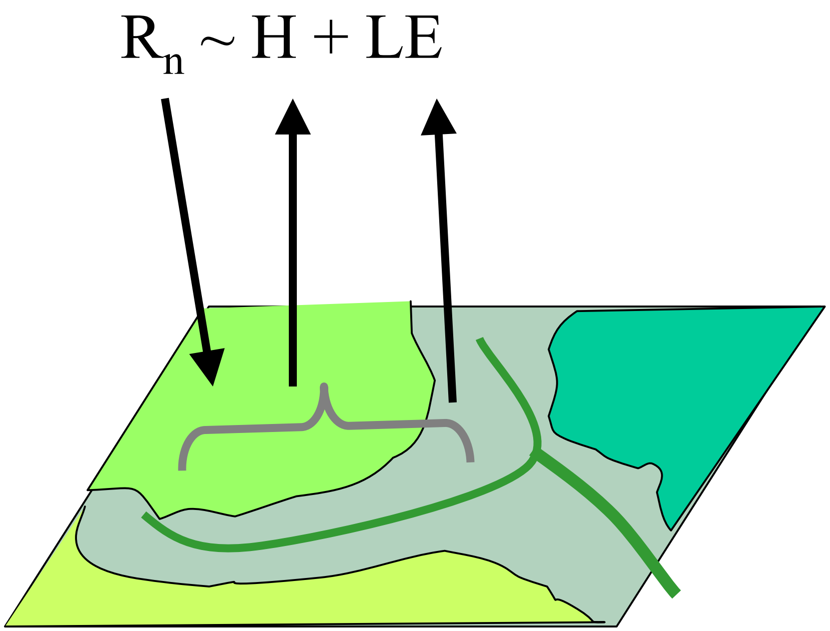
Surface water/energy budget coupling over heterogeneous terrain



$LE = f_{veg} LE_{veg} + (1 - f_{veg}) LE_{soil}$

$LE = f(R_n, T, g_c, g_a, g_{soil}, VPD)$

$g_a = f(\text{canopy structure, wind, ...})$

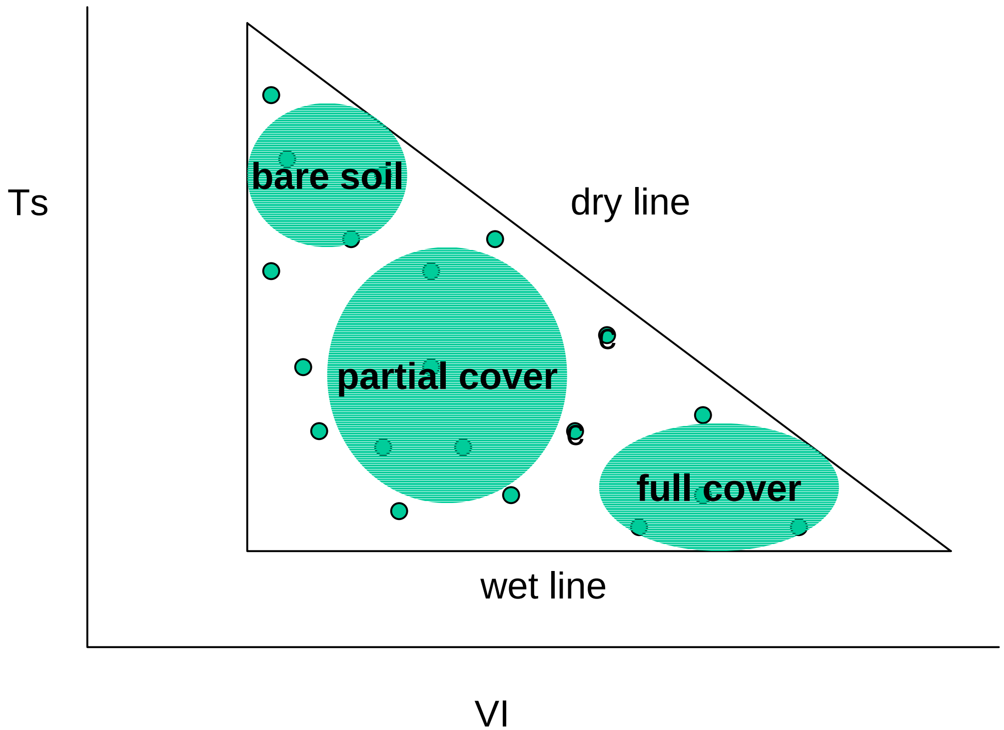
$g_c = f(\text{soil water, VPD, PAR, T, LAI})$

$g_{soil} = f(\text{soil water, ...})$

T_s lower with greater LE (evaporative cooling) as a function of soil water (other factors), greater canopy cover (higher NDVI)

T_s and NDVI estimated by a set of operational remote sensors

Interpretation of the VI-T_s Space

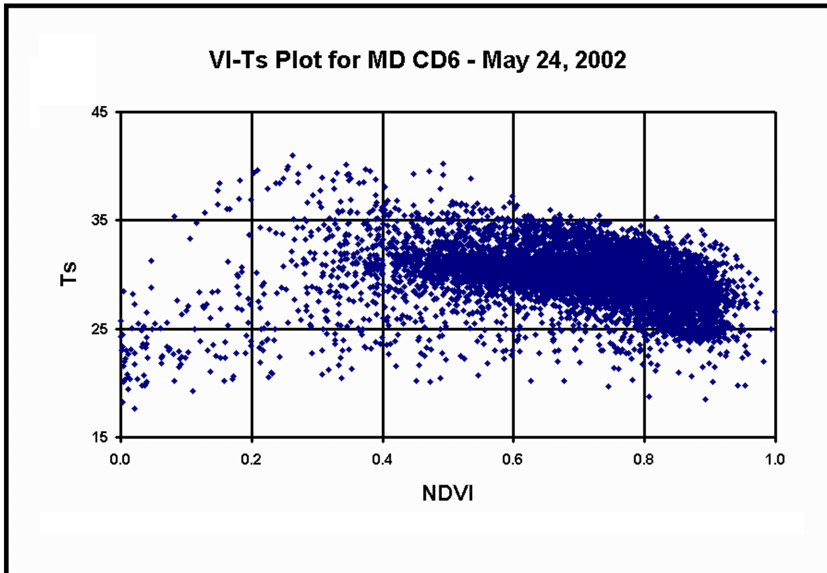


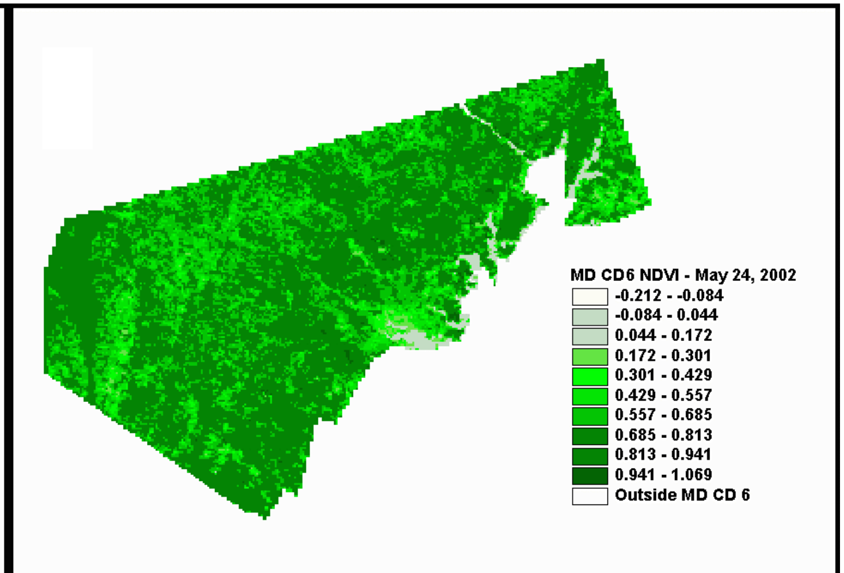
Adapted from Sandholt et al. 2002

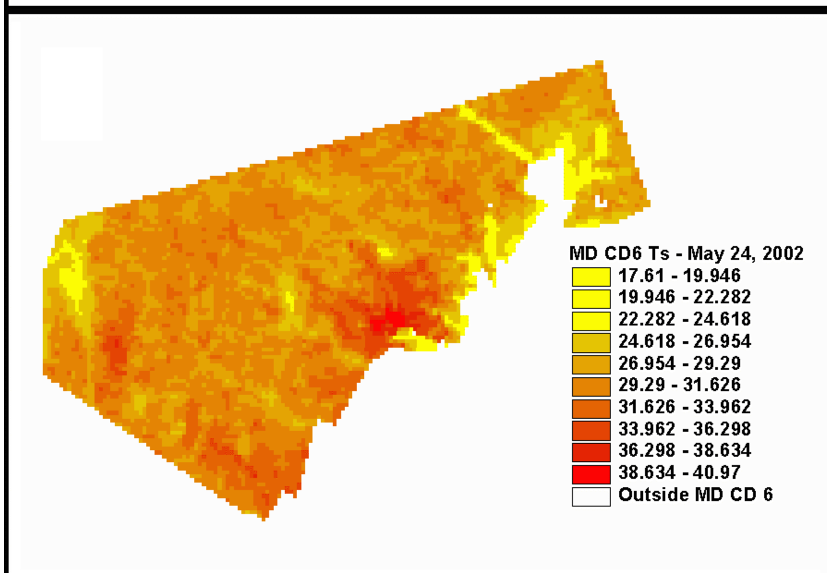
Dry Line Slope – Sigma (σ)

- Nemani and Running (1989) suggested, and later Nemani, Pierce, Running, and Goward (1993) demonstrated, that **the slope of the dry line** (symbolized using σ) is a good overall indicator of the **surface moisture condition** of a region (where the T_s and VI pixels that are drawn from to form the 2-D T_s-VI distribution) on the occasion when the imagery was collected
  - **Steeper, more negative** slopes represent **drier conditions** (where T_s disparities are greater)
- So **how** do we form the 2-D T_s-VI distribution and find the slope of the dry line?

