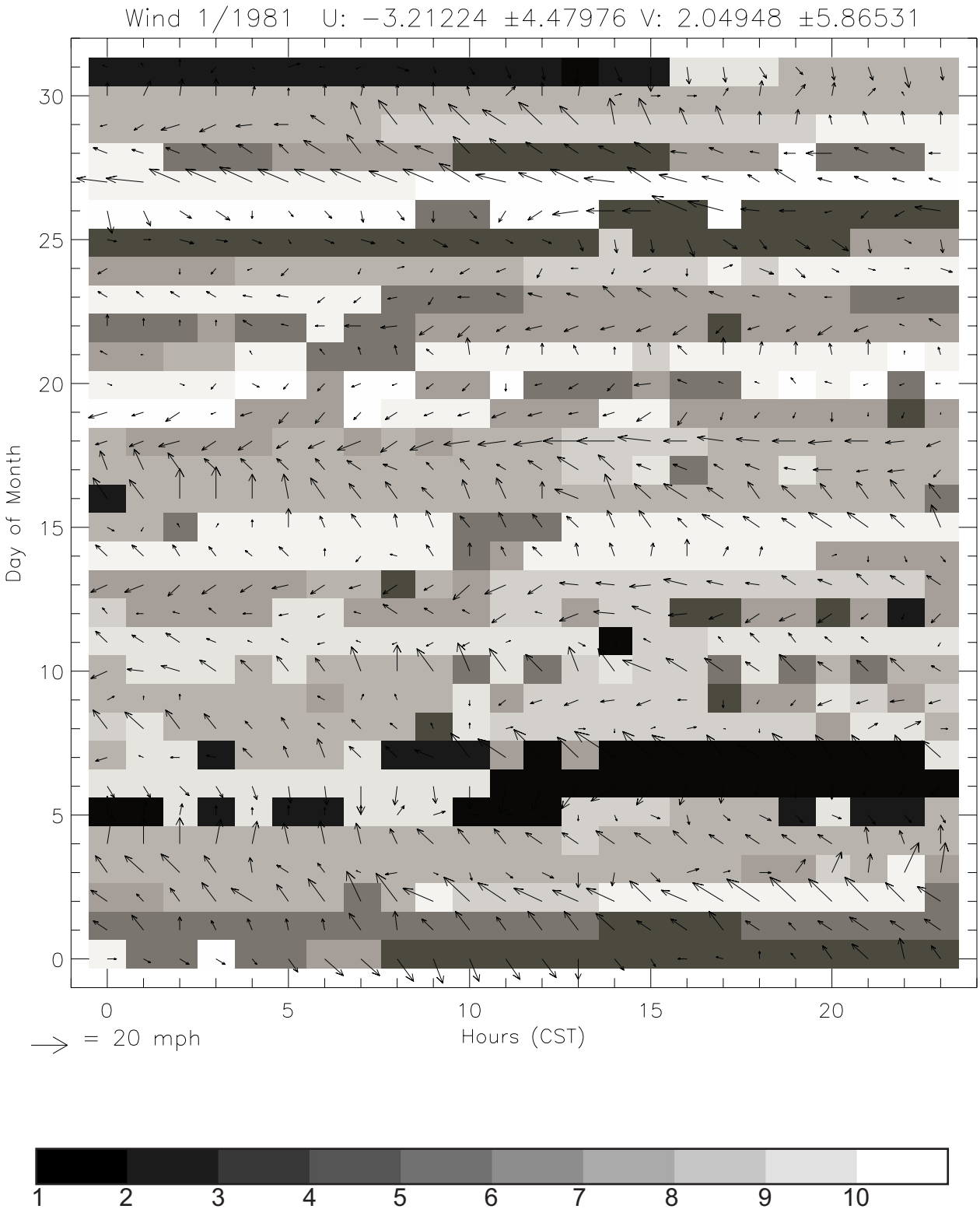


Figure 6. Same view as figure 3, except color of each box represents cluster assigned to that wind velocity (gray scale represents 10 clusters).

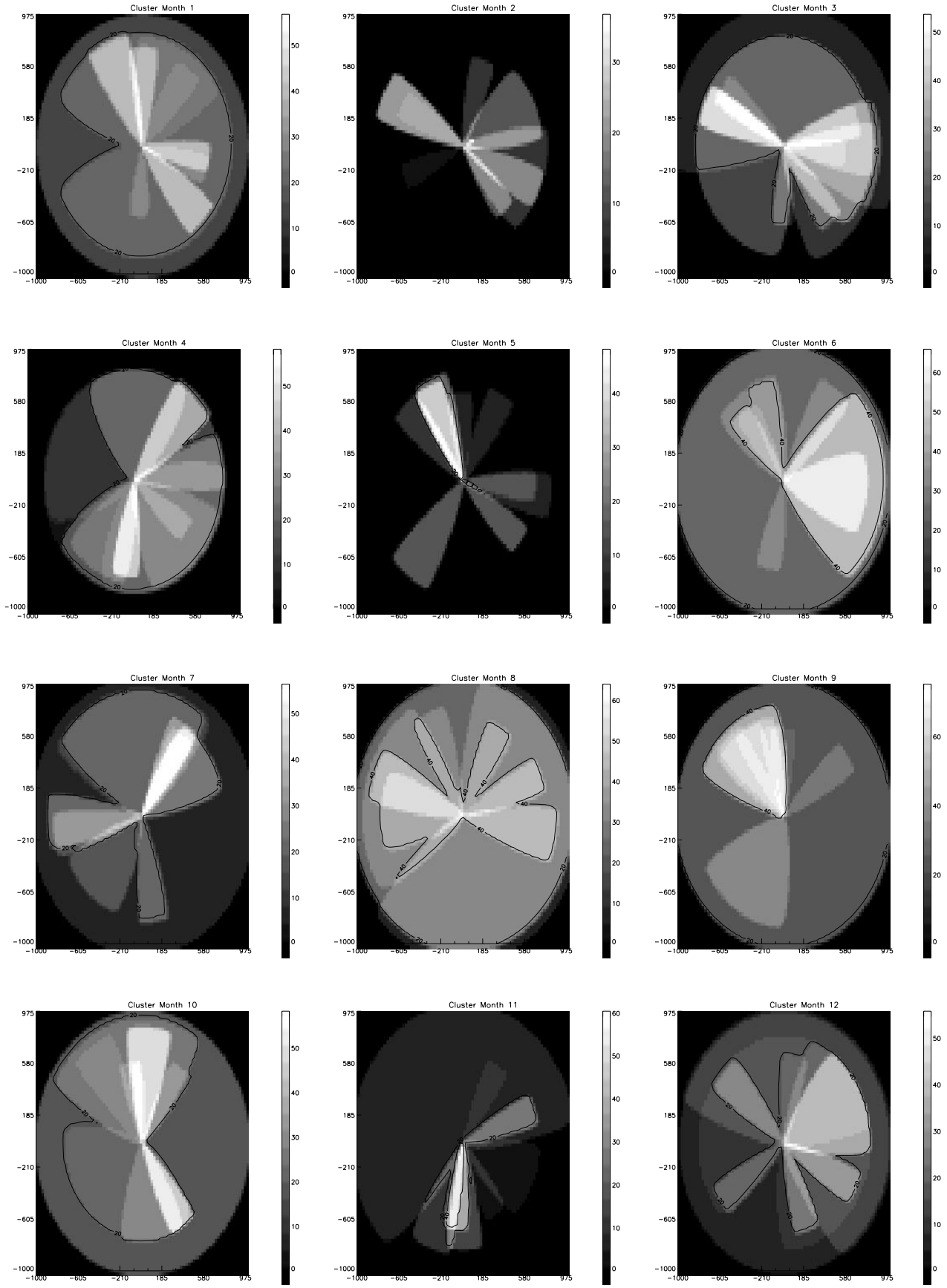


Figure 7. Plume risk buffer from clustering for each month. Whiter areas have higher risks from a pollution episode. X and Y axes are in meters.

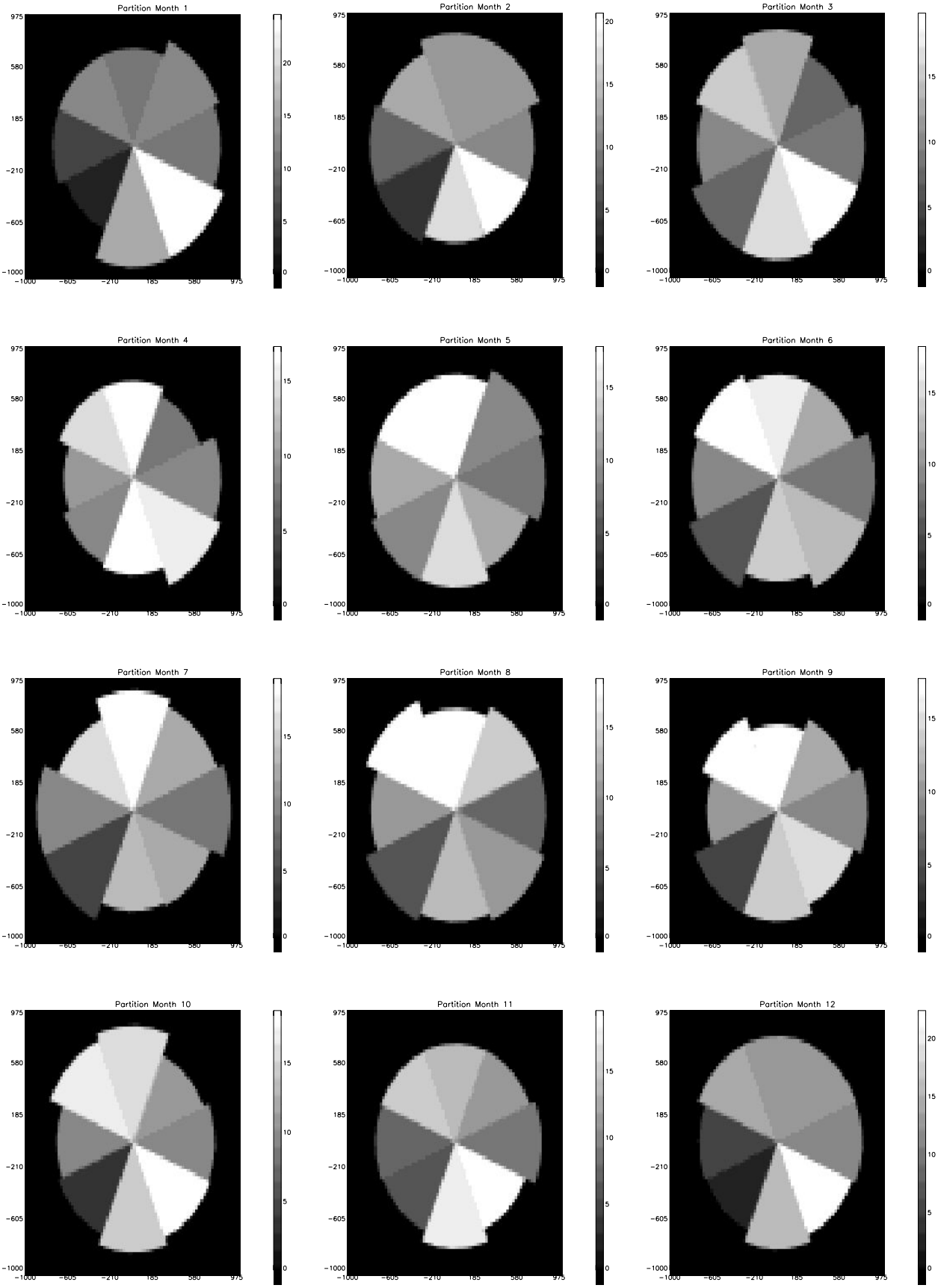


Figure 8. Plume risk buffer from pie chart compress rose partitioning for each month. Whiter areas have higher risks from a pollution episode.

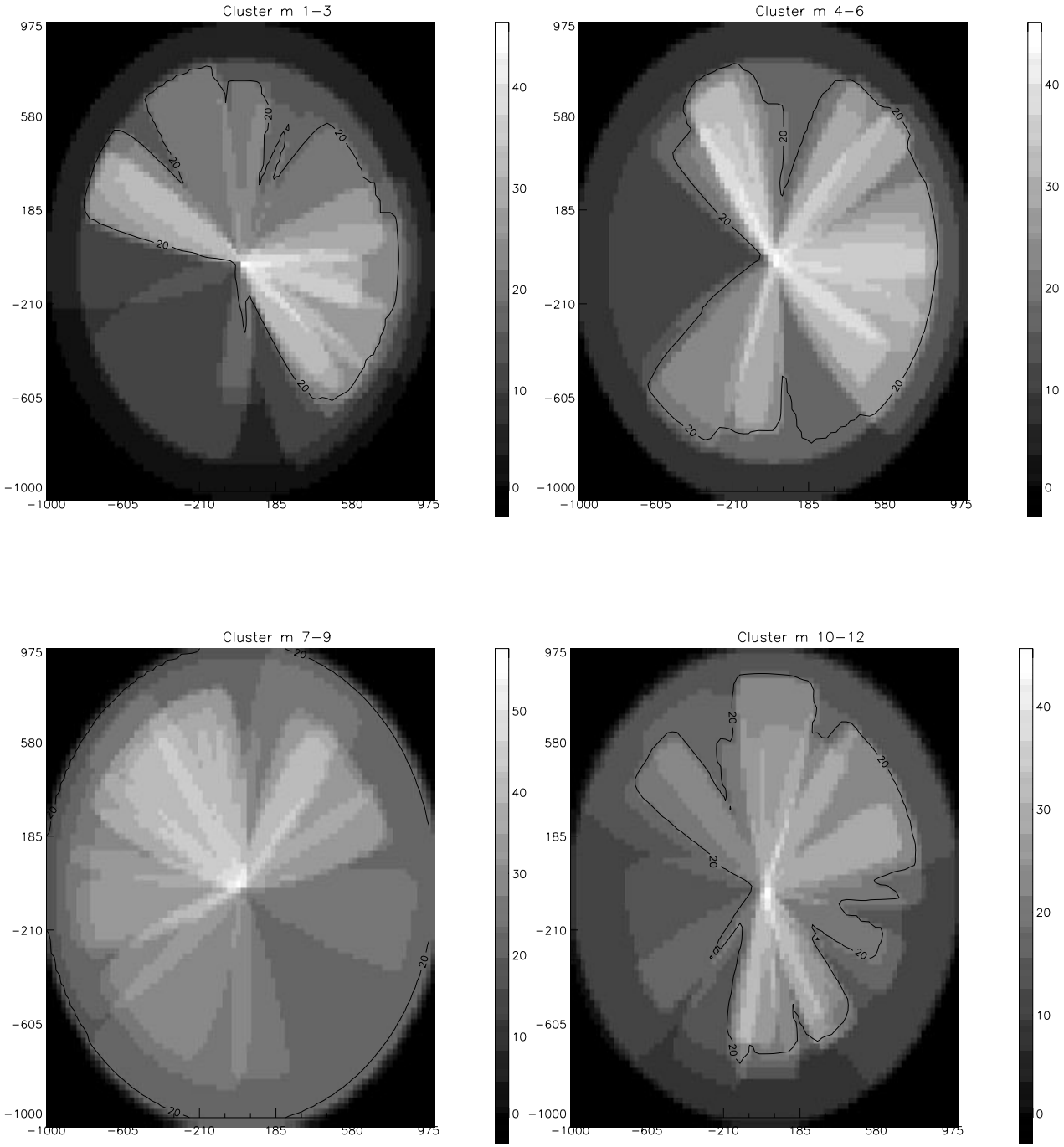


Figure 9. Seasonal aggregations of figure 7 cluster risk buffers.