Finding the Dry Line (σ) Slope





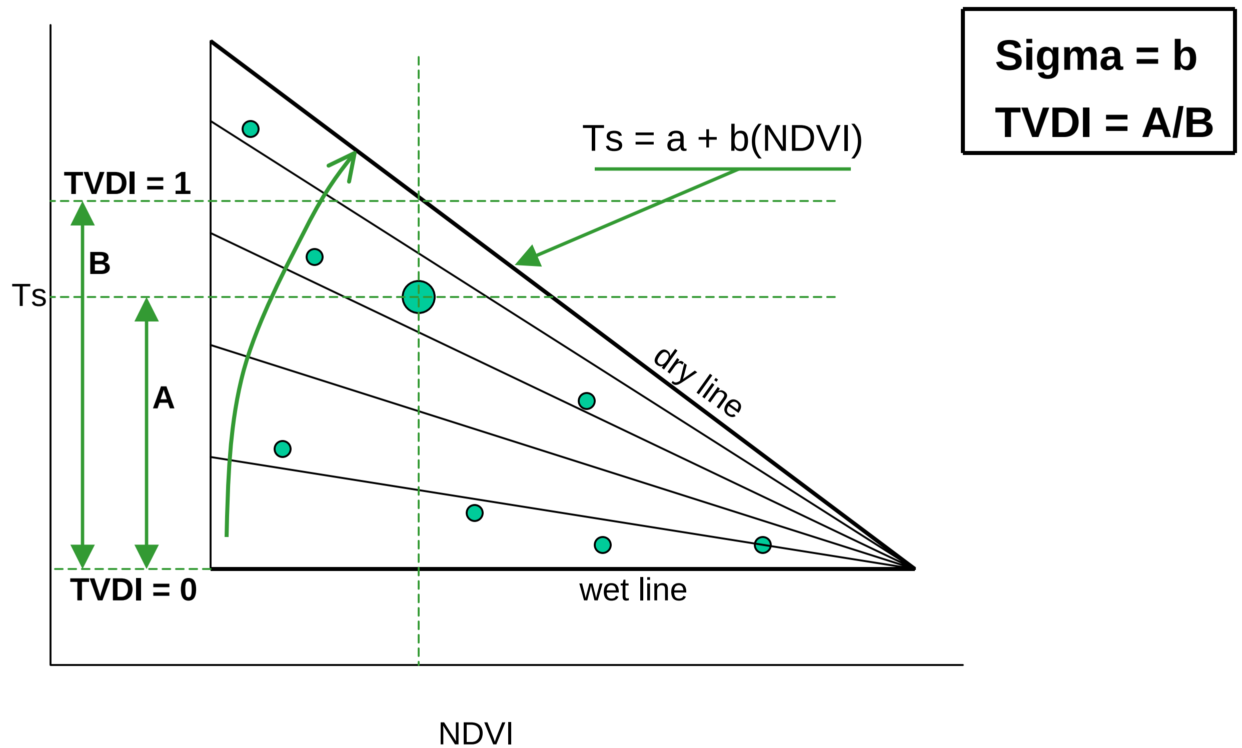


- We begin with **T_s and VI data**, ideally collected using the **same sensor at the same time** (e.g. from AVHRR bands 1, 2, 4, & 5)
- We then translate the values for each pixel into a **2-D parameter space**, the VI on the x-axis and the T_s on the y-axis

Obtaining Per Pixel Dryness Info

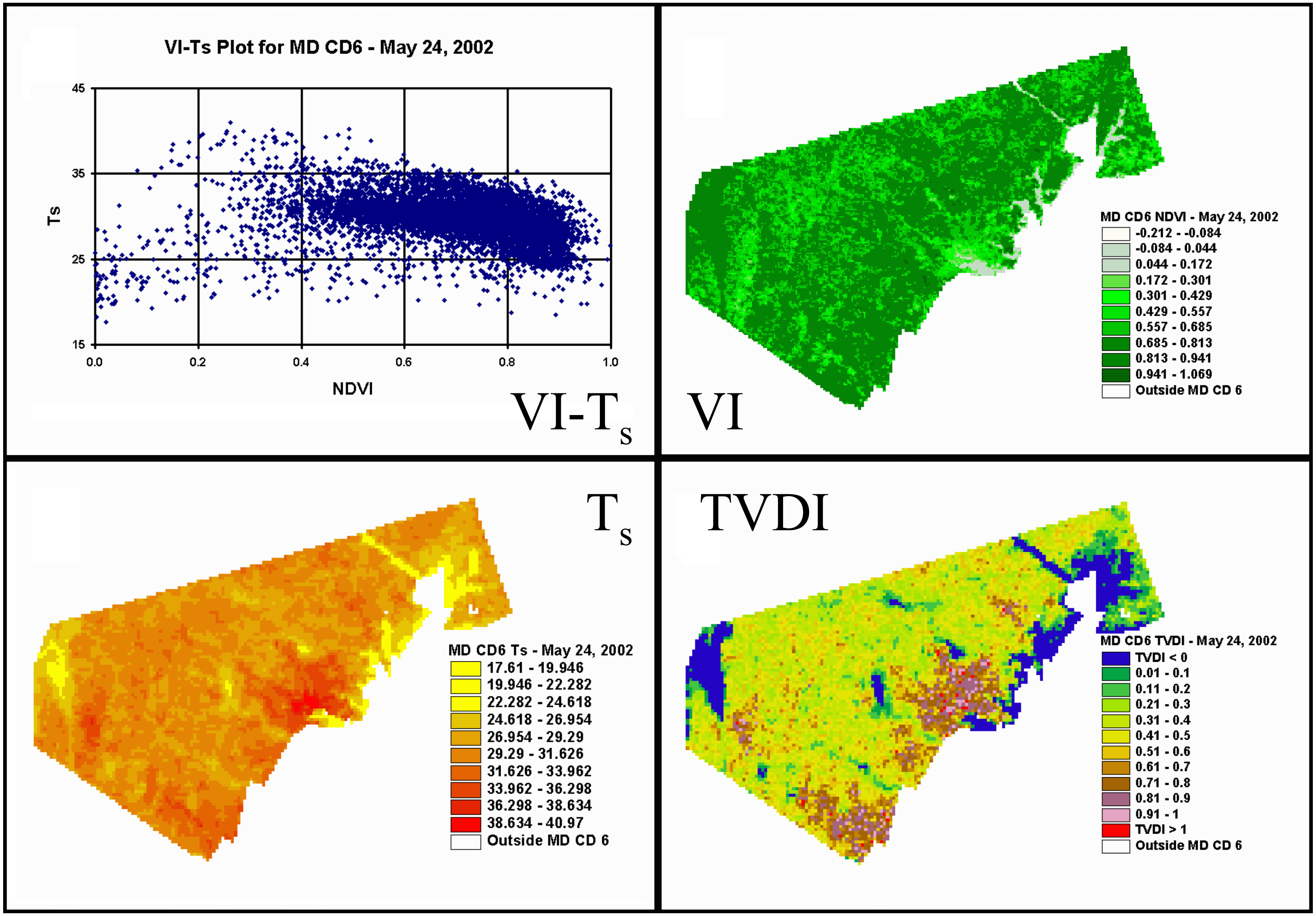
- **The slope of the dry line** (symbolized using σ) is a good overall indicator of the **surface moisture condition of a <u>region</u>** (where the T_s and VI pixels that are drawn from to form the 2-D T_s-VI distribution)
  - But it is just that, **a single number** that is a **regional descriptor** of the **surface moisture condition** of the **overall aggregate set of pixels**
- What if we want to know something about the **surface moisture condition of individual pixels**? How can we do this?
  - One way is to take an approach that **describes each pixel's position** in the distribution

Temperature Vegetation Dryness Index



Adapted from Sandholt et al. 2002

Generating TVDI Values



Temperature Vegetation Dryness Index

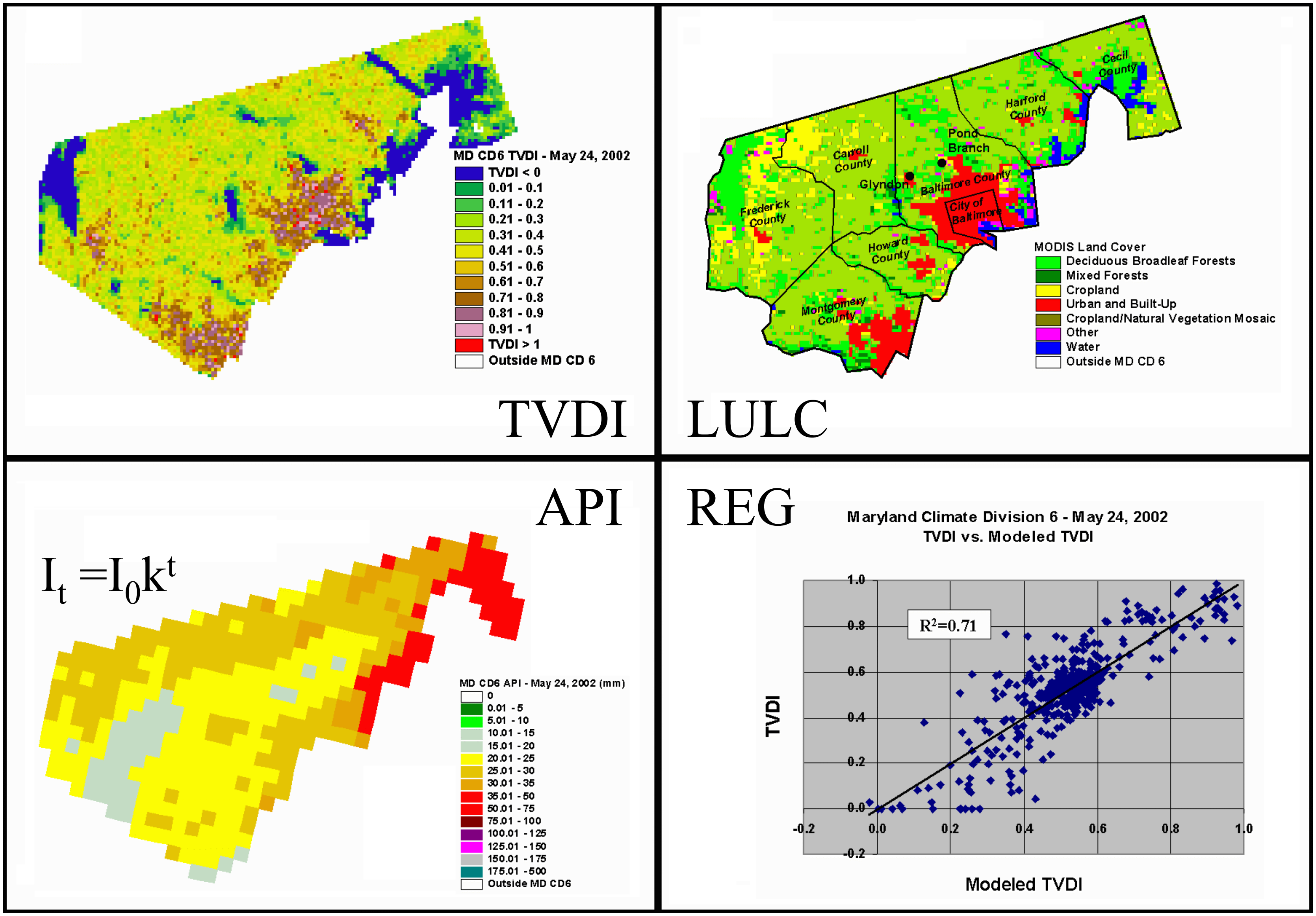
- The **procedure for creating TVDI** initially requires all the steps required to obtain σ :

  1. Form the 2-D T_s – VI distribution
  2. Calculate/find σ

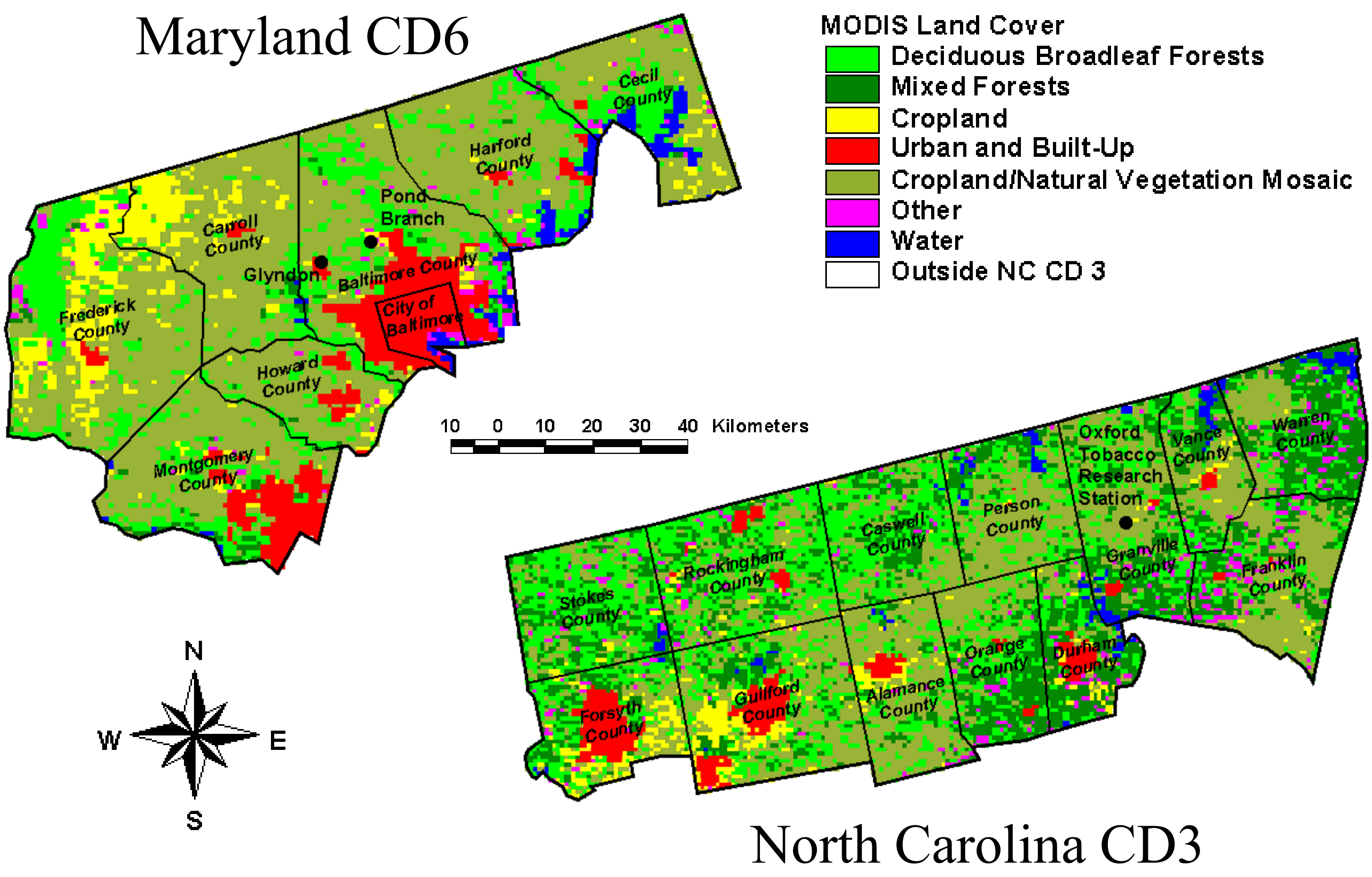
  followed by **a few further steps**:

  3. Define the wet line along the bottom the triangle (which can usually be safely done in a fairly unsophisticated fashion)
  4. Calculate TVDI as described (where is the point/pixel of interest positioned between the dry and wet lines at the given NDVI)
  5. Take the resulting values and map them back to their respective pixels