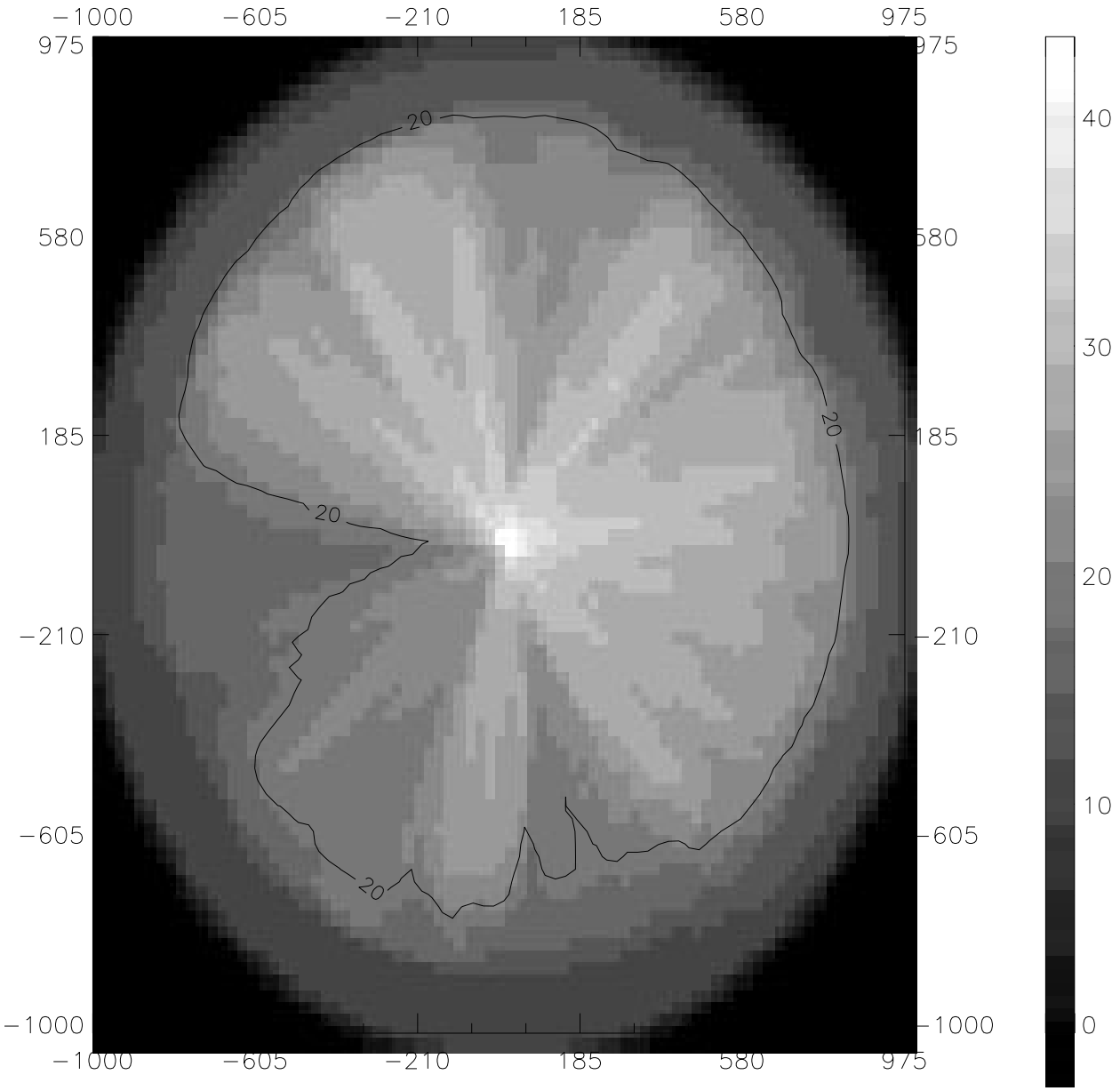


Figure 10. Annual plume buffer from aggregation of figure 9 seasonal buffers.

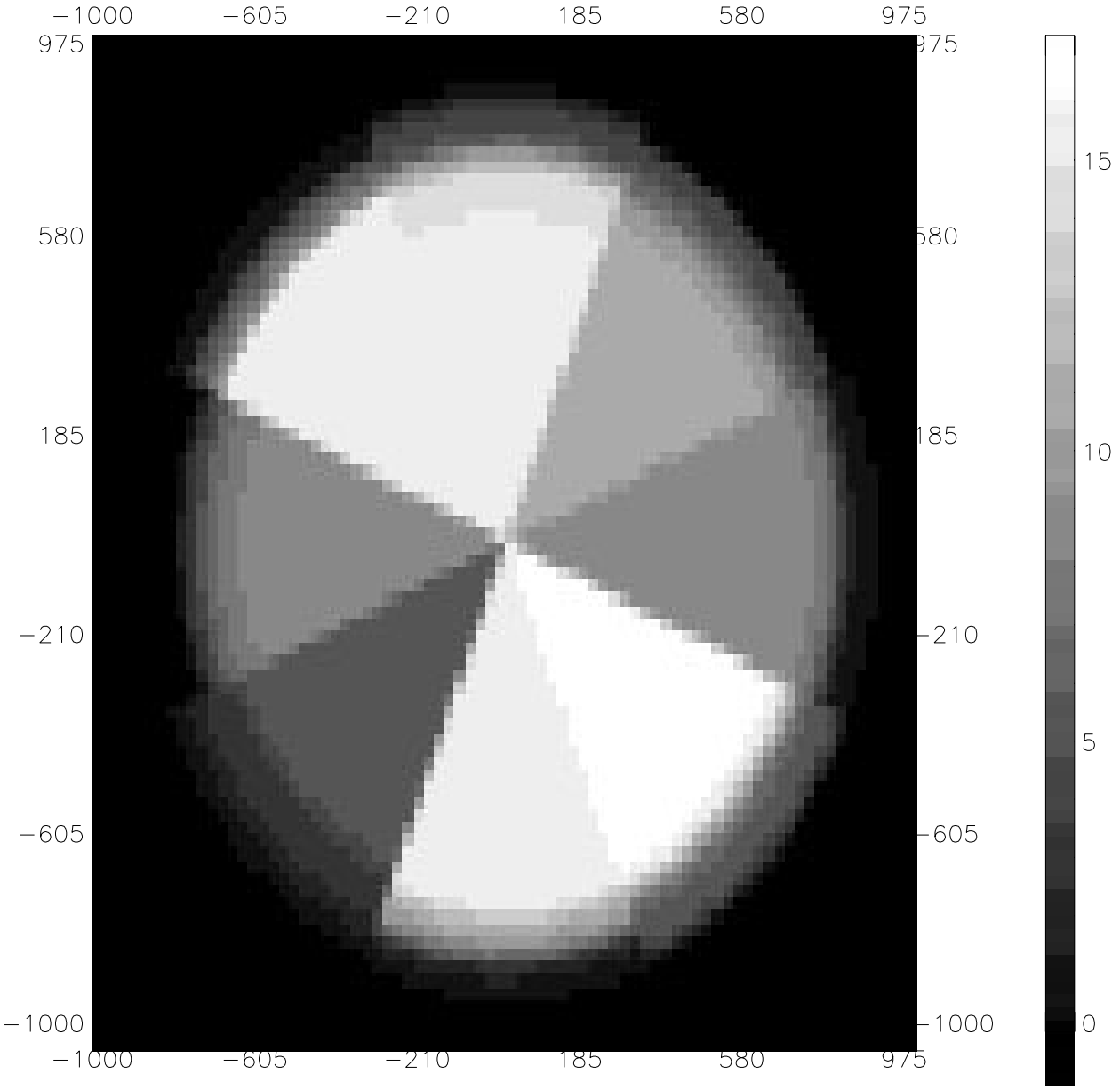


Figure 11. Annual compass rose partition plume buffer from aggregation of figure 8 monthly buffers.

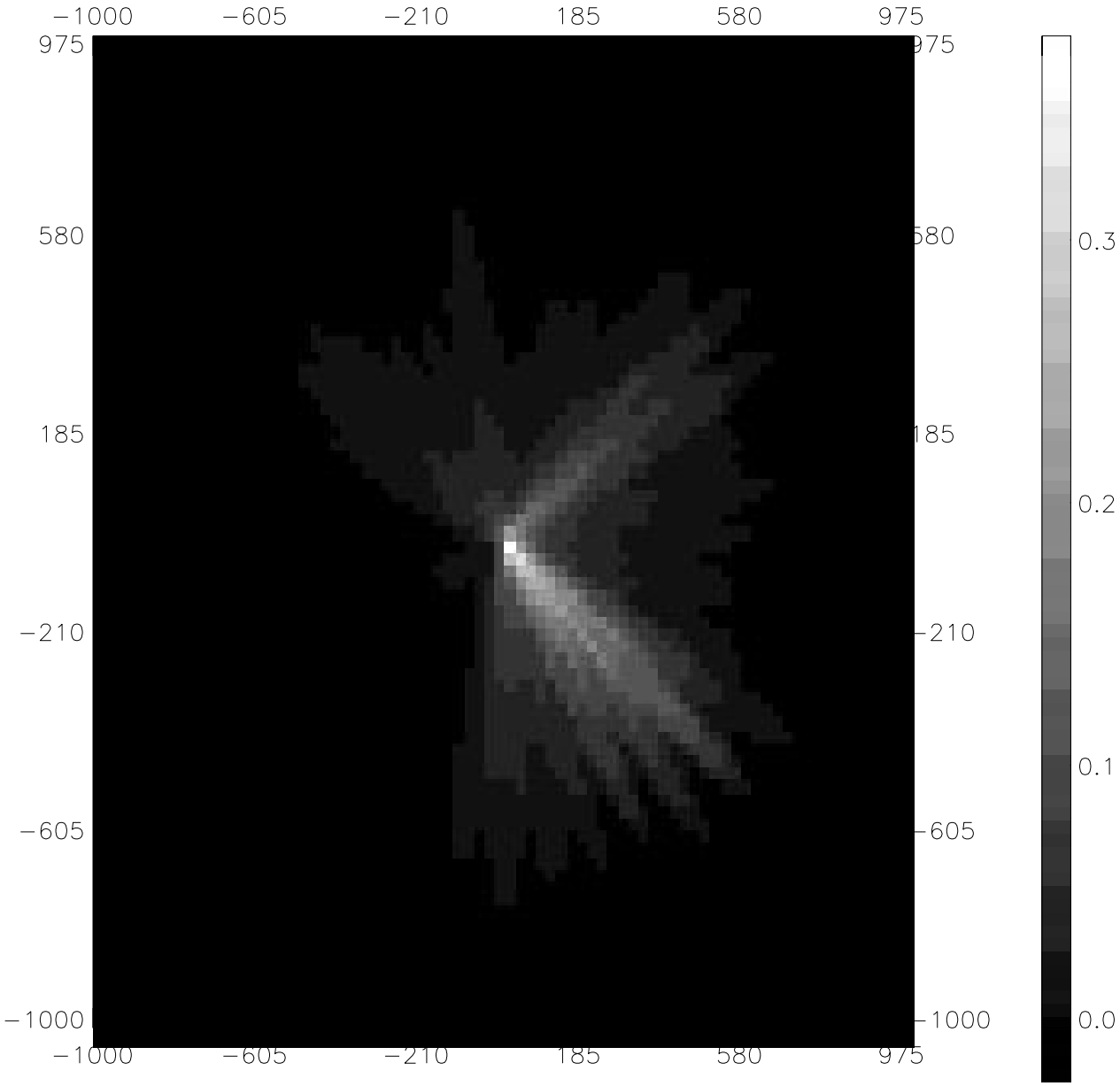


Figure 12. Every plume from January 1981 overlaid into an image to compare against the top right image in figure 7 or figure 8.

Pasquill-Gifford turbulence types for Daytime

Wind speed (m/s)	Solar Insolation Strong	Moderate	Slight
<2	A	A to B	B
2 to 3	A to B	B	C
3 to 4	B	B to C	C
4 to 6	C	C to D	D
>6	C	D	D

Pasquill-gifford turbulence types of Night time

Wind speed (m/s)	Cloud Coverage ≥ 4/8 low cloud or thin overcast	≤ 3/8
<2	G	G
2 to 3	E	F
3 to 4	D	E
4 to 6	D	D
>6	D	D

Table 1. Pasquill-gifford turbulence classes from Stull (1995)

Brigg's urban dispersion parameters for 100-10,000 meters

Pasquill stability	sigma(y)	sigma(z)
A-B	$0.32x(1+0.0004x)^{-0.5}$	$0.24x(1+0.001x)^{0.5}$
C	$0.22x(1+0.0004x)^{-0.5}$	$0.20x$
D	$0.16x(1+0.0004x)^{-0.5}$	$0.14x(1+0.0003x)^{-0.5}$
E-F	$0.11x(1+0.0004x)^{-0.5}$	$0.08x(1+0.00015x)^{-0.5}$

Brigg's rural dispersion parameters for 100-10,000 meters

Pasquill stability	sigma(y)	sigma(z)
A	$0.22x(1+0.0001x)^{-0.5}$	$0.20x$
B	$0.16x(1+0.0001x)^{-0.5}$	$0.12x$
C	$0.11x(1+0.0001x)^{-0.5}$	$0.08x(1+0.0002x)^{-0.5}$
D	$0.08x(1+0.0001x)^{-0.5}$	$0.06x(1+0.0015x)^{-0.5}$
E	$0.06x(1+0.0001x)^{-0.5}$	$0.03x(1+0.0003x)^{-1}$
F	$0.04x(1+0.0001x)^{-0.5}$	$0.016x(1+0.0003x)^{-1}$

Table 2. Briggs' sigma functions from Zannetti (1990)

Interaction of resolution and classification in high-resolution urban / suburban land cover imagery

Abstract

High-resolution urban remote sensing imagery acquired by aircraft is analyzed at multiple resolution levels to understand how resolution and classification accuracy interact. This experiment looks at 3-band 1-meter imagery from a low-density area of southwest Akron, Ohio and a high-density neighborhood in Minneapolis, Minnesota. A convolution filter was used to degrade these images to coarser resolutions. Aggregate statistics were computed on both the original images and per-pixel supervised and unsupervised classifications of the images. These statistics can be used to argue that 1.) "crossover" points, where classification accuracy of a component class decreases to below the accuracy for its aggregate class as resolution coarsens, occur in areas with high-frequency of spatial variation, 2.) traditionally grouped "classes" from hierarchical, taxonomic classification schemes tend to disaggregate in imagery at different resolutions, and 3.) component classes at high resolutions may "spectrally" be part of multiple low-resolution classes. Though none of these hypotheses are fully developed in this paper, the statistical methodology and literature review described may be of use in increasing the accuracy of classifying remotely sensed imagery.

Introduction

The current crop of satellites from both private and government sources are returning imagery at increasingly higher resolutions. These high resolutions allow for crisper images and the identification of smaller features (Veregin et al., in press). It consequently makes remote sensing of urban areas possible. Up until recently, the best publicly-available satellite imagery produced imagery at resolutions around 20-30 meters (i.e. SPOT, Landsat Thematic Mapper). While this level of resolution is adequate for land cover analyses of rural and forested areas, it is not necessarily effective for urban or suburban areas, where 30m is almost half of a city block. New satellites have been recently launched that are delivering imagery at resolutions from 1-4 meters (i.e. IKONOS). These satellites show the promise of being able to produce automated land cover maps of urban and suburban areas. Maps such as these would be useful in understanding patterns of land use change and growth in urban areas.

However, high resolution images face a problem when it comes to classification. Traditional classification, which occurs on a pixel by pixel basis, is only adequate for classifying those features on the ground that are 1/2 to 1 1/2 the size of the pixel (Jensen, 1996). In an urban area, this might include such classes as roof shingles, asphalt, grass, deciduous trees and shadows. These are not necessarily useful classes for understanding land cover in an urban area. Higher order classes such as residential, commercial, industrial, and open space require an understanding of how the detailed "component" classes aggregate into the higher order classes. It is a form of generalization (McMaster and Shea, 1992).

This is a relatively new problem for remote sensing, since images from rural or forested areas, where the frequency of spatial variation is low, are not as affected by this "class breakdown" the same way high spatial variability urban images are. This problem is not new, however, to those versed in air-photo classification or traditional land use map creation, where it is common to use taxonomic land cover classification schemes, such as the Anderson hierarchy developed at the United State Geological Survey (USGS) (Anderson, 1976). Taxonomic systems assign a hierarchy to land use classes based on level of detail. For example, the class urban may be broken down into residential and commercial, which can be further subdivided as needed. In remote sensing, hierarchical systems are sometimes used to classify rural or forested areas with a relative level of success. This paper intends to show that for urban imagery, the hierarchical classification scheme is complicated by the way component classes aggregate into their "mixed" pixels.

Aggregate classes are functions not only of the types of components, but also the spatial frequency and density of arrangement (Veregin, 1997). This is important because the same component class can be a member of multiple higher order classes. In highly variable images, it

may also be the case that the higher-order classes break down not in a hierarchical manner, but at different resolutions for different kinds of classes. A commercial class may disaggregate at a lower resolution than a residential class. This makes taxonomic hierarchical classification highly dependent on knowing what resolutions certain classes aggregate or disaggregate. There may be a point in a class where either increasing or decreasing resolution can decrease the accuracy of classifying that class. Knowing these accuracy curves could be used to identify what Veregin (1997) defines as the "crossover point." The crossover point is the resolution at which the accuracy of classifying a pixel in its component class or its aggregate higher order class is about the same. At a lower resolution, the pixel should be classified in its aggregate class. At a higher resolution, it should be classified by the component. Figure 1 shows different schematic ways Veregin (1997) understands the crossover effect. The identification of the crossover points would allow for the creation of a new kind of taxonomic classification system for high resolution urban remote sensing imagery.