Modeling TVDI



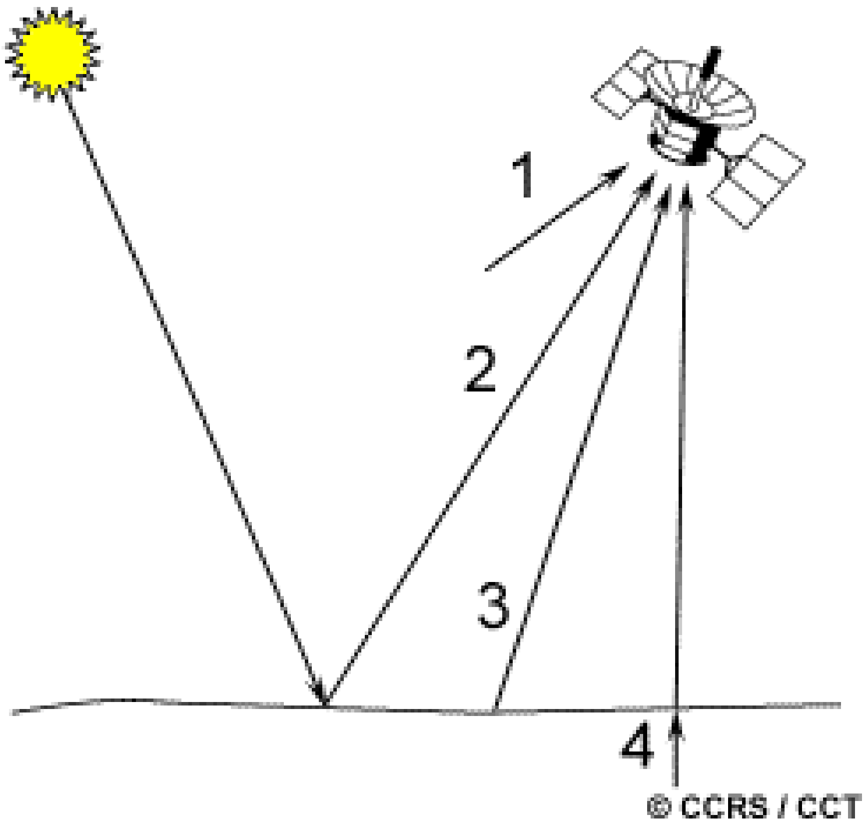
MODIS LULC In Climate Divisions



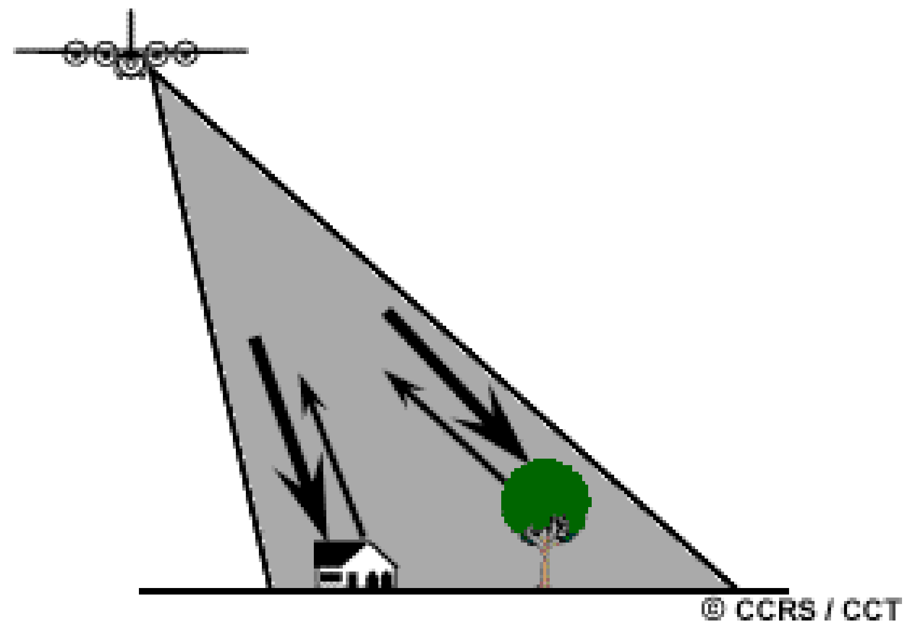
Two Types of Remote Sensing

- Based on the source of the energy, remote sensing can be broken into two categories:
- **Passive remote sensing**: The source of energy collected by sensors is either **reflected solar radiation** (e.g. cameras) or **emitted by the targets** (thermal imaging).
- **Active remote sensing**: The source of energy collected by sensors is actively **generated by a man-made device**. Examples include RADAR (RAdio Detection And Ranging, which uses microwave energy) and LIDAR (LIght Detection Imagery And Ranging, which uses a laser).

Passive vs. Active Remote Sensing



http://www.ccrs.nrcan.gc.ca/ccrs/learn/tutorials/fundam/chapter3/chapter3_1_e.html



Passive sensors receive **solar energy reflected** by the Earth's surface (2), along with energy emitted by the atmosphere (1), surface (3) and sub-surface (4)

Active sensors receive energy reflected from the Earth's surface that originally came from an **emitter other than the Sun**

RADAR Remote Sensing

- Remote sensing using RADAR can be active or passive:
  - Some earth materials do emit radiation in the **microwave range of wavelengths** (anywhere from a millimeter to a meter), and these can be sensed by a detector that operates just as many that we have already looked at does, sensing the energy passively
  - However today we're primarily going to look at active RADAR remote sensing, where the **source of the microwave energy** which returns to the sensor is a man-made source or emitter, and the characteristics of the emitter and sensor are both selected for the particular application (i.e. choose the wavelength and other factors based on what you want to capture in the imagery)

RADAR Remote Sensing

- The **platform/position** of the emitter and sensor can vary:
  - **Aircraft and ships** are routinely fitted with active RADAR systems for purposes of **navigation**, although we find research and geographic information oriented systems on these platforms as well
  - There are **satellite systems** that use active microwave sensing systems (e.g. Radarsat, Japan's Earth Resources Satellite JERS-1, and the SIR-C/X-SAR system that was flown on the space shuttle 1994 and again in 2000 - SRTM)
  - There are **land-based systems** like the Doppler RADAR network used to produce precipitation estimates (i.e. WRAL News' weather imagery)

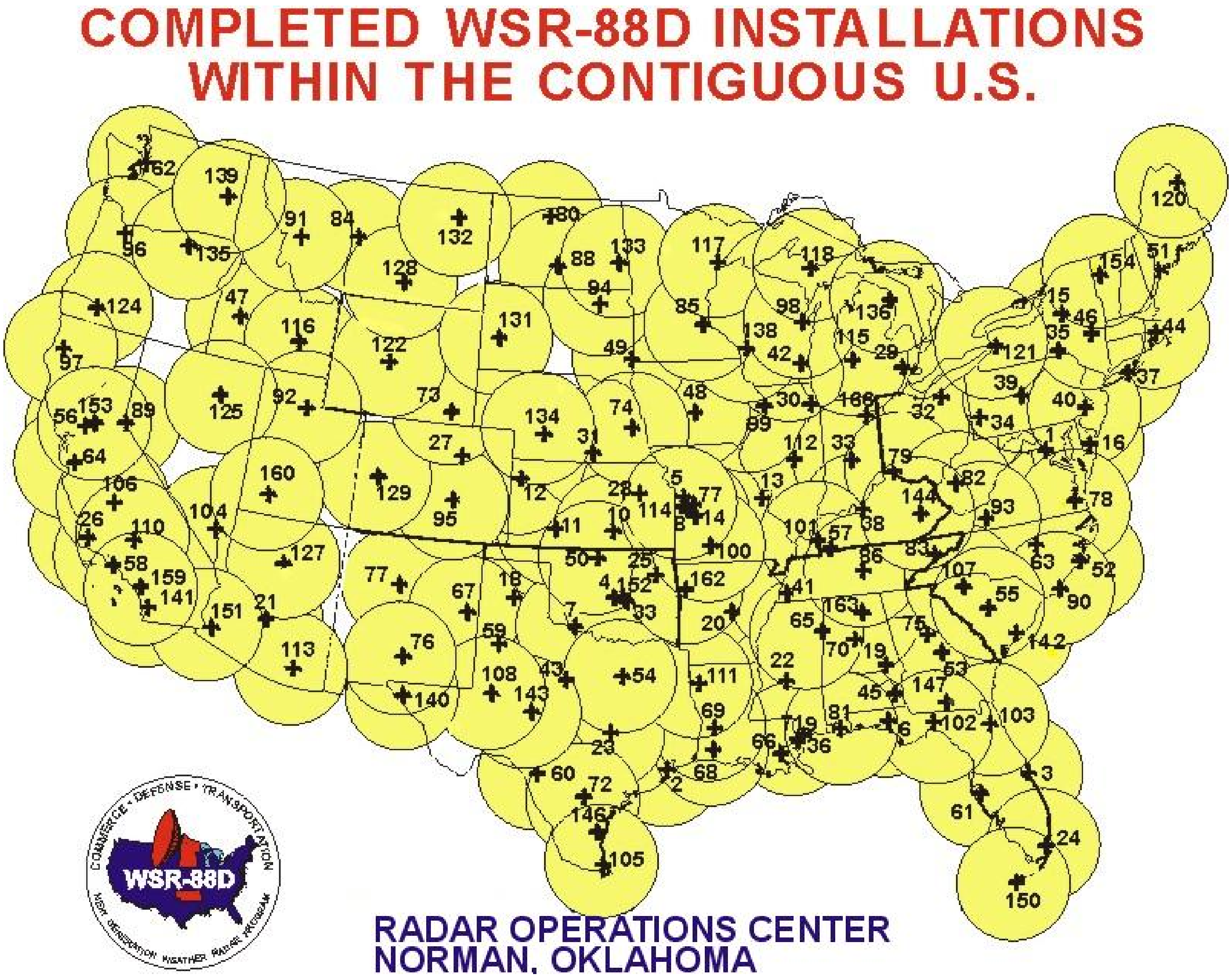
Nexrad Doppler Weather RADAR

- The **Nexrad** network of weather RADAR sensors consists of 158 radars that each have a maximum range of 250 miles that together provide excellent coverage of the continental United States

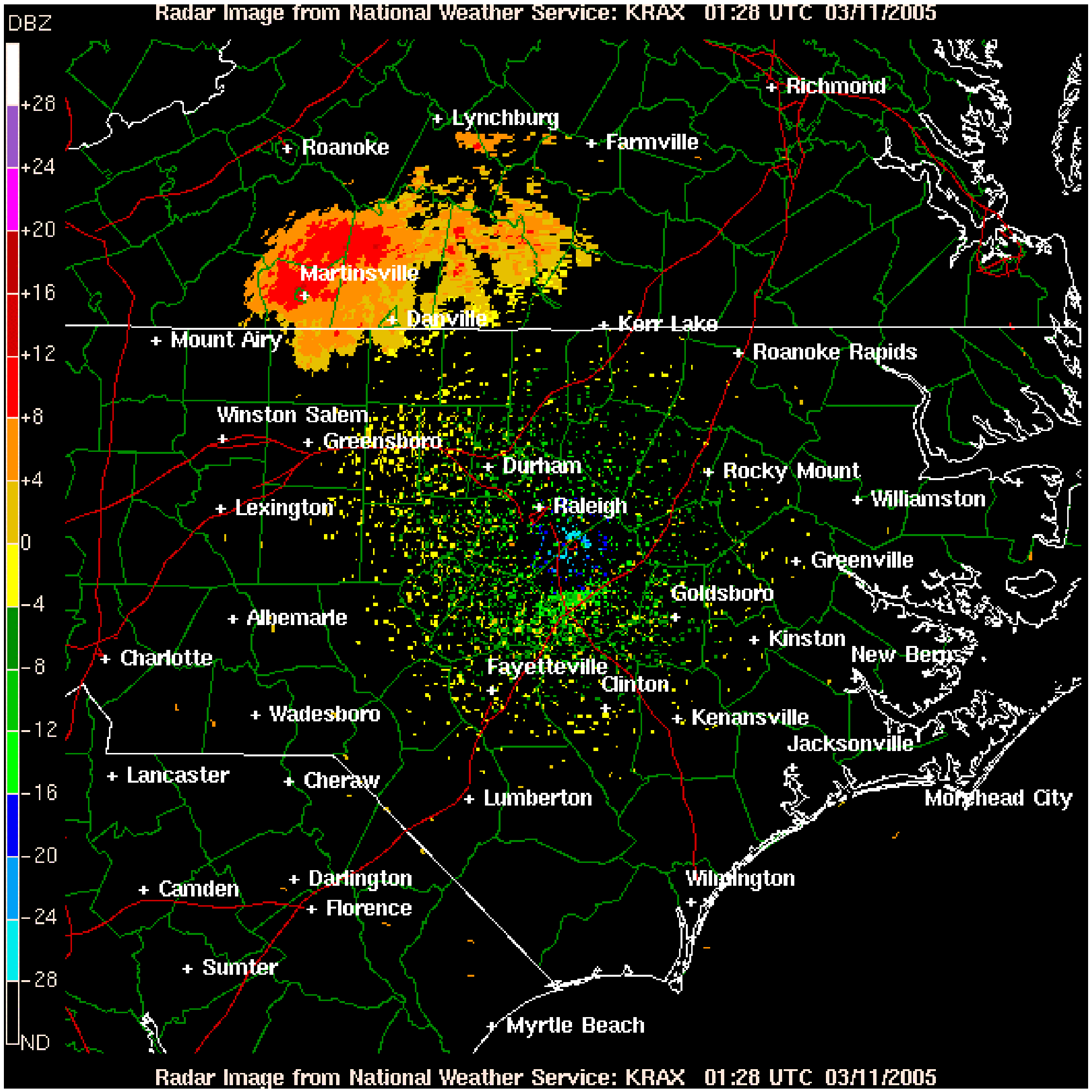
http://www.roc.noaa.gov/

The sensors are known by the designation **WSR-88D** (Weather Surveillance Radar 88 Doppler), and the station in this area is located at RDU airport is #64 - KRAX

Nexrad Doppler Weather RADAR



Nexrad Doppler Weather RADAR

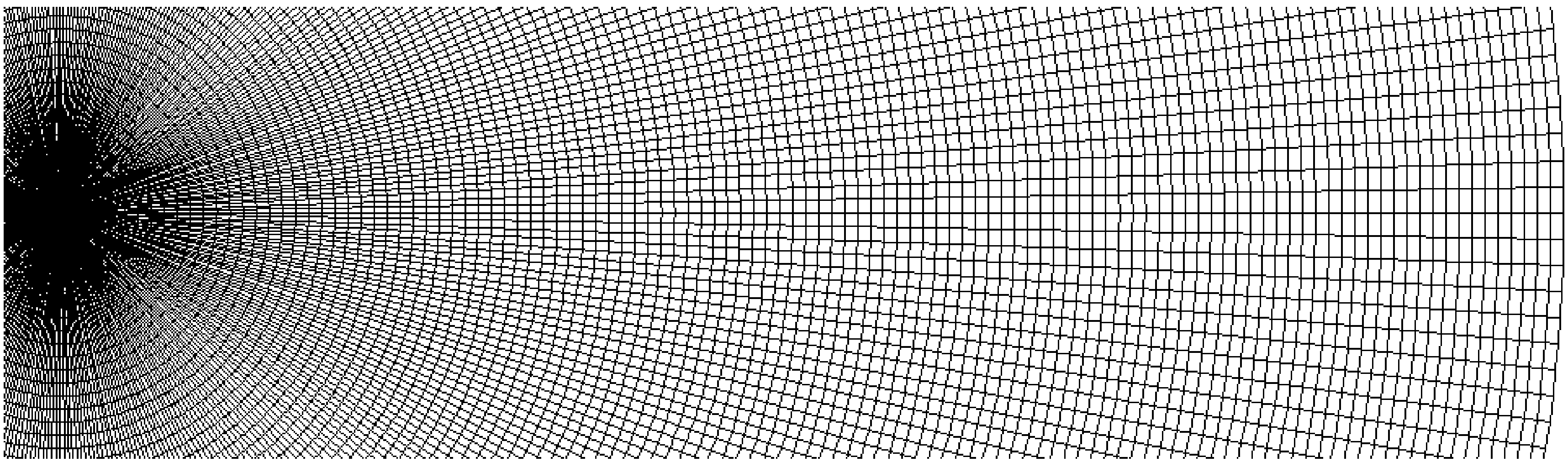


http://weather.noaa.gov/radar/latest/DS.p19r0/si.krax.shtml

- At any time, you can **go online and retrieve a weather RADAR image** for any of the 158 operational stations that is **no more than 10 minutes old** (this one is from KRAX at about 8:30 PM on March 10, 2005)
- Note the **scattered signal** from around the Triangle, and the strong, **organized return** from NW of the RADAR

Nexrad Coordinate Systems

- The **individual** sensors information is referenced using a **polar coordinate system**, the 250-mile radius circle that is sliced into chunks that are **1 degree of arc in width** and 2 **kilometers along the radius** in length:



- This produces units that are **smaller near the sensor**, and **larger as they get further away**, which is an accurate reflection of how a sensor that operates radially collects information about the world

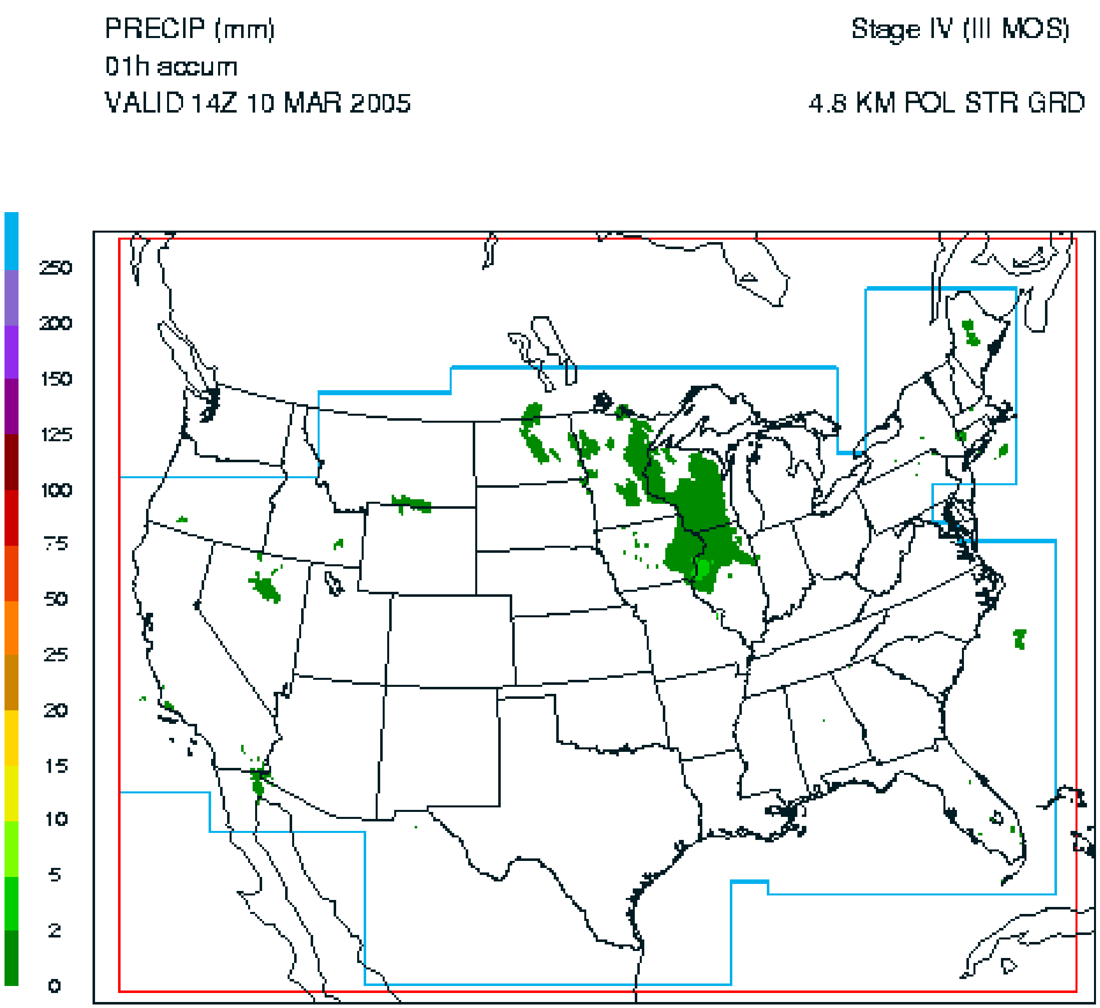
Nexrad Coordinate Systems

•To create regional or national mosaics of radar returns, the 158 RADARs' returns are combined into a **raster grid**, projected in a polar stereographic projection that covers the continental United States in either **4 km or 16 km** cells