This paper does not claim to come up with this system. Rather, it looks at the issues surrounding the interaction of resolution and classification in high resolution, high spatial variability images. I intend to show that the cross-over effect does exist, and occurs at different points for different classes. To do this, imagery is looked at from "both sides" of the point – the low and high resolution, the aggregate and component classes. This allows me to argue for the crossover effect by arguing two smaller hypotheses: 1.) higher order aggregate classes disaggregate at different resolutions, making hierarchical classification schemes not totally correct and 2.) high resolution component classes are part of multiple component classes and merge into the aggregate classes at different resolutions. Showing these two effects would

bolster the argument that the crossover effect occurs in urban images. A possible method for how to identify the crossover effect is suggested at the end of this paper.

Classifying Images

Before getting into the experiment, it is useful to understand the basic nature of per-pixel classification and to review the prior literature on the issue of classification, generalization and scale. The basic idea of per-pixel classification comes from multidimensional mathematical clustering. Each pixel is composed of bands which have discrete responses in different sections of the electromagnetic spectrum , i.e. red, green, blue, infrared. Pixels with similar spectral responses in its bands to other pixels are clustered together as the same class. The motivation behind this method is that similar land cover types respond with similar reflectance in the electromagnetic spectrum. The more bands, the more fleshed out the spectrum and the greater the chance of accurate classification. The clustering can follow any of a number of popular methods, including minimum-distance, k-means, or phenograms. Newer classification methods use advanced clustering methods such as fuzzy classification and neural networks. While traditional classification methods tend to classify pixels based on class statistics (mean or centroid and variance), some newer methods use Bayesian probability statistics and all pixels within each cluster to calculate class membership (Hardin, 1994). The ultimate goal of classification is to reduce within-class variance and increase between-class variance, thus maximizing classification separability and accuracy.

In supervised classification, the user identifies "training" classes in the image with known classes, based on field work or interpretation from other imagery or maps. At the high resolution scale, training classes should be spectrally meaningful to be accurately classified. Thus, physical

land cover classes such as asphalt, grass and shingle are chosen over land use classes such as residential and commercial (Veregin et al., 1995). Clustering methods then assign each pixel to one of these classes. In a minimum-distance-to-means classification, the spectral "distance" between the unclassified pixel and the means of all training classes is computed. The pixel is assigned to the class with the shortest distance. This distance measure can be simple Euclidean distance in n-dimensional space, or more complex measures. An improvement to this kind of classification is to incorporate variance information about each training class. In maximum-likelihood classification, the statistical probability that a pixel is a member of every training class is computed based on class mean and variance. The pixel is assigned to the class with the highest probability. Minimum-distance and maximum-likelihood are the two most common methods of supervised classification (Jensen, 1996). More advanced methods of clustering rely on non-linear distance measures (neural networks) and multiple class designations for each pixel (fuzzy classification) (Veregin et al., in press). Current research involves using multiple kinds of classification methods for one image at the same time and "conflating" the best classifications with heuristic rules (expert systems) (Veregin et al., in press).

Unsupervised classification does not require the user to identify training classes, but instead to select the number of classes and a threshold divergence level. Unsupervised algorithms attempt to find the optimal clustering of an image. The user then assigns a class to each identified cluster, assuming the clusters have some identifiable class. The Iterative Self-Organizing Data Analysis Technique (ISODATA) algorithm is the most common unsupervised classification algorithm (Jensen, 1996). The algorithm requires the users to select the maximum number of clusters, the maximum percentage of pixels whose class values are allowed to be unchanged between iterations, the maximum number of times to run the iterative algorithm and

the minimum number of pixels in a cluster. The algorithm starts with arbitrary clusters based on the mean and variance of each band and assigns a cluster to each pixel based on minimum-distance-to-means calculation. New centroid statistics are computed and the algorithm is run over and over again until the maximum percentage of unchanged pixels value or the maximum number of iterations is reached. Small clusters are deleted and reassigned to the nearest cluster. Often, unsupervised classifications are used to decide which classes to use in a supervised classification. The main problem in unsupervised classification is trying to decide if a given cluster is a physically meaningful land cover class and what that class is.

Classification accuracy, which is measured by comparing the assigned class of a pixel in random or non-random selected set of pixels with their actual classes, is highly dependent on resolution: spectral, radiometric and spatial. Spectral resolution is the number or width of bands and radiometric resolution is the amount of discrimination ability in each band. Both have a direct relationship between increasing resolution and increasing clustering ability, and hence classification accuracy (Toll, 1985). Spatial resolution is a different story. Spatial resolution fundamentally affects what classes the user can see and chooses. Unlike air photo classification, which is highly dependent on spatial information, per-pixel classification classifies only based on spectral response of each pixel (sometimes, spatial information is brought in with spatial filtering, but overall the information is per-pixel). Typically, what the user can classify is on the order of the size of the pixel to about one-half the size of the pixel.

Selecting classes based on pixel size is not too hard to do when pixels are "pure," that is, each pixel is comprised of one relatively spectrally homogenous land cover class. However, in urban scenes, a large number of pixels are "mixed." A pixel can be made up of components that have quite different spectral responses (i.e. road, tree, roof) than the overall scene (residential).

In this case, increasing resolution may decrease classification accuracy (Veregin, 1997). Aggregate classes like "residential" may or may not have a homogenous spectral responses among pixels classified residential. Moreover, two classes which seem like they should have maximum classification accuracy at the same resolution, such as residential and commercial, or even new residential and old residential, may have different optimal resolutions. It may be the case that different classes that fall in the same "hierarchy" of traditional classification tend to disaggregate at different resolutions.

A common taxonomic classification scheme is the Anderson classification system developed at the USGS (Anderson, 1976). The scheme was designed with both remote sensors and air photo systems in mind, but was originally based on 1:24,000 air photos of New York state. Classes are defined on one of four levels. Level I classes are broad, aggregate classes designed to be classified at coarse resolutions. The Level I classes are urban, agricultural, rangeland, forest, water, wetland, barren land, tundra and perennial snow or ice. Level II classes were designed for classification at the 1:8000 scale. For the urban level I class, the Level II classes are: residential, commercial, industrial, transportation-communication-utilities, industrial and commercial complexes, mixed-urban and other. Level III and Level IV classes were designed to be more fluid in class components and are based on using high resolution classification along with secondary information. Level III urban residential classes include: single family, multi-family units, group quarters, residential hotels, mobile home parks, transient lodgings and other. Anderson classes rely more on land use than land cover, especially for the urban classes. Remote sensors, however, are better at classifying land cover than land use. This is one reason urban classification has been difficult using taxonomic classification schemes.

The Problem of High Resolution Urban Classification

There has been some literature on this issue of mixed vs. pure pixels and resolution dependent classification. Broadly, this work falls under the realm of issues of scale in cartography and remote sensing. Most scale work in remote sensing has been more related to how scaling of remote sensing images affects biophysical models and vegetation indices (Goodchild and Quattrochi, 1997). This kind of work tends to occur more at the coarse resolution scale and on global Earth observing systems. Less literature exists on issues of scale in high resolution images over smaller areal extents.

It is important to define scale for the purposes of remote sensing since it is an abused word in geography. Scale in remote sensing applies specifically to spatial resolution (as opposed to spectral and radiometric resolution) and not to display scale (cartographic scale) or areal extent (geographic scale). Scale in remote sensing can sometimes also refer to the smallest observable object in the scene (Goodchild and Quattrochi, 1997), but for the purposes of this paper it refers to the average size of one pixel. Pixel resolution is a function of Earth curvature, sensor optics and the analog-to-digital conversion process (Jensen, 1996). The goal of understanding scale in remote sensing is figuring out what properties of scale are invariant and what measures of scale are invariant properties of geographic detail in the data (Goodchild and Quattrochi, 1997). For this paper, this goal translates to finding out how invariant is classification accuracy with scale in urban and suburban imagery.

There is a relatively small group of researchers looking specifically at scale and classification in urban imagery. Some of these studies and their methodologies are reviewed below. Most of this work began in earnest in the early 1980's. Early work included that of Monmonier (1983), who was looking at how raster data could be used to create land cover maps.

He argued that Anderson Level I classes rasterized from vector polygons tend to keep their frequency of occurrence over the image as resolution degrades. Monmonier, however, did not look specifically at remote sensing imagery and classification accuracy.

Cushnie (1987) wrote one of the seminal pieces on the issue. He looked at 5, 10 and 20 meter resolution imagery of a small town and nearby rural area in Blewbury, Oxfordshire. Cushnie used spectral plots where the brightness values of each pixel in one band are plotted against the brightness value in the other band. The results showed that a high variability residential area increased in classification accuracy with coarsening resolution by 16% at 10 m and 23% in 20 m. A low variability wheat area did not increase similarly. Cushnie argued that classification accuracy is not just a function of resolution and scene noise (variability), but also the relative spectral locations of the categories chosen. The whole map (classified with nine classes) had decreasing classification accuracy with coarsening resolution. Finally, Cushnie showed that 5x5 median smoothing filters help classification accuracy more in high resolution imagery than low resolution imagery.

Woodcock and Strahler (1987) used another statistical method to look at spatial variability in images with coarsening resolution. They computed the "local variance" of an image, which is a measure of the average size of objects in a scene. The local variance consists of computing the mean value of the standard deviation around a 3x3 kernel of each pixel. The mean of all these values is the local variance for that image at a specific resolution. They showed that for an urban image of Canoga Park, CA, local variance peaked at around 10-15 m and declined as resolution coarsened. A Thematic Mapper (TM) image of Washington, D.C. maintained high local variance from 30 m through 60 m and then started to decline. Woodcock and Strahler suggested that the high resolution image had a local variance peak at a resolution

just smaller than average object size, whereas the TM image had no distinct group of objects of specific size that dominated the scene. They also suggested that in urban images, a local variance peak, if well defined, will vary from 8-10 m for residential areas to 20-30 m for denser downtown areas. Cultural influences on building shape and size also affect this value. Cao and Lam (1997) argue that local variance essentially measures the amount of correlation between pixel neighbors, but that the validity of using local variance as a way to predict classification accuracy has not been verified.

Wang et al. (1991) came up with a novel way to aggregate classes from SPOT imagery of the Toronto area. The image was classified in seven physical land cover types: asphalt, concrete, bare ground, soil, deciduous trees, coniferous trees, low density grass and high density grass. Then, land use masks of old residential, new residential, industrial/commercial, land under construction, open space and woodland were defined as various combinations of the six land cover types. A 9x9 entity-frequency filter was used to assign each 9x9 entity as one of the six land use masks based on various statistical methods. They noted that old residential was often confused with open space. The use of their method was more for cartographic generalization than classification accuracy, but does show promise as a way to generalize land use classes given prior information about how component classes clump into aggregate classes.

On the image of suburban Akron, OH, Veregin et al. (1995) computed a low pass filter statistic to look at how individually classed pixels at the 1 m resolution converge to aggregate classes. The filter calculates the mean distance of all pixels of a certain class to the centroid of the other classes as resolution degrades. Figure 2 shows an example result from this filter. It was hypothesized that the mean distance to the centroid will tend to converge for all classes with decreasing resolution. The results verified the hypothesis.

Recently, Veregin (1997) has suggested new ways to understand resolution and classification accuracy. He has proposed the crossover effect model. The goal of this model is to find the point where classification accuracy for a component class decreases to below that of its aggregate class as resolution is degraded. Knowing where these points occur give information about what resolutions aggregate classes start to appear, and what resolutions are optimal for classifying aggregate land use classes instead of component land cover classes. This model is untested.

Resolution may not be the only factor inhibiting classification accuracy. Toll (1985) studied the classification accuracy of Landsat Thematic Mapper (TM) (30 m) versus Landsat Multi-Spectral Scanner (MSS) (80 m) in an urban fringe area of Washington, D.C.. He wrote that analysis of variance (ANOVA) statistics show the increase in spectral and radiometric resolution of TM worked more to increase classification accuracy than spatial resolution. Increased spatial resolution actually decreased classification accuracy due to reduction of within class variability. The overall result was a decrease in classification accuracy. Toll suggested using a low pass filter to degrade data before classifying urban areas.

Even with non-urban images, Foody (1994) showed that statistical models of component class frequency can be used to retrieve information at the high resolution. Foody developed a regression model between Thematic Mapper forest cover computations to AVHRR reflectance. He used his model to find a good correlation between forest cover and AVHRR reflectance. A similar kind of experiment could be done for urban areas.

All of the statistics generated above from resolution changes and classification accuracy are broadly forms of generalization. McMaster and Shea (1992) provide a hierarchical model for kinds of generalization that can occur in raster data. The simplest, structural generalization,

involves any form of spatial resampling, tessellation changes or matrix redefinition. Numerical generalization involve applying low or high pass filters to the image, creating masks, calculating band ratio indices, or applying histogram changes (contrast-stretching). Numerical categorization involves all the per-pixel classifiers described in the previous section. Finally, categorical generalization involves the merging of classified categories, the aggregation of cells, attribute generalization and dissolving features. Most of the work in resolution and classification accuracy tend to find ways of numerical generalization and categorization methods that can make categorical generalization possible. This generalization framework is handy for understanding how statistics and calculations on component pixels are used to improve classification accuracy of aggregated categories. The question of spatial resolution and classification/generalization of urban imagery will become more important as high resolution remote sensing imagery becomes regularly used.

Data Description

Data for this experiment came from airplane imagery over a low-density residential and forested area in Ohio and a dense residential urban area in Minnesota. The Ohio image is from the Mud Run Creek area of southwest Akron, OH (Veregin et al., 1995). The imagery was collected from airplane with a video camera in the red, green and blue spectral bands. Scenes were then digitally mosaiced and standardized to 1 meter resolution. The image was clipped to subcatchments since the data was collected originally for modeling stormwater runoff. Figure 3 shows subcatchment 24, the image used in this paper. This image has also been used by Veregin et al. (in press) to look at ways of increasing high resolution urban classification accuracy. The imagery has been classified at the 1 meter resolution using a hybrid classification technique that

combined the best classifications achieved with fuzzy sets, fuzzy neural networks and a modular neural network. The final classification, which the authors call "conflation," uses a set of heuristic and statistical rules to combine the three classification methods. Figure 4 shows the best classification achieved with this method.

The data for Minnesota cover the Whittier neighborhood of Minneapolis, MN at 1 meter resolution. This image was collected by Veregin (1997) for use in studying the crossover effect hypothesis. The image was produced by scanning an airphoto in the green, red and infrared bands. Figure 5 shows the original image, contrast-stretched to increase viewability. No supervised classification of this image has been done. I performed a 10 class ISODATA unsupervised classification on the image in the method outlined by Jensen (1996). Figure 6 shows the unsupervised classification of the image. Note how the classifier picked up what has previously been identified as physical land use classes in urban areas: shingles, grass, asphalt, etc...

To understand how agglomerated classes spectrally decompose, a hand masking of subsets of pixels into broad land use classes equivalent to Anderson Level I was performed by the author in a manner similar to Wang et al. (1991). Masks chosen were residential, commercial, park and road, which were the four primary urban land use classes visible in the images. For the Akron image, additional masks of agricultural field and forest were added, since they dominate a large part of the scene. For the Minneapolis image, an additional mask of high density residential was added for apartment buildings but was not used since it was overall very similar to the residential mask. Figure 7 shows the masks for Minneapolis and Figure 8 shows the masks for Akron. Only specific regions of interest were masked based on field work and airphoto interpretation.

Additional data for the Minneapolis metropolitan area was available from Landsat Multi-Spectral Scanner (MSS) and SPOT satellites. Figure 9 and 10 show the satellite images from MSS and SPOT, respectively. These images were not used in the classification analysis, but are used to validate the resolution degradation methodology.

Methodology

The first step in understanding the interaction of classification and resolution is to have imagery at multiple resolutions. The simplest method of image degradation is square (2x2, 3x3, etc…) averaging. However, this kind of averaging does not accurately simulate the response of a remote sensor (Justice et al., 1989). Simple averaging assumes that the sensor has a "square wave" response to incoming radiation, which is not the case (Woodcock and Strahler, 1987). A better filter would take into account how a sensor physically scans reflectance in the along and across track direction. Sadowski and Sarno (1976) developed a 5x5 convolution filter to degrade high resolution airborne multispectral scanner data to 80 meter resolution Landsat Multi Spectral Scanner (MSS). This filter is asymmetric in the across track direction to account for how MSS works. This is same filter used by Cushnie (1987). Justice et al. (1989) designed a comparable filter for degrading MSS data to even lower resolutions using a 7x7 filter based on the inverse Fourier transform of a sensor transfer function. This filter was compared and shown to be mostly similar to the 5x5 filter.

In this experiment, the Minneapolis image was degraded from 1 meter resolution down to 1024 meters in powers of two (2, 4, 8, 16, 32, …) using the method developed by Sadowski and Sarno (1976) and outlined by Cushnie (1987). Prior to degradation, a histogram equalization contrast stretch was performed on the Minneapolis image since it was relatively dark. The

Akron image had too few pixels to degrade all the way down to 1 kilometer and so was degraded only down to 64 meters. No contrast stretching was necessary on this image. Figures 11 and 12 show the Minneapolis image and the Akron image at 32 meters resolution, respectively. Figure 11 has similar variability with the SPOT image of Minneapolis in Figure 10, though Figure 11 tends to have a bit more noise and lower overall brightness values (less blue and more black).