•Products are produced at a range of **time scales**: Hourly, 6-hourly, or daily precipitation mosaics for the CONUS can be downloaded from various web sites

•Of course, in addition to the coordinate system transformation, the RADARs' measurements of **returned microwave energy** need to converted into an **estimated amount of precipitable water** in the atmosphere, which is further improved by comparison to rain gauge data

CONUS Hourly Nexrad Rainfall



http://wwwt.emc.ncep.noaa.gov/mmb/ylin/pcpanl/stage4/images/st4.6hrloop.gif

- Here is Nexrad gauge-corrected for **six one-hourly periods** for the afternoon and evening of March 10, 2005

- Note the changes in shape of the **blue bounding box**, which show that some RADARs were offline where no overlapping coverage was present, thus no information was available

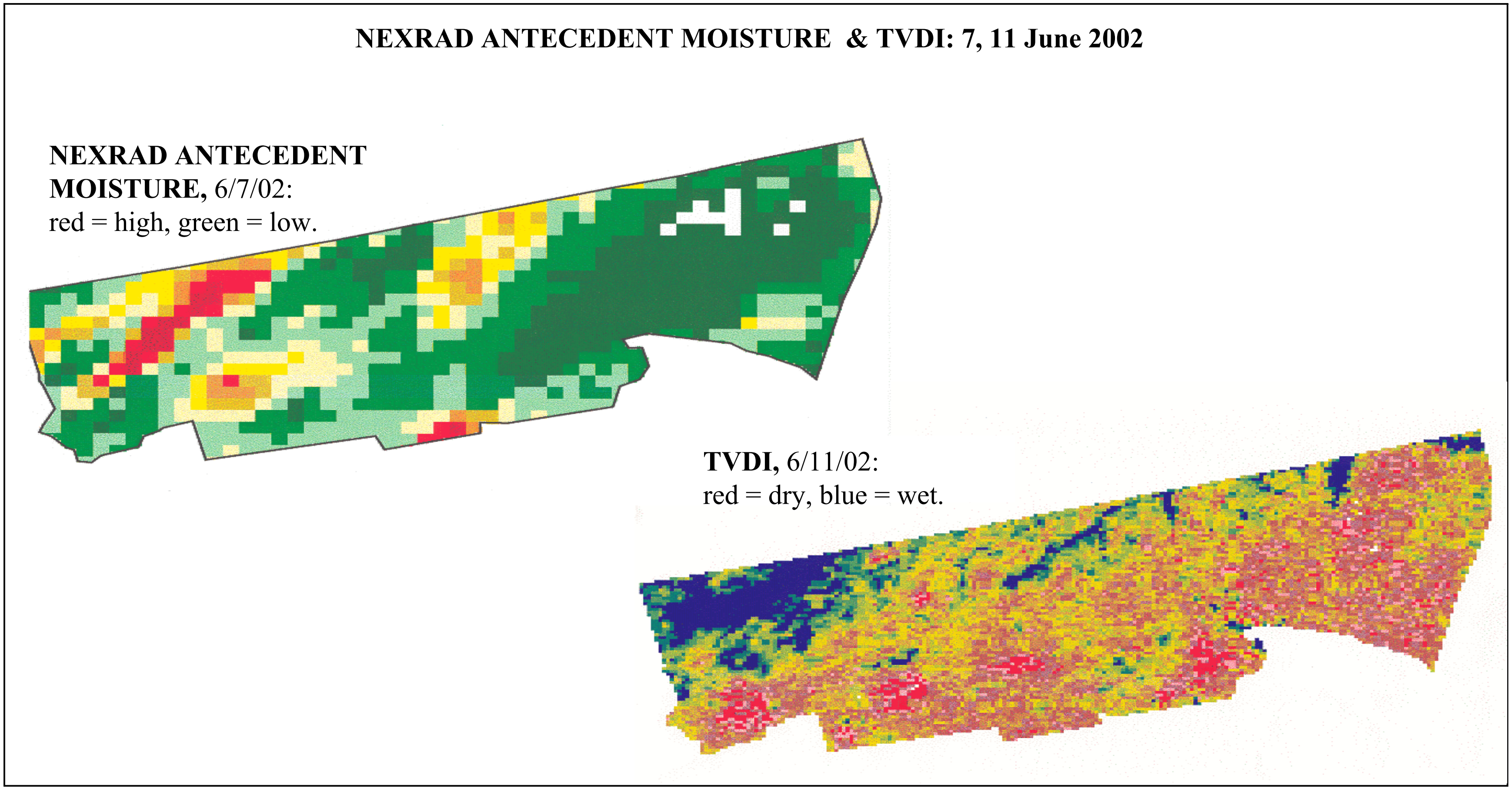
Antecedent Precipitation Index (API) from Stage IV Nexrad Data

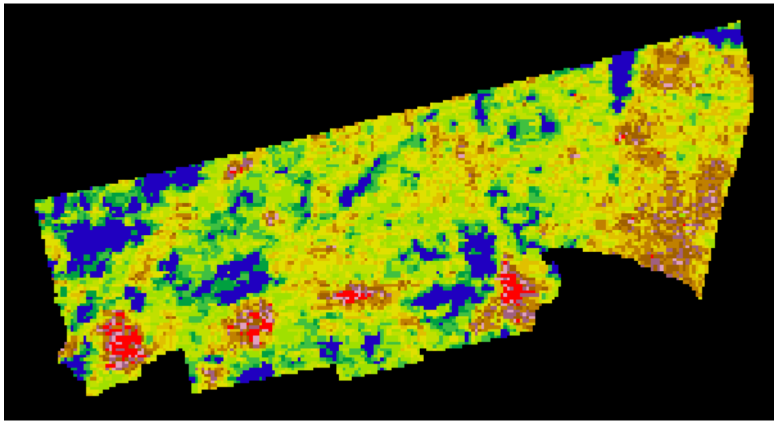
•Successive **daily Stage IV Nexrad rainfall data** were accumulated into an antecedent precipitation index (API) for the study climate divisions for the study period

•The **API** is of the form $I_t = I_0 k^t$ where I_0 is an initialization value, and k is a decay constant (0.9 is a typical value from Dunne & Leopold)

- •For example, assume $I_0 = 5$ mm and $k = 0.9$
- •On $t = 0$, $I_t = 5 \text{ mm} * (0.9^0) = 5 \text{ mm}$
- •On $t = 1$ it rains 1.5 mm, $I_t = 5 \text{ mm} * (0.9^1) + 1.5 \text{ mm}$
  $= (5 \text{ mm} * 0.9) + 1.5 \text{ mm}$
  $= 4.5 \text{ mm} + 1.5 \text{ mm} = 6 \text{ mm}$
- •On $t = 2$ it does not rain, $I_t = 6 \text{ mm} * (0.9^1) = 5.4 \text{ mm}$

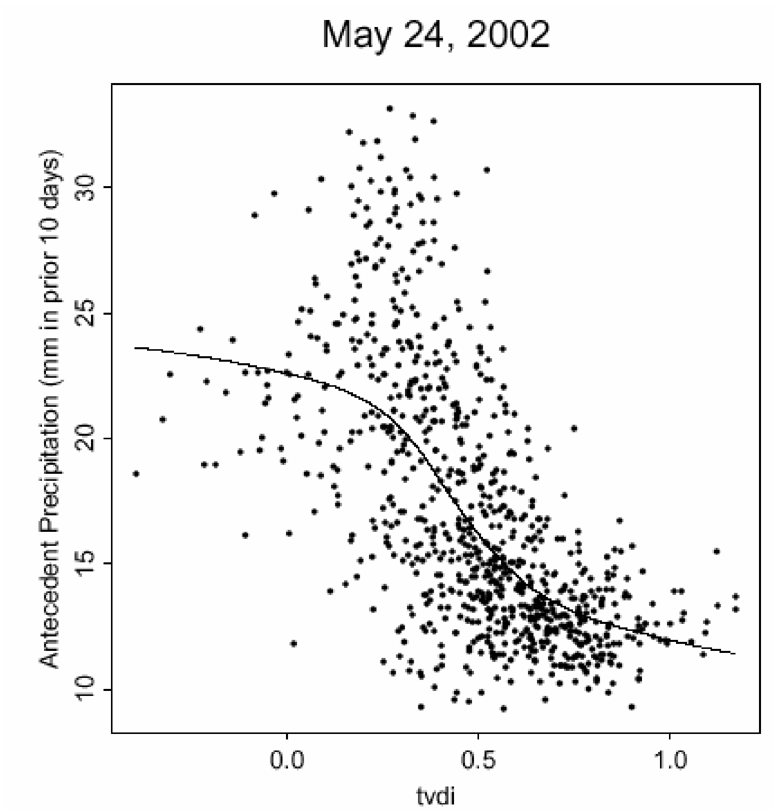
Antecedent Moisture from NEXRAD

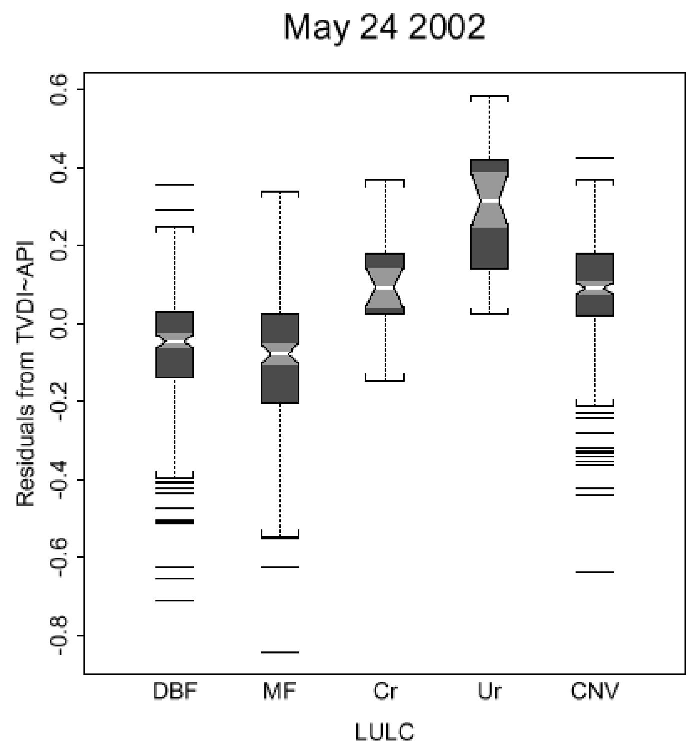


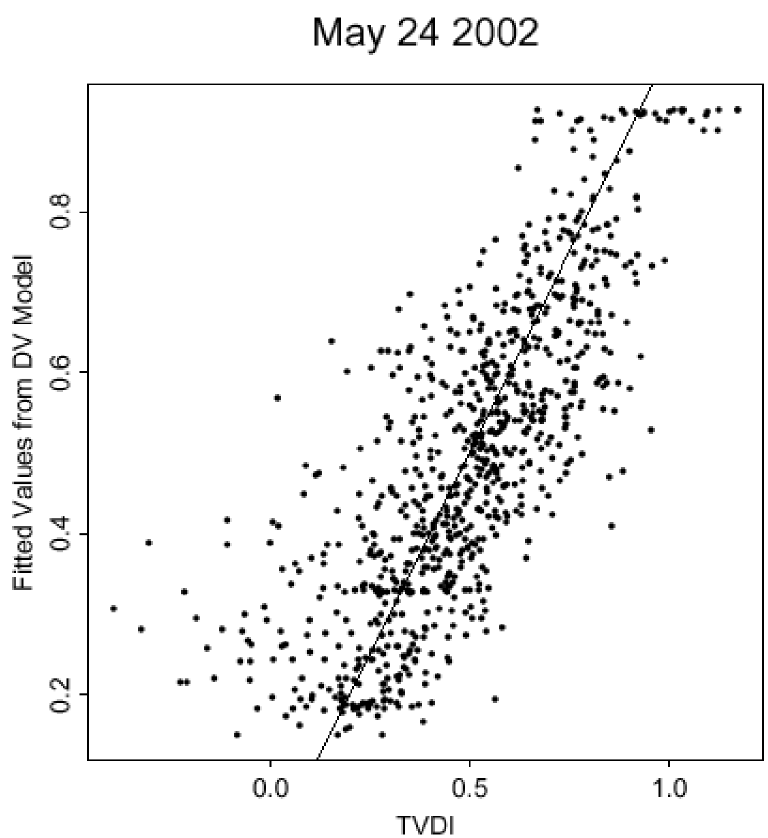


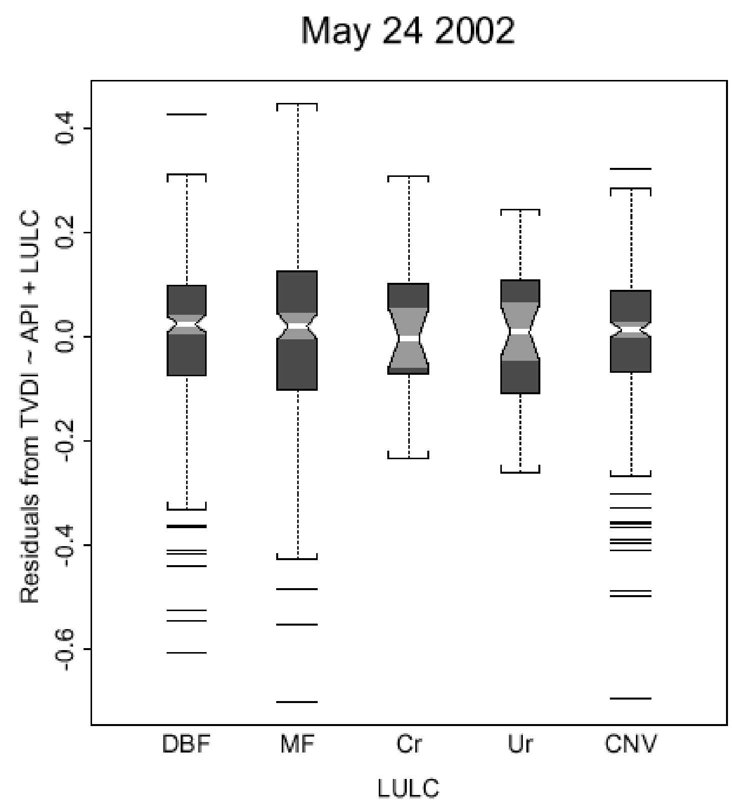
Significant explanation of residuals of plot based on land use/land cover

TVDI variation with API









May 24 2002, Dummy Variables Model, Water Omitted

Coeff.	Value	SE	t-value	Pr(>|t|)
Int	1.5103	0.0651	23.1966	0.0000
API	-0.0811	0.0071	-11.4511	0.0000
API^2	0.0014	0.0002	7.7445	0.0000
DBF	-0.1454	0.0130	-11.1485	0.0000
MF	-0.1778	0.0137	-12.9461	0.0000
Urb	-1.6600	0.7310	-2.2708	0.0234
Urb:API	0.2417	0.0970	2.4927	0.0129
Urb*API^2I	-0.0074	0.0031	-2.3619	0.0184

Residual standard error: 0.1545 on 830 degrees of freedom
Multiple R-Squared: 0.6041
F-statistic: 180.9 on 7 and 830 degrees of freedom, the p-value is 0

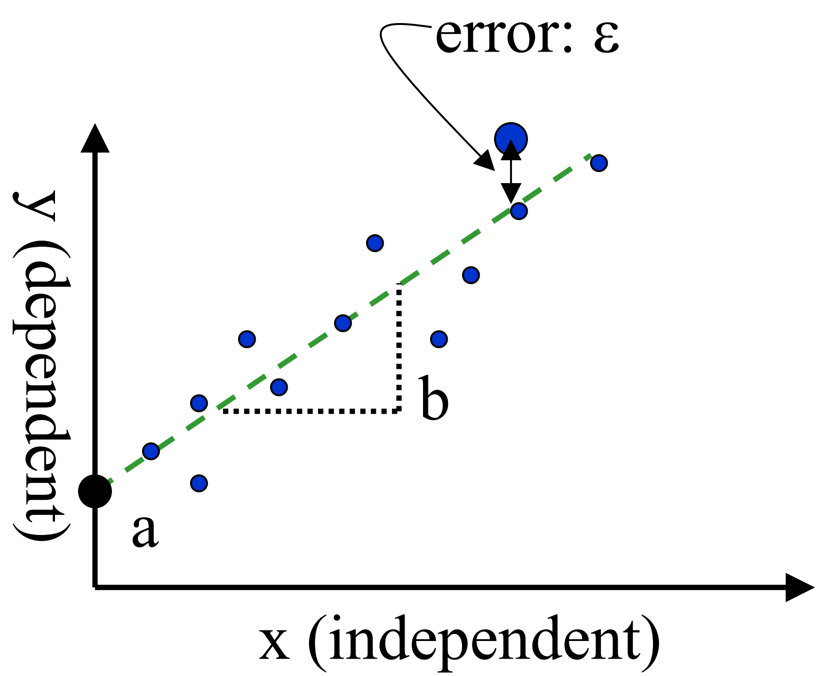
Simple vs. Multiple Regression

- Today, we are going to examine **simple linear regression**, where we estimate the values of a **dependent variable (y)** using the values of an **independent variable (x)**
- This concept can be extended to **multiple linear regression**, where **more** explanatory independent variables (x_1, x_2, x_3 … x_n) are used to develop estimates of the dependent variable's values
- For purposes of **clarity**, we will first look at the simple case, so we can more easily grasp the mathematics involved