Since both the Akron and Minneapolis image have a high amount of correlation among the three bands, a principal components transformation was applied before calculating some of the image statistics and for making it easier to visualize these statistics in two-dimensional graphs. The goal of a principal components transformation is to translate and rotate the original axes of the data (centered at 0 and orthogonal to each band). The new axes are orientated such that the first-axis (principal component 1) has its origin at the centroid of the data set and is rotated in the direction that represents the maximum variance of the data. The remaining principal components are orthogonal to this first axis (Jensen, 1996). There are as many principal components as the number of bands, but typically the first two components can explain the variance and hence separability of 95% of the data (Jensen, 1996). Figure 13 shows a schematic from Jensen (1996) of how principal components works in two dimensions. The calculation of the principal components is essentially an orthogonalization of the matrix of data. This is equivalent to calculating the eigenvalues and eigenvectors of a matrix. The diagonal matrix of eigenvalues is the equal to the original matrix multiplied by the eigenvector coefficients (Kolman, 1991). Fuller treatment of the principal components transformation is available in Jensen (1996).

Once the image has been classified, masked, contrast fixed, degraded to multiple resolutions and axes transformed, statistics of classification can be performed. Various statistics

are performed on both the high resolution classified classes and the aggregate masked classes, including several previously described. Briefly, each method is described here:

Statistics on component classes

The component classes (the unsupervised 1 meter classes in the Minneapolis image and the classified 1 meter classes in the Akron image) are assumed to be "pure" classes, which appears to be a reasonable assumption in the case of urban and suburban imagery. These classes were plotted out spectrally with the axes as principal component 1 vs. principal component 2. The color of each dot was set equal the original classified color of the corresponding pixel in the image. Thus, each pixels' first and second principal component were computed. It was then plotted as a dot on a graph with the principal components as axes. The color of this dot was set to the color of its classification (green for deciduous, etc…). To aide viewability of this plot, the dots were rasterized onto a grid and the intensity of each grid cell was set equal to the number of dots that fell in each grid cell, based on a logarithmic scaling. Additionally, each set of dots of the same color were separated out in photo editing software and transparency levels increased so that a grid point with more than one class type of dots would still show all kinds of dots. This spectral plot was done for each resolution. The goal of these plots is to visualize how the individual component classes are merging and separating with coarsening resolution.

To get a quantitative measure of how component classes are merging with resolution, divergence statistics were calculated. Divergence statistics involve calculating how "separable" two classes are from each other based on mean and variance statistics of the brightness values within each class. Jensen (1996) reviews two basic divergence statistics and notes that they are relatively well correlated to classification accuracy.

The first measure computed is called transformed divergence. Transformed divergence first requires calculating the untransformed divergence of two classes, which relies on the covariance matrix of each class. The covariance between two bands is calculated as:

$$\mathrm{cov}_{kl} = \frac{\sum_{i=1}^{n} (BV_{ik} - \mu_k)(BV_{il} - \mu_l)}{n-1} \quad (1)$$

(Jensen, 1996)

where the BV_{ik} is the brightness value of pixel i in band k, μ_k is the mean of all values in band k, and n is the total number of pixels. The numerator is the squared sum of products between two bands. The covariance matrix is the n x n matrix consisting of the covariance between every pair of bands combinations for all pixels assigned a specific class. This covariance matrix is used in the divergence calculation. Divergence between class c and d is calculated as:

$$Divergence_{c,d} = \tfrac{1}{2} tr\left[(V_c - V_d)(V_d^{-1} - V_c^{-1})\right] + \tfrac{1}{2} tr\left[(V_c^{-1} + V_d^{-1})(\mu_c - \mu_d)(\mu_c - \mu_d)^T\right] \quad (2)$$

(Jensen, 1996)

where, V_c is the covariance matrix of all pixels assigned to class c, μ_c is the mean vector of class c and tr is the trace function (sum of diagonals). The divergence, however, varies among classes and images. The transformed divergence rescales divergence onto a logarithmic scale of 0 to 2000. Transformed divergence is calculated as:

$$TDiver_{c,d} = 2000\left[1 - e^{\frac{-Divergence_{c,d}}{8}}\right] \quad (3)$$

(Jensen, 1996).

Another similar measure of divergence is the Jeffreys-Matusita distance, which is based on the Bhattacharya distance between two classes. The Bhattacharya distance is calculated as:

$$Bhat_{cd} = \tfrac{1}{8}(\mu_c - \mu_d)^T \left(\frac{V_c - V_d}{2} \right)^{-1} (\mu_c - \mu_d) + \tfrac{1}{2} \ln \left(\frac{\left| \frac{V_c - V_d}{2} \right|}{\sqrt{|V_c|}\sqrt{|V_d|}} \right) \quad (4)$$

(Jensen, 1996)

where, μ_c is the mean vector of class c, V_c is the covariance matrix of class c, and |X| is the determinant of matrix X. The Jeffreys-Matusita distance is a saturating transform for Bhattacharya and scales it from 0-2. It is calculated as:

$$JM_{c,d} = \sqrt{2\left(1 - e^{-Bhat_{c,d}}\right)} \quad (5)$$

(Jensen, 1996).

Another way to look at component aggregation in the images is to look at coarse resolution pixels and calculate statistics based on the types of classes that were in the original high resolution classification that a coarse resolution pixel encompasses. This is a form of categorical generalization. Two kinds of statistics are developed in this manner, the pureness statistic and the aggregate statistic. The pureness statistic calculates the percentage of pixels in each class that remain "pure" with coarsening resolution. A pure coarse pixel is one in which all the 1 meter components are the same class. This kind of classification can be expanded to included those pixels that are not 100% of one class, but are still dominated by that class. This dominance is broken down into totally pure (100% same class), mostly pure (dominant class is > 50% of component classes), mostly mixed (dominant class is > 25% of component classes), and totally mixed classes.

The aggregate statistic breaks down the pureness statistic further. This statistic calculates what percentage of pixels within each component class falls into various 2x2, 4x4, 8x8, etc... aggregations of pixels that are dominated by one or another component. For example, one may

find that in the 1 meter image, analyzed by 32x32 clumps of pixels, 20% of all pixels classified shingle at 1 meter are in 32x32 clumps that are mostly deciduous and 40% are in clumps that are mixed but still deciduous dominant. In theory, this might suggest that somewhere between 20% - 60% of shingle pixels would most likely be misclassified as deciduous at 32 meters resolution. The aggregate statistic is computed for all components, for clumps from 2x2 to 32x32. Since the pureness and aggregate statistics make more sense with well classified components as opposed to unidentified unsupervised classified components, these two statistics were only computed on the Akron image.

Statistics on aggregate classes

Similar kinds of statistics were computed on the aggregate masked classes as were the component classes. Both spectral plots of the principal components and divergence statistics of the aggregate classes were computed at each resolution. Additionally, the spectral plots were transformed into spectral ellipses to understand class divergence better. Ellipses were drawn around two standard deviations around the centroid in each class. The major axis of the ellipse is in the direction of maximum variability of pixels in a specific aggregate class and the minor axis is orthogonal to the major axis. Thus, the ellipses are a way of computing principal component subsets (for each class) of the image.

Aggregate and pureness component statistics do not have an analog for the aggregate classes. Instead, another measure, a measure of spatial variability of spectral responses is calculated for each aggregate class at each resolution. The measure used is local variance, as described by Woodcock and Strahler (1987). Local variance computes the standard deviation around each 3x3 kernel of every pixel and takes the mean of all the standard deviations. Local

variance was designed for used on one band imagery and so for this experiment, local variance is computed for principal component 1. Unlike Woodcock and Strahler, local variance is computed for each masked aggregate class in addition to the entire image.

Other Statistics

An ideal statistical measure of the interaction of classification and resolution would incorporate information of both the component and aggregate classes. One visual method used in this experiment involved overlaying the spectral plots of component classes on top of the masked aggregate classes spectral ellipses. This shows how aggregate classes move in spectral space with respect to the spectral locations of the high resolution component classes. A quantitative measure of this is the component statistic. This statistic measures the percentage of component classes that are a member of the aggregate classes. In this experiment, this statistic is only computed for the 1meter component classes within the masked aggregate classes. In theory, this statistic could be extended to the other resolutions, too.

Obviously missing from the lists of statistics is plain old classification accuracy. Unfortunately, classification accuracy is a labor intensive statistic. It involves identifying training classes at each resolution, computing a classification at each resolution, identifying the class of a randomly selected set of pixels at each resolution, calculating the errors of commission and omission for each class in each resolution and finally aggregating those statistics into an overall measure of classification accuracy. Future work in this project will involve computing this classification accuracy. The goal of this paper is to develop statistical methods to understand classification and resolution interaction that do not necessarily require a full classification

accuracy test to be done at every resolution. As such, the methodology described above serve as proxy measures for understanding classification accuracy.

Results

From the methods described above, it appears obvious that there are quite a few graphs to consider. The goal is to understand how all these graphs tell part of the story of how classes intermingle with resolution. The results will be presented looking at the aggregate (masked) classes first, then the interaction of aggregate and component (1 meter) class, and finally statistics on just the component classes.

Figures 14 and 15 show the spectral and ellipse plots for the aggregate classes of Minneapolis and Akron, respectively. The ellipses represent two standard deviations around the centroid of each class while the spectral dots represent the brightness value (scaled to 1.0) in principal component 1 on the x-axis and principal component 2 on the y-axis. Overlaid on top of each other and repeated from high to low resolution, these spectral plots form the "base" of the interaction story. They show how "different" aggregate land use classes are from each other in land cover reflectance at different resolutions. The dots show the actual values of the various components within a class while the ellipses show the average value and variance of their reflectance. One obvious problem with the whole experiment becomes apparent from these graphs. The total number of pixels decreases with lower resolution, and so the spectral/ellipse plots at coarse resolutions show low variance by default. With so few pixels, separability between classes becomes a function of small changes in location of one or two pixels. A future experiment might try using a much larger dataset – such as a mosaic of airphotos, or incorporation of Landsat-like remote sensors. For the purposes of this experiment, it is assumed

that the pixels plotted at the coarse resolution mimic the actual average variance of all pixels in classes at that resolution.

The spectral plots show that aggregate classes move around a bit in spectral space. In figure 14, it appears that at 8 meters and again at 512 meters, the classes appear to be most separable, though not by much. The lower variance park class appears to always be more separable than other classes. It also appears that the commercial class separates out farther than the residential or road class. Figure 15 also shows residential and road not separating out at any particular resolution, while the commercial complex class maintains a wide spectral variance among various components, even past 8 meters. The water class remains separable at high resolutions, but appears to mix more at coarse resolutions.

Quantitative measures of aggregate class separability are available from the divergence statistics. This experiment uses transformed divergence and the Jeffreys-Matusita (JM) distance. Figures 16 and 17 show transformed divergence and the JM distance for Minneapolis, respectively, while figure 18 and 19 show the same for Akron. The graphs are plotted as log base 10 of resolution (since experiment values were tested in logarithmic base 2 increments – 2, 4, 8, 16, etc...) versus divergence value. The uppermost divergence value (2000 for transformed, 2.0 for JM) represents maximum separability. Each graph shows the separability of one class against all other classes. The color of each line is the same as the color mask. The thick black line in each graph is the average separability for that class against all other classes. The results appear to show that transformed divergence is more easily fooled by the lowering of variance at higher resolution due to smaller number of pixels in the sample size. However, both measurements tend to have similar curve shapes. Except for road, the Minneapolis average curves (figures 16 and 17) all appear to peak around 128 m before either leveling off or actually

decreasing in the case of commercial. The residential class has a mini-peak around 64 m. The Akron classes do not show any characteristic peak of divergence. Both transformed and JM divergence stay relatively constant with resolution with a slight increase in divergence occurring at the coarsest resolution possibly due to the small number of pixels. The Akron data only goes out to 64 m, so it is possible that at higher aggregations (with enough pixels), some peak value may occur. What the results do show is that more homogenous classes like water and deciduous have high class separability at all resolutions, while variable classes like commercial and residential slightly increase in divergence with coarsening resolution.

To get a quantitative handle on average component size within the aggregate classes, the local variance was computed. Figures 20 and 21 show the local variance for Minneapolis and Akron, respectively. In both cases, more variable classes tend to have a signature peak at some resolution, typically between 4-16 meters. The more urban Minneapolis image has an overall local variance peak around 8 meters, while the more suburban Akron image was a mini peak closer to 32 meters. The overall local variance on most classes is lower on the Akron image than the Minneapolis image. It also appears the both the Akron and Minneapolis residential classes have two peaks, one around 4-8 meters, and another appearing at a coarser resolution. The second peak might be indicative of when the residential class is surrounding entirely by non residential class pixels (large number of "border" pixels). Also, as classes become more mixed, even low variability classes like Field, Forest and Water have a jump in local variance as the number of border pixels of different classes increase.

To get a handle on how the components affect these aggregate classes, the next set of figures look at how the components aggregate over resolution. Figure 22 shows the percentage coverage the various component classes have within each of the masked aggregate classes for

both images. For Akron, this figure shows the pureness of the deciduous and water classes compared to the mixed classes of field, residential and commercial. The field class has more stubble and grass, while the commercial class has more concrete and the residential class has an equal mix of many classes. The Minneapolis components graph shows less obvious differences between the masks, partly because the component classes are simple unsupervised, unlabelled classes. It appears that areas with more buildings (high density residential) have more shadow / road, parks have more trees and grass, and roads and residential areas have almost the same set of components.

Figures 23 and 24 show the spectral graphs of the components. They also show that the components have more meaning and coherence over degrading resolution in the Akron image versus the Minneapolis image. In these figures, the ellipses of the aggregate classes are still plotted, but the color of the dots underneath are representative of the component classes instead of the masked aggregate classes. Figure 23 (Minneapolis) shows very clearly the effect of the minimum distance classifier at the 1 meter resolution. The 1 meter Akron image in figure 24 shows how a hybrid supervised classification creates differently shaped and overlapping classes. The components in figure 23 aggregate quickly, to only a few classes at 2 meters, and quite mixed from then on. The only except is the dark green class, class 8, which maintains an unique signature at least through 8 meters. In figure 24, most of the classes stay together at least through 8 meters. It appears that water and deciduous classes tend to stay as components even past that, while the remaining classes quickly blend together.

Figures 25 and 26 show the transformed and JM divergence for the Minneapolis components, while figures 27 and 28 show the same for Akron. The Minneapolis divergence graphs seems to suggest that aggregation occurs at even earlier resolutions than the spectral plots

suggest. Also at 128 meters resolution, some of the classes have a very slight increase in divergence, especially class 1. The Akron divergence graphs show how Akron components tend to stay divergent for a longer time than the Minneapolis components. They also show how water and surprisingly stubble are generally more divergent from the remaining components that other classes.

The final set of figures look at the results of the categorical generalization of the Akron component classes. Figure 29 shows the pureness generalization for each component class. It shows the percentage of all pixels at each resolution that are dominated by one class. The graphs show how those aggregations that are mostly pure deciduous dominate the image at coarse resolutions, suggesting that at a Landsat like resolution, low density suburban areas might be confused with forested areas. Classes that are involved in mixed areas, such as asphalt and concrete are less likely to be the dominant component in any pixel aggregation, but do show up in similar percentage mixtures (minus the pure class) at coarser resolutions. Water is interesting since at 16 meters it switches from a trend of increasing mixing to more mostly pure classes. The grass component is also interesting in how it increases in dominance at 16 meters. This may suggest that grass is best classified at 16 meters. The pureness graphs also validate the results of local variance. They show that most components lose pure areas around 8 meters, which is the same location that local variance for aggregate classes peaked.

The aggregate statistics in figure 30 are an expansion of the pureness graphs and are consequently more difficult to interpret. The graphs show what components are making up each kind of aggregation shown in the pureness graph (pure deciduous, mostly pure deciduous, pure water, mostly mixed shadow, etc…). For example, the graph titled Deciduous Dominance shows how a growing percentage of pixels that are of the member class asphalt at 1 meters exist in

classes that are mostly deciduous at coarser resolutions. At 2 meters, around 5% of all asphalt pixels are in classes that are mostly pure deciduous and another 6% are parts of aggregations that are mostly mixed deciduous. With coarsening resolution, more asphalt pixels are part of conglomerations that are mostly deciduous. The graphs show how most classes increase in likelihood that they are in mostly deciduous aggregations. Meanwhile, classes like shadow and shingle disappear, and stubble mixes into grass, while concrete mixes into asphalt, all typically after 8 meters resolution.

Discussion

What did the preceding 30 figures have in common? They all basically tried to show different dimensions of how changing the resolution of an image affects how "pure" component land cover classes tend to mix and aggregate into characteristic land use classes, and how "mixed" aggregate land use classes similarly disaggregate in different ways. The results show that the reality of this claim is complicated by a large amount of variation in urban and suburban images, and the lack of a necessary unique spectral response for different land use classes, or even component classes. Within the noise, however, there are a few results that are interesting.

The first of these is how urban component classes began to aggregate around 6-10 meters resolution, as evidenced by the ellipse plots, aggregate divergence, local variance and the pureness statistics. This was the case even for the suburban residential area of Akron and the urban residential area of Minneapolis, despite how different they look from the air. Road and commercial classes appear to have a similar component aggregation peak in both images.

The small number of pixels at coarser resolutions prevent any real understanding of what is happening to the aggregate classes. The components statistic showed that especially for

Akron, different aggregate classes had unique components. But the spectral and divergence figures showed that these components quickly lost their unique signature at coarser resolutions. Moreover, the unique signatures do not appear to meld into a unique aggregate signature that is different from other urban classes (i.e. separating residential from commercial). This may suggest that a multi-resolution hybrid classification method may be the best way to classify high spatial variability urban land use classes.

What of the crossover effect? The results do show component classes aggregating after 8 meters resolution. They don't necessarily show them aggregating into various land use classes. Other than for spatially pure areas like forests and water areas, the component classes mix spectrally in relatively similar ways. This may be a function of the small amount of data at the coarse resolution, or it may suggest that there is no one optimal urban land use classification resolution. Consequently, the point at which classification accuracy of a component falls below that of classification accuracy for the aggregate class may never occur in this experiment, since both probably decline rather quickly.

Future work with this data will require doing a classification accuracy analysis at high, low and intermediate resolutions to understand the crossover effect. This paper only lays out the groundwork for that. The paper did show that land use classes with high variability (residential, commercial) have their components aggregating at relatively high resolutions (8 meters). They potentially have a second peak at resolutions near 100 meters. With a larger data set, it might be possible to show that component classification accuracy peaks around 10 meters and aggregate class accuracy peaks around 50-100 meters. Thus, the crossover resolution occurs somewhere in between these two resolutions, perhaps right around the resolution of current Landsat satellites (20-30 meters). The component class spectral plots overlaid with aggregate class ellipses

(figures 23 and 24) show what appears to be the largest level of confusion in either sets of classes around that resolution (i.e. the ellipses separate out more at coarser resolutions, while the components separate out more at finer resolutions).

The divergence plots clearly show that the effect described above is far more prominent in high variability classes than low variability classes. Deciduous and water component classes, and forest and water aggregate classes maintained high separability at all resolutions, thus suggesting that there is no real crossover effect in relatively homogenous classes, at least for the resolutions examined in this experiment. This can best be understood from the components graph (figure 22) which clearly shows that homogenous classes are made up of far fewer components than heterogeneous classes. Thus, the component and aggregate classes essentially have the same spectral response.

While the results hinted at a crossover effect and clearly showed how component classes can be parts of many aggregate classes, they do not easily show the hypothesis that hierarchical classification schemes may fail in urban images. The results showed that all the high variability aggregate classes had local variance peaks around 8 meters. More data might show the location of the second peak. A resolution in between those two would form the area where a taxonomic classification scheme may move from classifying components to classifying aggregations. The results did not show, for example that commercial and residential classes have definite peaks at different resolutions from each other. A larger data set might have looked at a greater variety of land use types (industrial and commercial complexes, downtown commercial areas, different kinds of roads) and might be able to show a claim otherwise, since it intuitively makes sense that larger building complexes or narrower roads would ultimately have different peaks of local variance and optimal classification resolutions. A shopping center has different sets of

components (homogenous area of road and rooftop, some water and grass) compared to an urban commercial strip (highly variable area of road and rooftop, less water and grass).

The images used in this experiment are also subject to other kinds of errors. Clearly, the Landsat and SPOT images of the Minneapolis area show that the degradation methodology did not produce the exact same results. This is primarily from the low contrast and low brightness values in that image. The Akron image suffers from spectral variation among the various mosaics. The images either were scanned airphotos or digitized video and so spectral separation is a function of scanner spectral resolution.

In both images, the classification at 1 meter was assumed to have occurred on "pure" classes and 100% classification accuracy was assumed. Obviously, neither is the case. This assumption may affect how the component classes aggregated. It was also assumed that the randomly chosen hand-masked aggregate classes were representative members of their classes as a whole. Further research should check that the members of any one aggregate class are in fact similar land use types. Otherwise they should be separated out (i.e. two residential masks may be separated if one comes from a different kind of residential area than the other – high density vs. low density).

Despite the assumptions described above, and others (i.e. divergence and local variance measures are representative proxies of classification accuracy and the first two principal components are representative of all three bands), the results showed the complications that arise in classifying high resolution urban images. The ability to properly classify remotely sensed urban images will continue to perplex analysts until researchers understand how automated classification and spatial resolution fully interact on high variability images. One such option posited in this paper is the use of multiple resolution images together in a hybrid classification.

This classification should also takes advantage of local variance information, since the spatial context (density and arrangement of components) becomes very important in classifying urban images.

Conclusion

One meter resolution images from Akron, OH and Minneapolis, MN acquired by aircraft were degraded to multiple resolutions using a method developed by Sadowski and Sarno (1976). The original images were classified using a hybrid supervised classification technique for Akron and an standard unsupervised classification for Minneapolis. Additionally, randomly selected hand-drawn masks of aggregate land use classes were made. Statistics on the first and second principal components of both images were computed at multiple resolutions. These statistics included spectral plots, spectral ellipses, transformed and Jeffreys-Matusita divergence, local variance, components plots, pureness statistics and aggregate statistics.

The results showed that high variability urban classes (i.e., residential, commercial) had different responses to these statistics than homogenous classes, such as agriculture, forest and water. The urban classes had components aggregating all around 8 meters, and masked classes disaggregating past 100 meters (or 60 meters in the case of Akron). This suggests that a crossover resolution occurs somewhere in between these two values. The results did not necessarily show that taxonomic hierarchies breakdown at different resolutions, except that urban classes breakdown differently than homogenous classes. The results did show how components are parts of many, widely varying aggregate classes and that density and pattern of arrangement is as important as simple occurrence of a component class within any aggregate

class. This was more the case in the Minneapolis image, where multiple aggregate classes all had similar amounts of the various component classes.

The primary goal of this paper was to develop the statistical methodology to test classification and resolution interaction. Future work with this experiment should look at the details of aggregate and component classes to be able to start to develop actual methods of improving urban classification. Classifications at intermediate levels would be useful to predict crossover points. A larger data set may be of use to account for the smaller number of data points available at coarser resolutions within the same image. The results from these experiments should be used to create predictive taxonomic class systems for urban remotely sensed images. Finally, the idea of multiple resolution hybrid classification should be considered and tested. Somewhere within the complications of the reality of urban remote sensing classification, a method that incorporates both spectral and spatial information will prevail in providing better land cover and land use assessment of urban images.

Acknowledgements

I would like to thank my advisor, Dr. Howard Veregin in the Department of Geography, University of Minnesota, for providing just about everything needed to do this project, from the data, to the theories to the brainstorming of methods. This project was also aided by the support of members in the Spring 2000 *Seminar is GIS and Cartography*. Finally, I would like to thank Dr. Marv Bauer in the Department of Forest Resources, University of Minnesota and the entire staff at the Earth Resources Analysis Center (ERSAC) for providing technical support.

Works Cited

Anderson, James R., Ernest E. Hardy, John T. Roach and Richard E. Witmer. 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data.* United State Geological Survey, Professional Paper 964.

Cao, Changyong and Nina Sio-Ngam Lam. 1997. Understanding the Scale and Resolution Effects in Remote Sensing and GIS, in Dale A. Quattrochi and Michael F. Goodchild, eds. *Scale in Remote Sensing and GIS*. Boca Raton, FL: Lewis Publishers.

Cushnie, Janis L. 1987. The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies. *International Journal of Remote Sensing*, v. 8, n. 1: 15-29.

Foody, Giles M. 1994. Ordinal-Level Classification of Sub-Pixel Tropical Forest Cover. *Photogrammetric Engineering and Remote Sensing,* v. 60, n. 1: 61-65.

Goodchild, Michael F. and Dale A. Quattrochi. 1997. Scale, Multiscaling, Remote Sensing, and GIS, in D. Quattrochi and M. Goodchild, eds. *Scale in Remote Sensing and GIS*. Boca Raton, FL: Lewis Publishers.

Hardin, Perry J. 1994. Parametric and Nearest-Neighbor Methods for Hybrid Classification: A Comparison of Pixel Assignment Accuracy. *Photogrammetric Engineering and Remote Sensing,* v. 60, n. 12: 1439-1448.

Jensen, John R. 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Upper Saddle River, NJ: Simon & Schuster, Inc.

Justice, C. O., B. L. Markham, J. R. G. Townshend and R. L. Kennard. 1989. Spatial degradation of satellite data. *International Journal of Remote Sensing*, v. 10, n. 9: 1539-1561.

Kolman, Bernard. 1991. *Elementary Linear Algebra*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

McMaster, Robert B. and K. Stuart Shea. 1992. *Generalization in Digital Cartography*. Washington, D.C.: Association of American Geographers.

Monmonier, Mark S. 1983. Raster-Mode Area Generalization for Land Use and Land Cover Maps. *Cartographica*, v.20, n. 4: 65-91.

Sadowski, F. A. and J. E. Sarno. 1976. *Forest classification accuracy as influenced by multispectral scanner spatial resolution*. Report No. 109600-71-F. Ann Arbor, MI: Environmental Research Institute of Michigan.

Toll, David L. 1985. Effect of Landsat Thematic Mapper Sensor Parameters on Land Cover Classification. *Remote Sensing of the Environment,* v. 17: 129-140.

Veregin, Howard. 1997. *Resolution-Dependent Effects in Digital Image Classification*. unpublished class handout for Geography 5900, Earth Imagery I, University of Minnesota Department of Geography

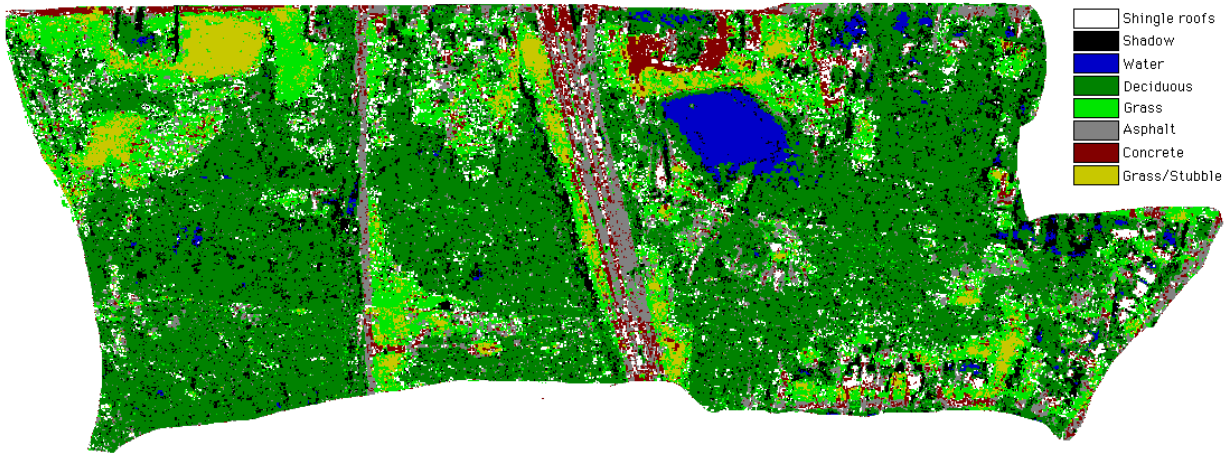
Veregin, Howard, Peter Sincák, Kirk Gregory, and Lawrence Davis. 1995. Integration of High Resolution Video Imagery and Urban Stormwater runoff Modeling. *15th Biennial Workshop on Videography and Color Photography in Resource Assessment:* 182-191.

Veregin, Howard, Peter Sincák and Norbert Kopco. in press. Enhancement of Classification Accuracy using Conflation Procedures. *GeoCarto International*.

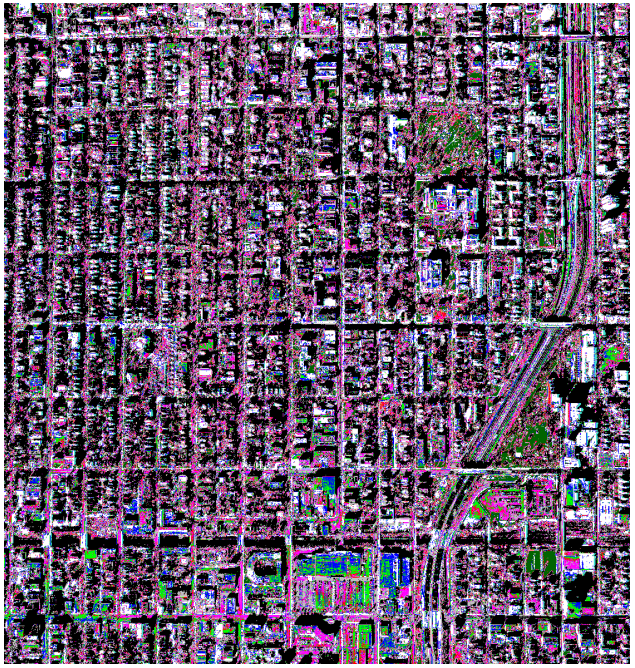
Wang, Minhua, Peng Gong and Philip J. Howarth. 1991. Thematic Mapping From Imagery: An Aspect of Automated Map Generalization. Proceedings: Tenth International Symposium on Computer-Assisted Cartography, Auto-Carto 10: 123-132.

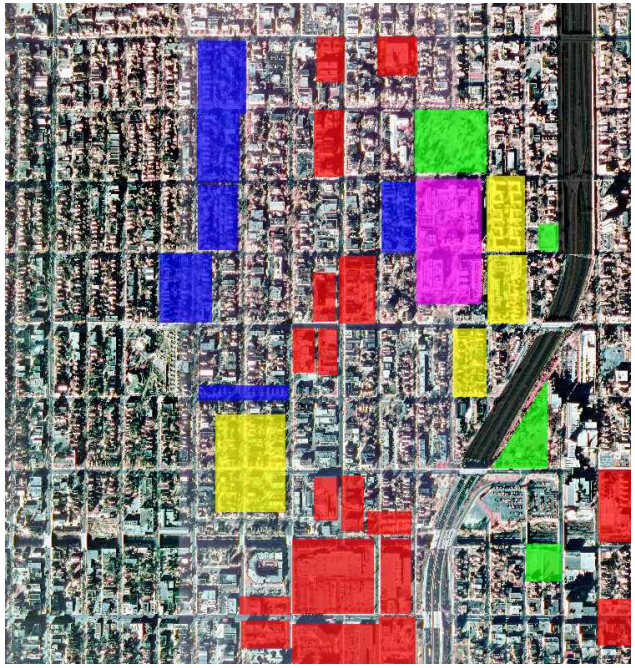
Woodcock, Curtis E. and Alan H. Strahler. 1987. The Factor of Scale in Remote Sensing. *Remote Sensing of the Environment,* v. 21: 311-332.

List of Figures and Tables


Shingle roofs
Shadow
Water
Deciduous
Grass
Asphalt
Concrete
Grass/Stubble

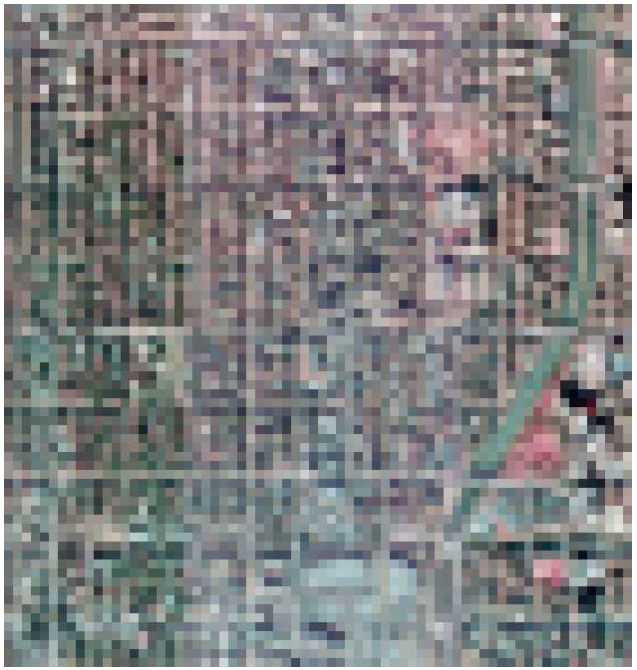

Legend
Road
Field
Commercial complex
Water
Forest
Residential

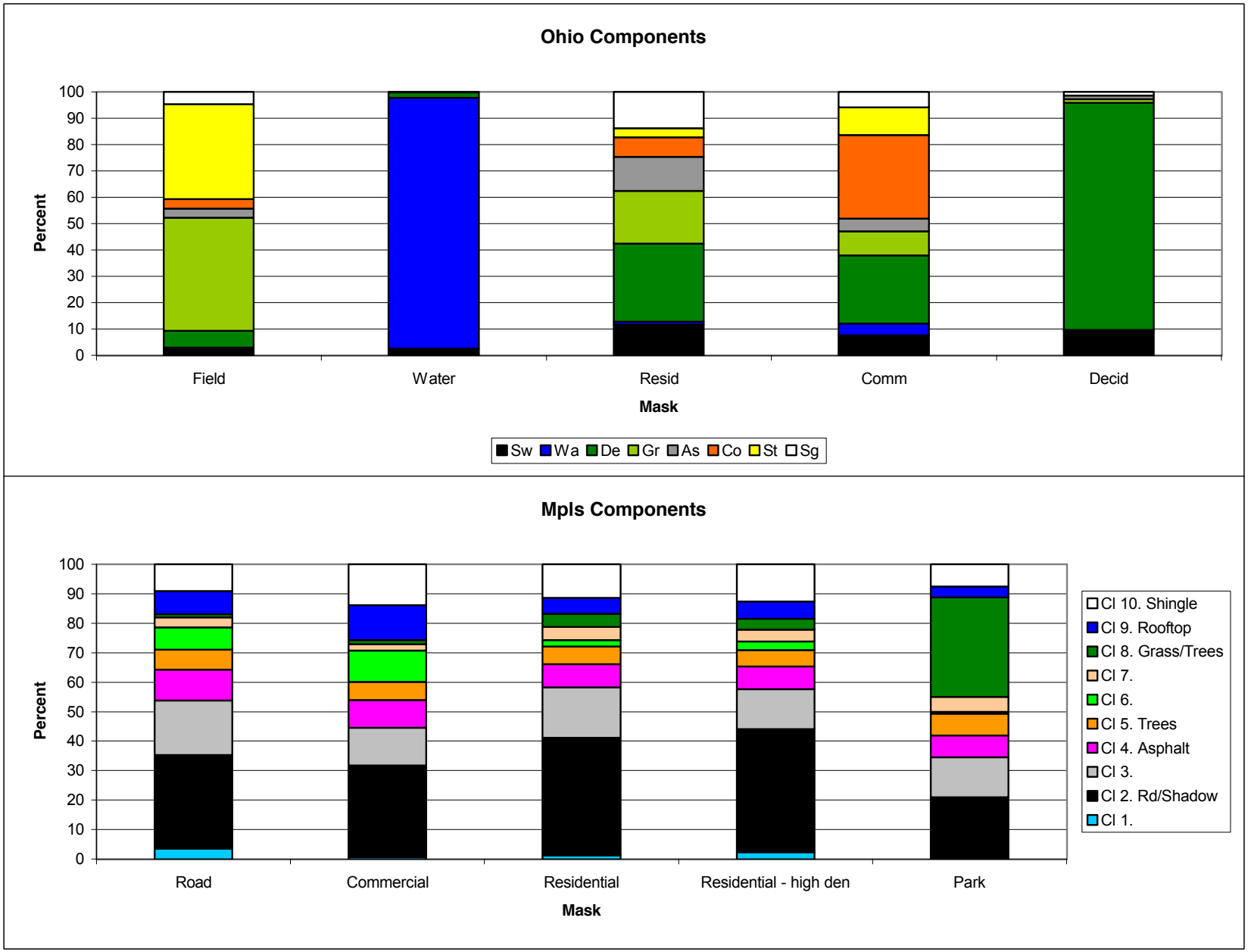


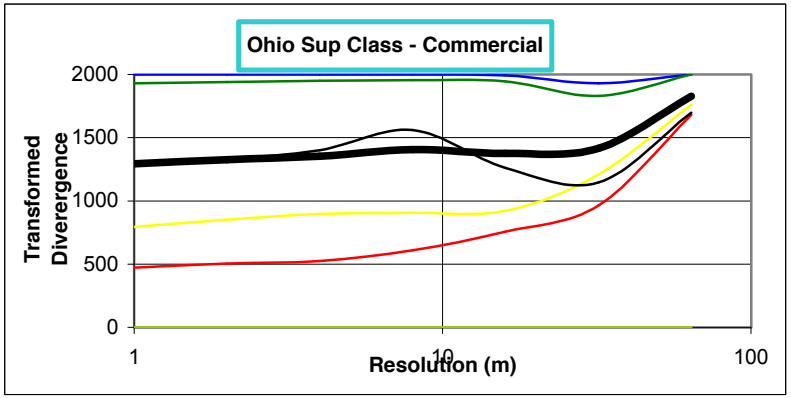


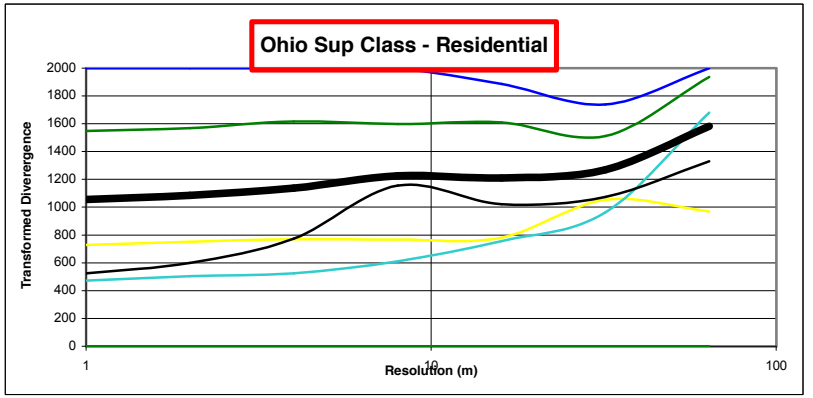
Legend

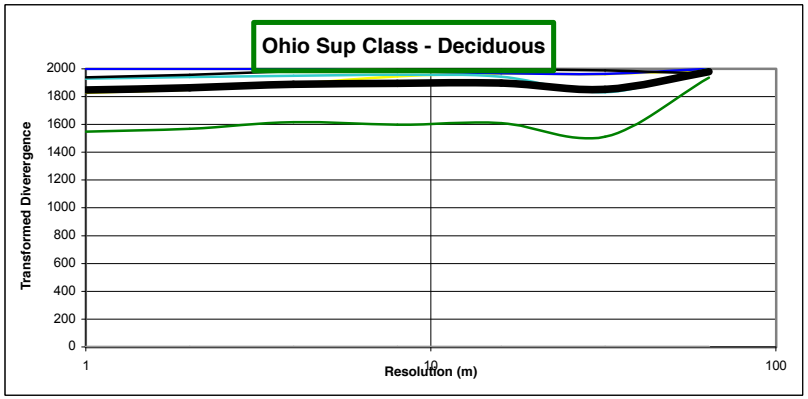
- Road
- Parkland
- Residential
- Commercial
- Residential (high density)
- Cultural

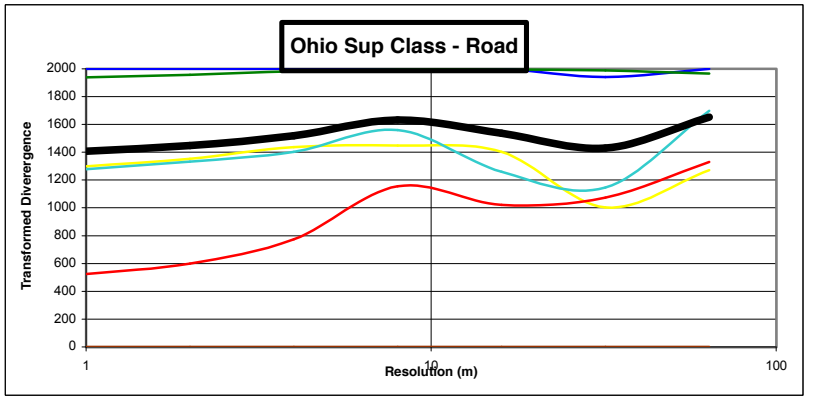


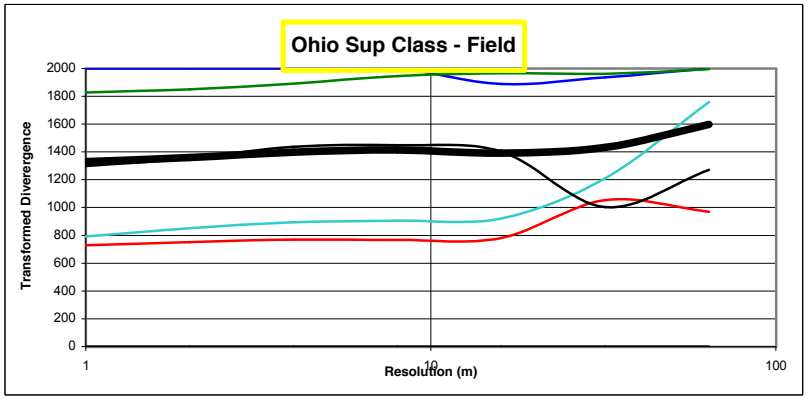

Ohio Components
Percent
100
90
80
70
60
50
40
30
20
10
0
Field
Water
Resid
Comm
Decid
Mask
Sw
Wa
De
Gr
As
Co
St
Sg
Mpls Components
Percent
100
90
80
70
60
50
40
30
20
10
0
Road
Commercial
Residential
Residential - high den
Park
Mask
Cl 10. Shingle
Cl 9. Rooftop
Cl 8. Grass/Trees
Cl 7.
Cl 6.
Cl 5. Trees
Cl 4. Asphalt
Cl 3.
Cl 2. Rd/Shadow
Cl 1.

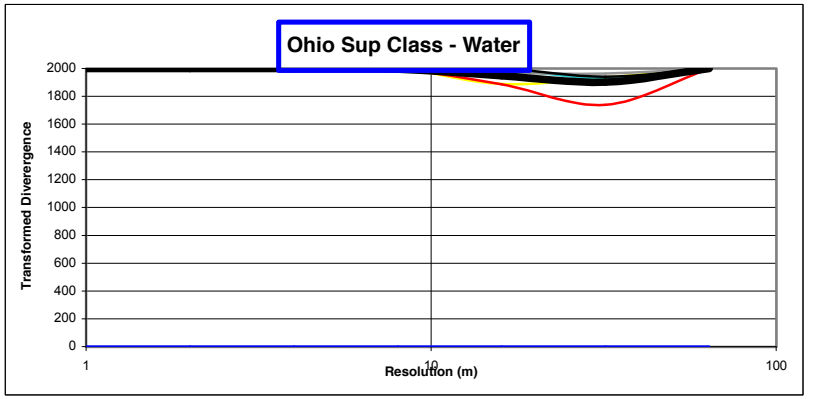

Ohio Sup Class - Commercial
Transformed Diverergence
2000
1500
1000
500
0
1
10
100
Resolution (m)

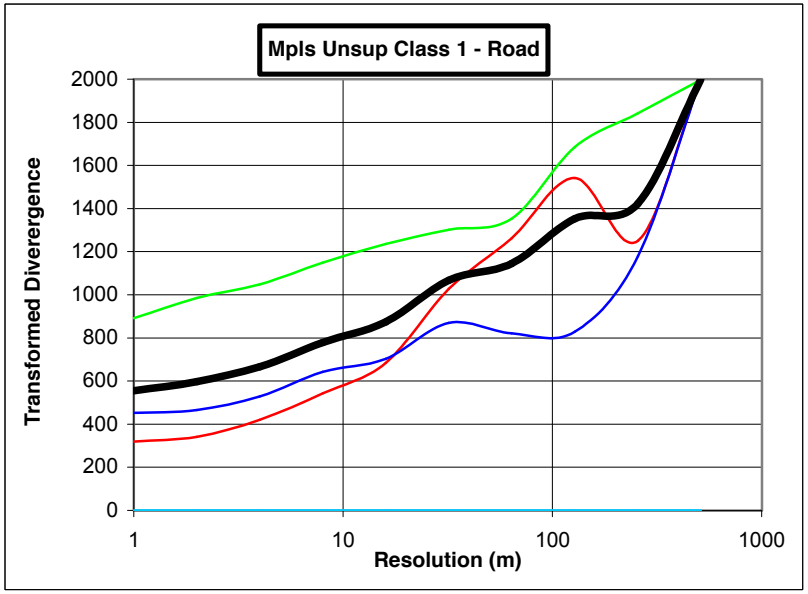

Ohio Sup Class - Residential
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
Resolution (m)

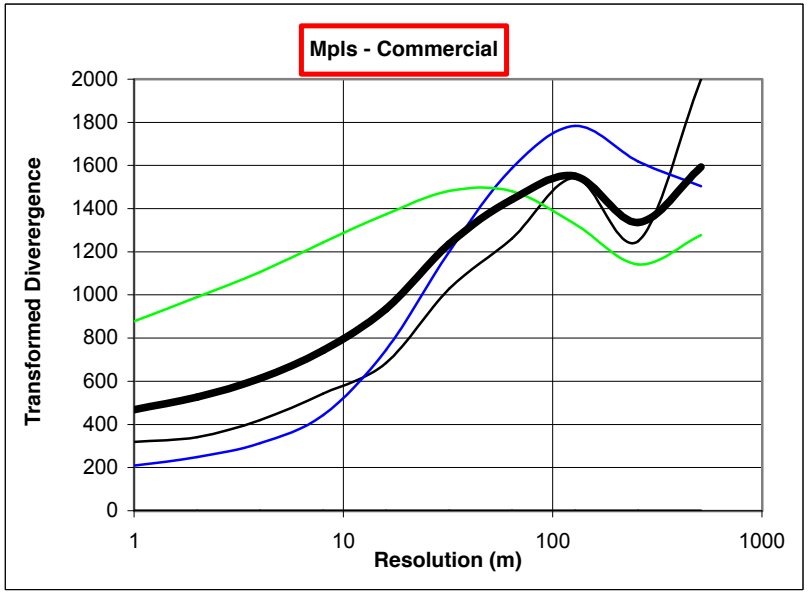

Ohio Sup Class - Deciduous
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
Resolution (m)

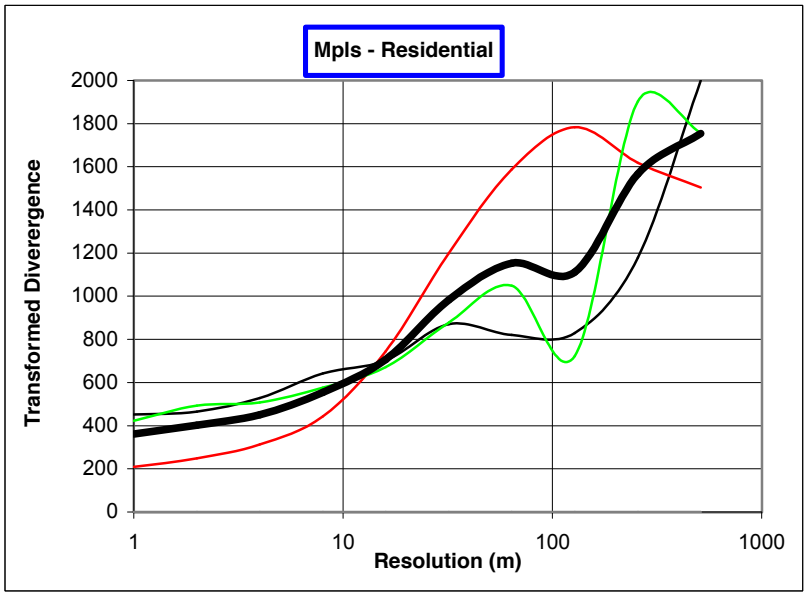

Ohio Sup Class - Road
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
Resolution (m)

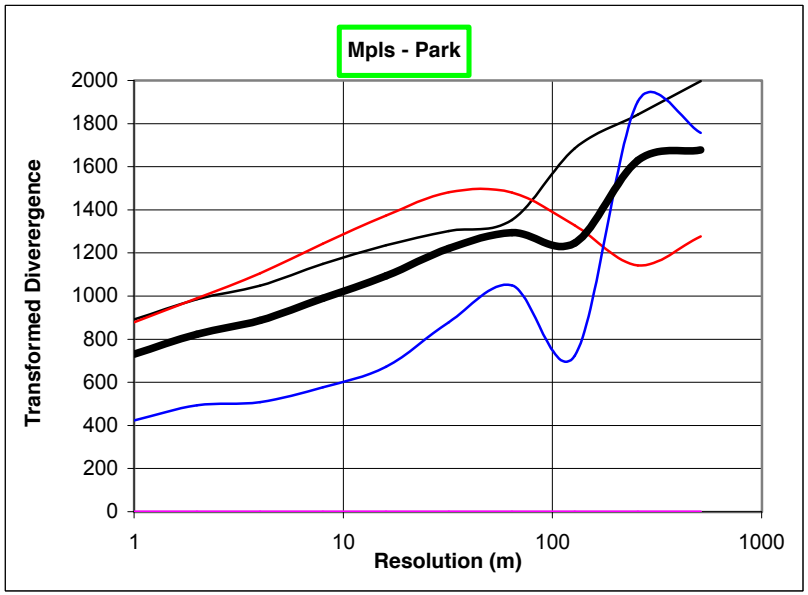

Ohio Sup Class - Field
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
Resolution (m)

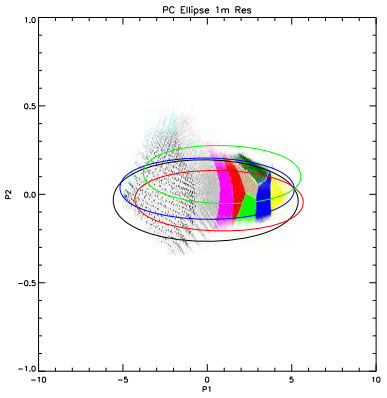

Ohio Sup Class - Water
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
Resolution (m)

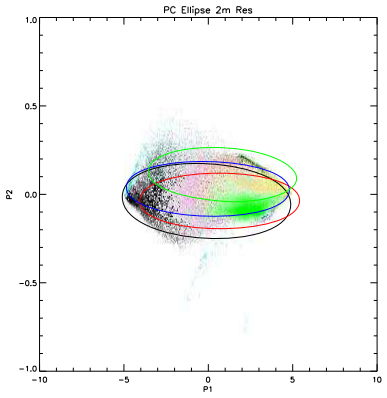

Mpls Unsup Class 1 - Road
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
1000
Resolution (m)

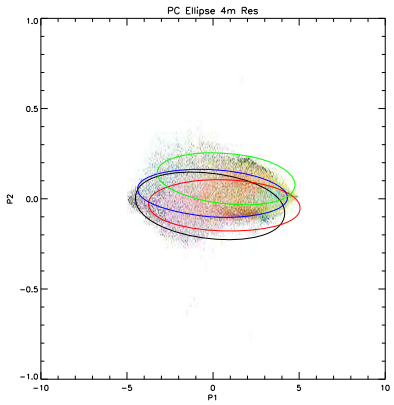

Mpls - Commercial
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
1000
Resolution (m)

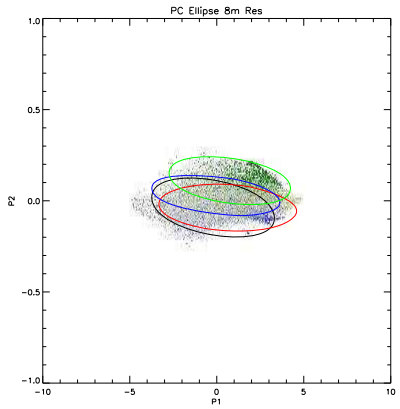

Mpls - Residential
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
1000
Resolution (m)

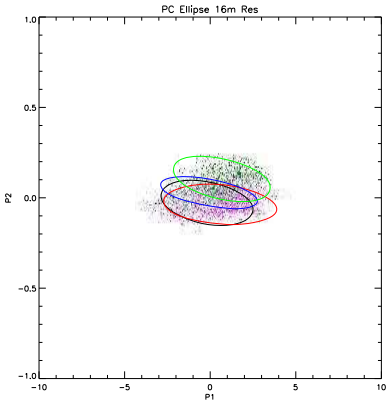

Mpls - Park
Transformed Diverergence
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1
10
100
1000
Resolution (m)

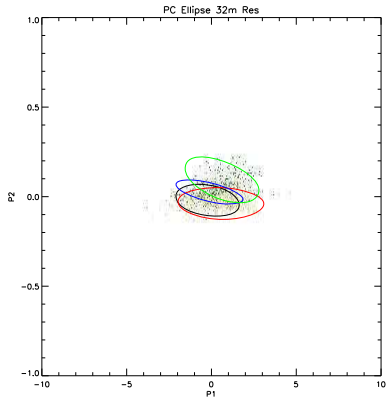

PC Ellipse 1m Res
P1
P2

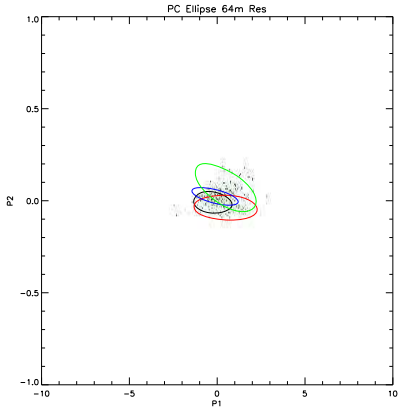

PC Ellipse 2m Res
P1
P2

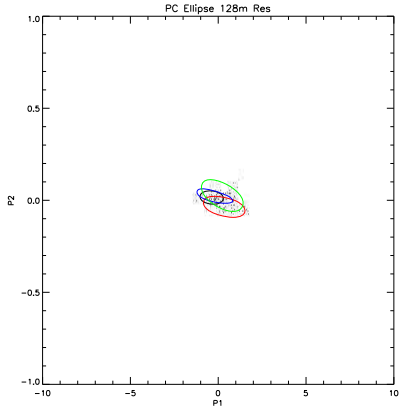

PC Ellipse 4m Res
P1
P2

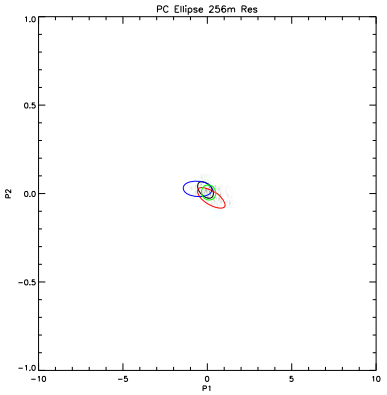

PC Ellipse 8m Res
P1
P2

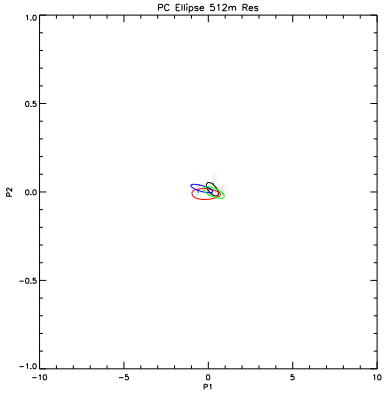

PC Ellipse 16m Res
P1
P2


PC Ellipse 32m Res
P1
P2

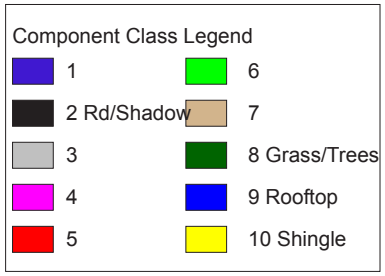

PC Ellipse 64m Res
P1
P2

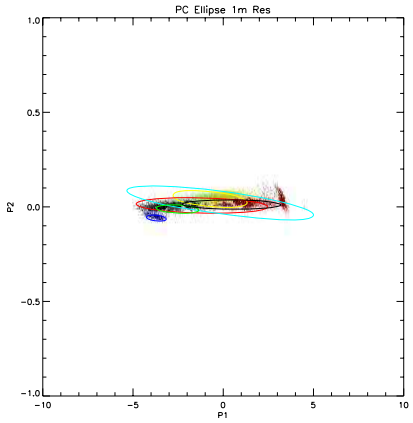

PC Ellipse 128m Res
P1
P2

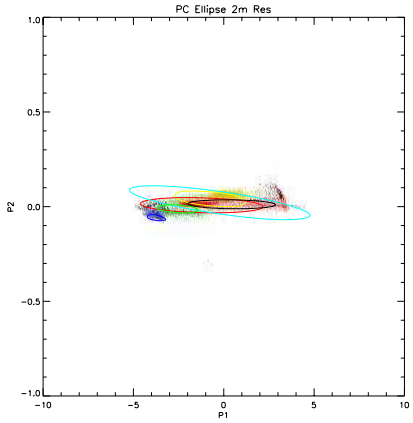

PC Ellipse 256m Res
P1
P2

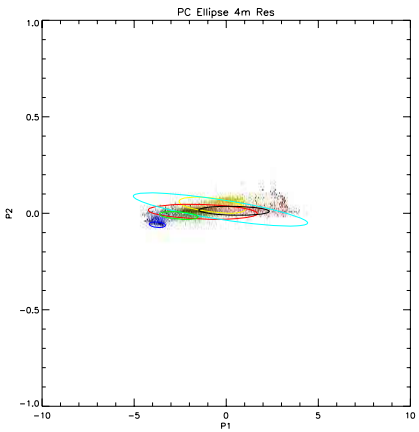

PC Ellipse 512m Res
P1
P2

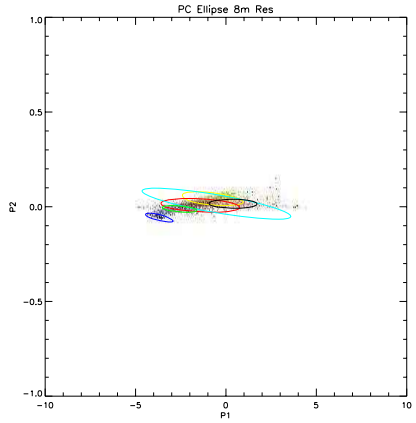

Ellipse / Mask Legend
Residential
Commercial
Park
Road

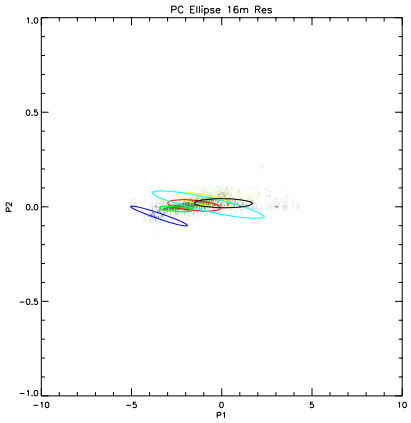

Component Class Legend
1
2 Rd/Shadow
3
4
5
6
7
8 Grass/Trees
9 Rooftop
10 Shingle

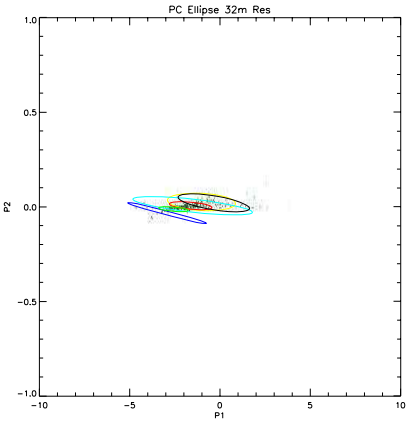

PC Ellipse 1m Res
P1
P2

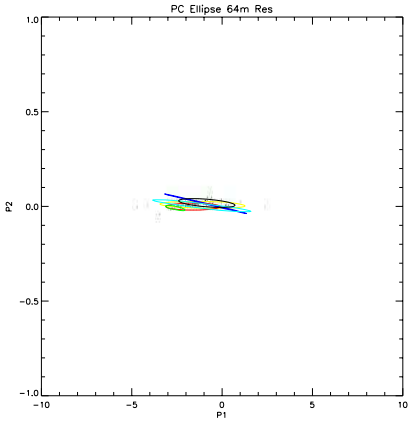

PC Ellipse 2m Res
P1
P2

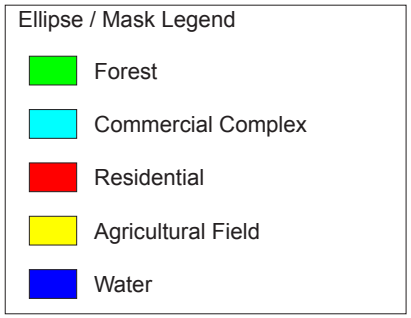

PC Ellipse 4m Res
P1
P2

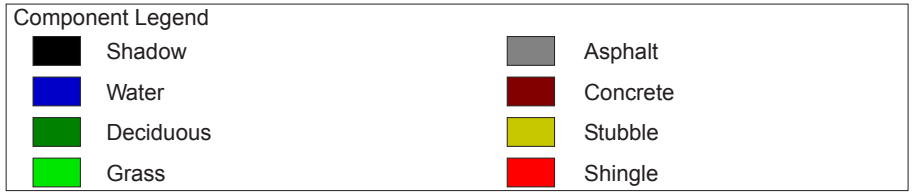

PC Ellipse 8m Res
P1
P2

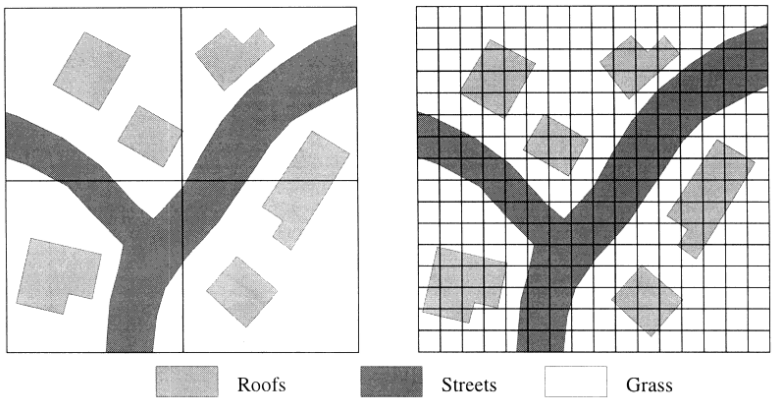

PC Ellipse 16m Res
P1
P2

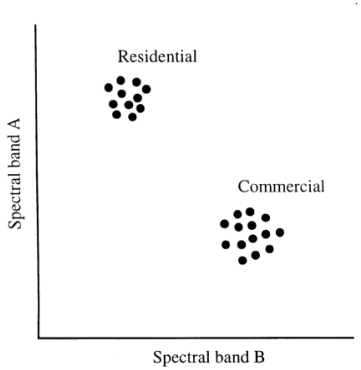

PC Ellipse 32m Res
P1
P2


PC Ellipse 64m Res
P1
P2

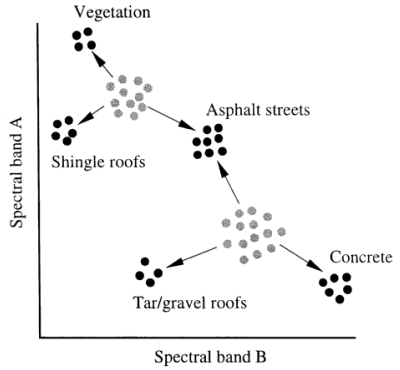

Ellipse / Mask Legend
Forest
Commercial Complex
Residential
Agricultural Field
Water


Component Legend
Shadow
Water
Deciduous
Grass
Asphalt
Concrete
Stubble
Shingle

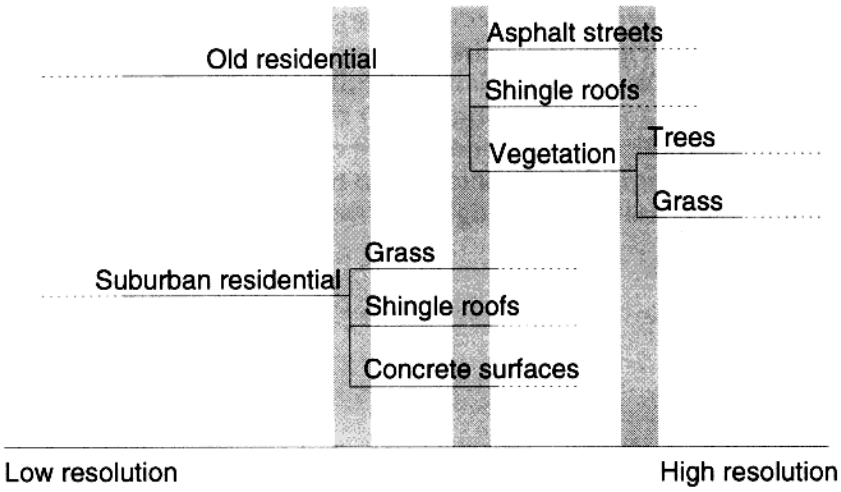


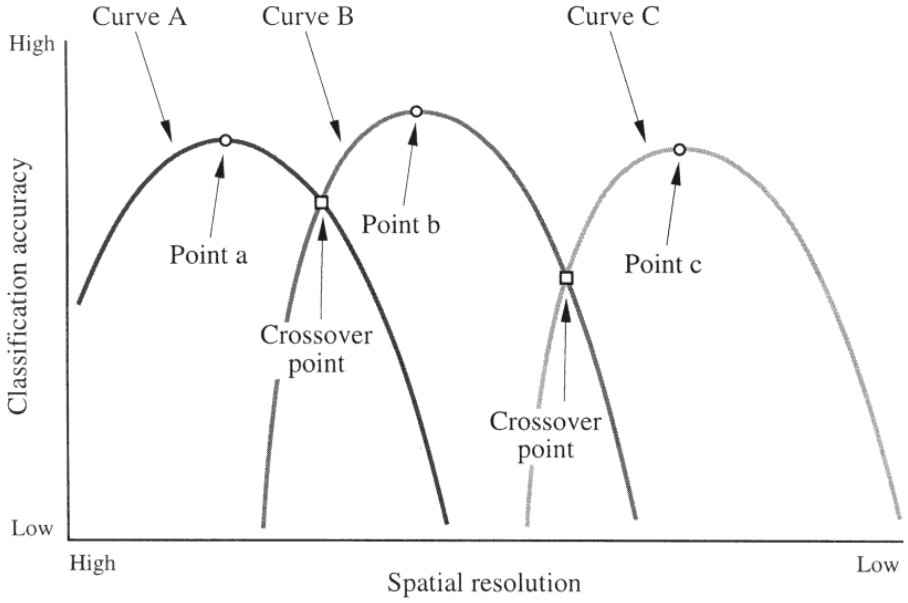


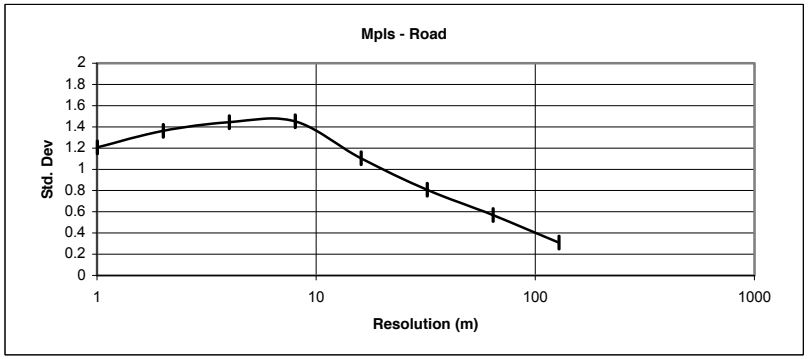
Low Resolution

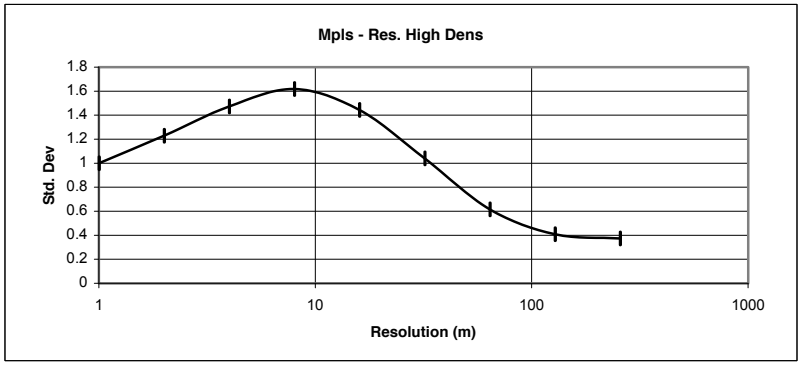


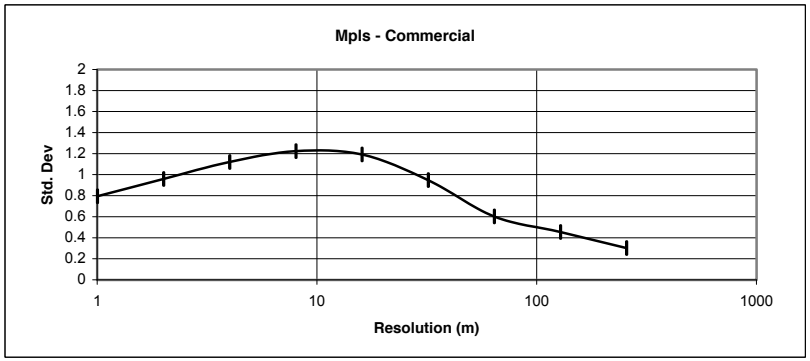
High Resolution



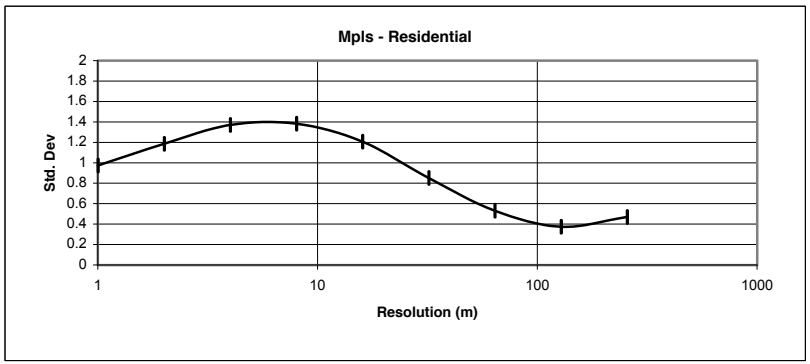


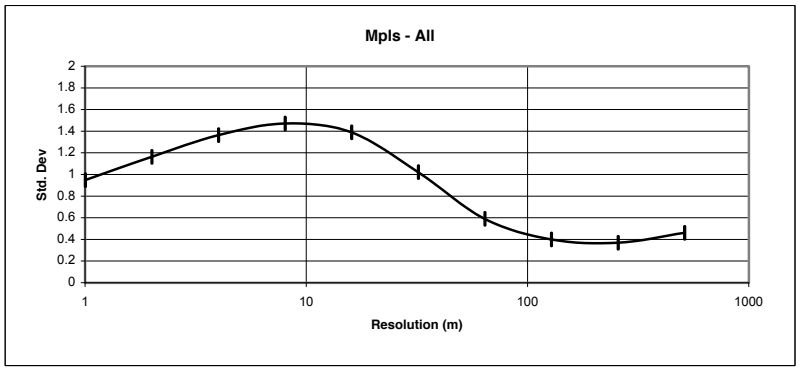

Mpls - Road
Std. Dev
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
1
10
100
1000
Resolution (m)

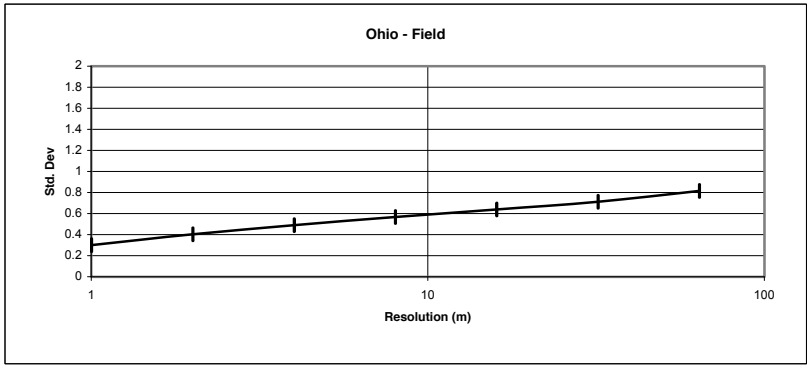

Mpls - Res. High Dens
Std. Dev
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
1
10
100
1000
Resolution (m)

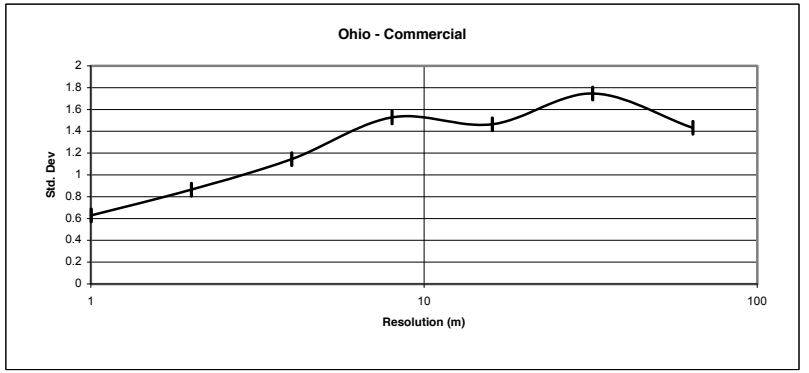

Mpls - Commercial
Std. Dev
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
1
10
100
1000
Resolution (m)

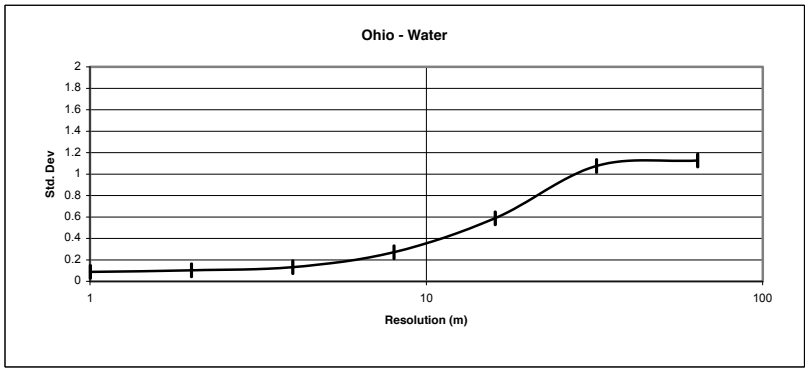

Mpls - Park
Std. Dev
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
1
10
100
1000
Resolution (m)

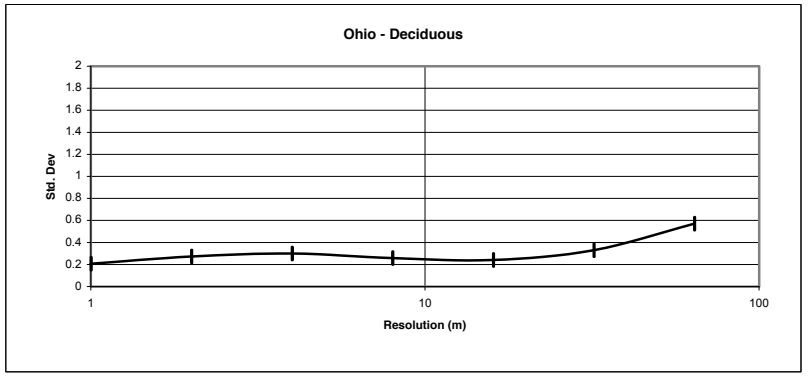

Mpls - Residential
Std. Dev
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
1
10
100
1000
Resolution (m)

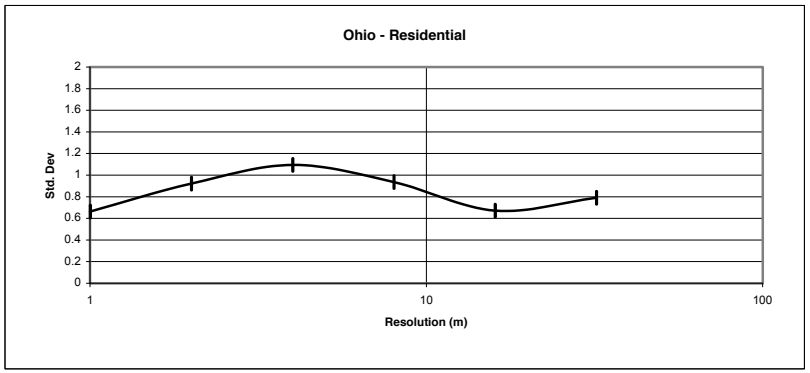

Mpls - All
Std. Dev
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
1
10
100
1000
Resolution (m)

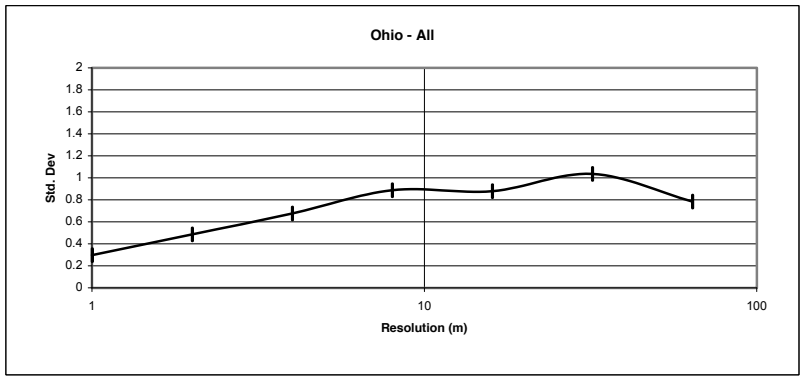

Ohio - Field
Std. Dev
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
10
100
Resolution (m)

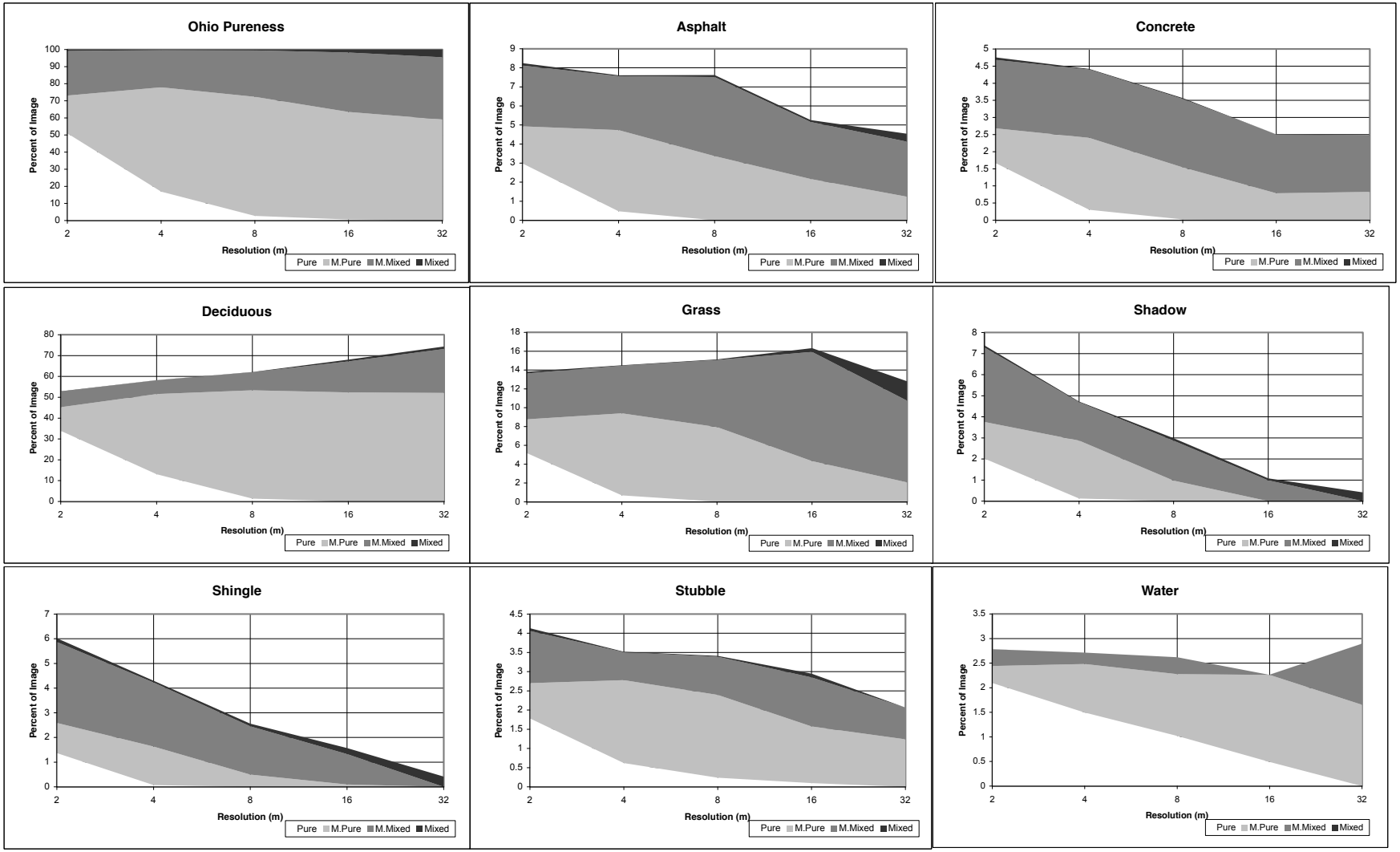

Ohio - Commercial
Std. Dev
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
10
100
Resolution (m)


Ohio - Water
Std. Dev
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
10
100
Resolution (m)


Ohio - Deciduous
Std. Dev
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
10
100
Resolution (m)


Ohio - Residential
Std. Dev
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
10
100
Resolution (m)


Ohio - All
Std. Dev
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
10
100
Resolution (m)


Ohio Pureness
Asphalt
Concrete
Deciduous
Grass
Shadow
Shingle
Stubble
Water
Percent of Image
Resolution (m)
2
4
8
16
32
Pure
M.Pure
M.Mixed
Mixed