Simple Linear Regression

- **Simple linear regression** models the relationship between an independent variable (x) and a dependent variable (y) using an equation that expresses y as a linear function of x, plus an error term:

$$y = a + bx + e$$



x is the **independent** variable

y is the **dependent** variable

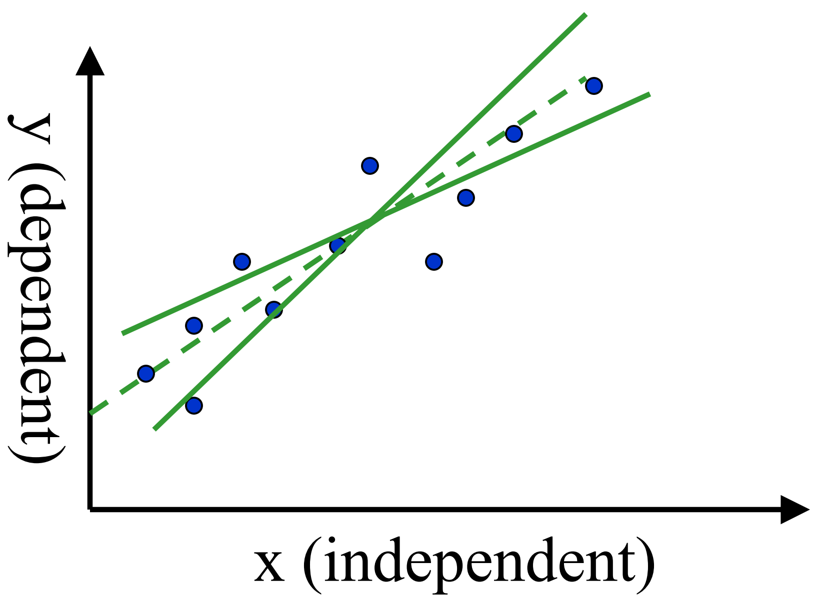
b is the **slope** of the fitted line

a is the **intercept** of the fitted line

e is the **error** term

Fitting a Line to a Set of Points

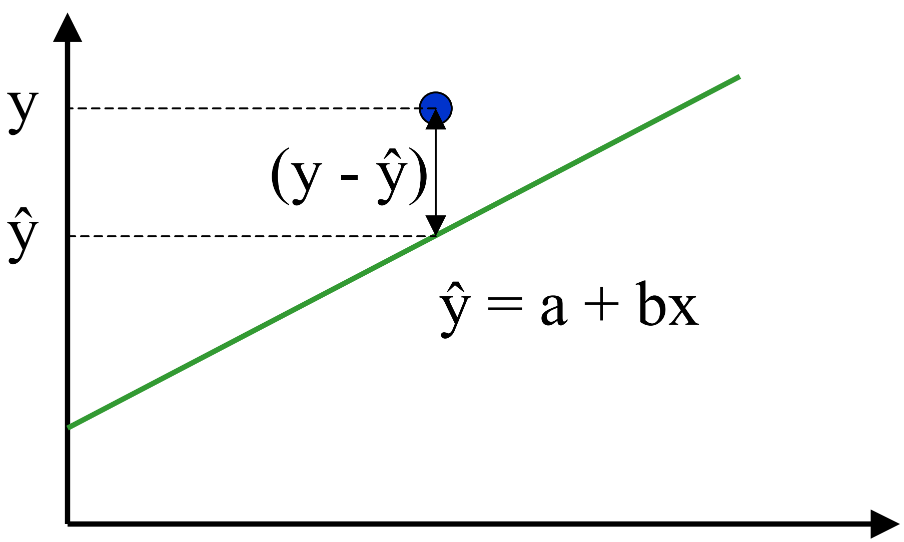
- When we have a data set consisting of an independent and a dependent variable, and we plot these using a scatterplot, to construct our model between the relationship between the variables, we need to **select a line** that represents the relationship:



- We can choose a line that fits best using a **least squares method**
- The least squares line is the line that **minimizes** the **vertical distances** between the points and the line, i.e. it minimizes the **error term** ε when it is considered for all points in the data set

Least Squares Method

- The least squares method operates mathematically, **minimizing the error term e** for all points
- We can describe the line of best fit we will find using the equation $\hat{y} = a + bx$, and you'll recall that from a previous slide that the formula for our linear model was expressed using $y = a + bx + e$



- We use the value $\hat{y}$ on the line to estimate the true value, y
- The **difference** between the two is $(y - \hat{y}) = e$
- This difference is **positive** for points **above** the line, and **negative** for points **below** it

Error Sum of Squares

- By squaring the differences between y and ŷ, and summing these values for all points in the data set, we calculate the **error sum of squares** (usually denoted by SSE):

$$SSE = \sum_{i=1}^{n} (y - \hat{y})^2$$

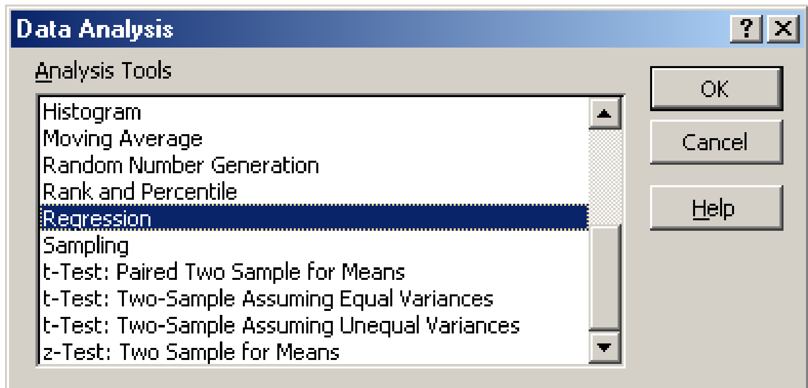
- The **least squares method** of selecting a line of best fit functions by finding the parameters of a line (intercept a and slope b) that **minimizes** the error sum of squares, i.e. it is known as the least squares method because it finds the line that **makes the SSE as small as it can possibly be**, minimizing the vertical distances between the line and the points

Simple Linear Regression in Excel

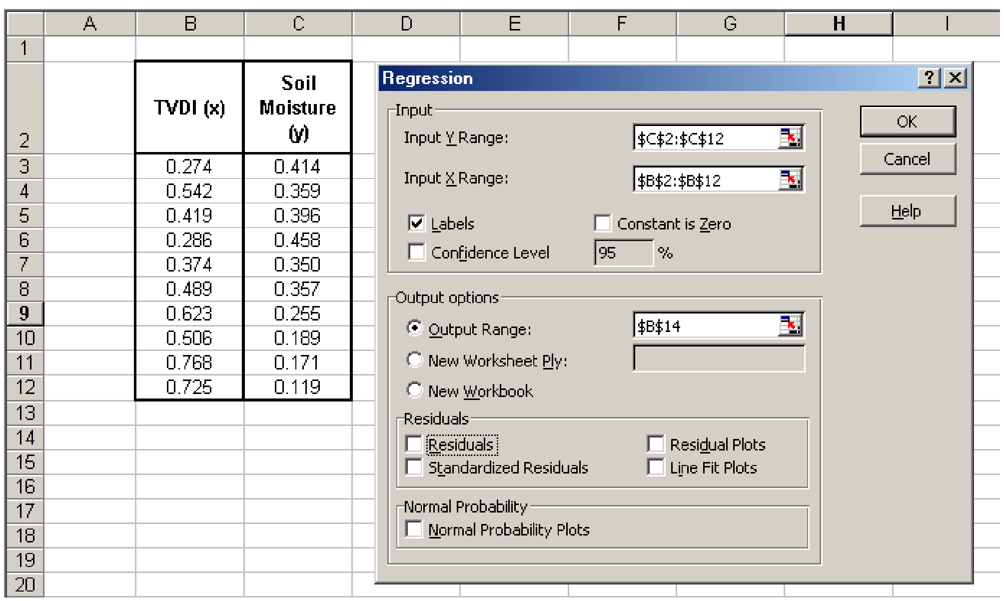
- Excel can calculate regression parameters in **two** ways:
  - There are **built-in functions** that can be entered into a cell to specify the calculation of a regression slope or regression intercept:
    - SLOPE(array1, array2) can be used to calculate the **slope** of the least squares regression line, specifying the y values in array1 and the x values in array2
    - INTERCEPT(array1, array2) can be used to calculate the **intercept** of the least squares regression line, specifying the y values in array1 and the x values in array2
- There is also a **Data Analysis Tool** that can be used to calculate the regression parameters

Regression Analysis Tool

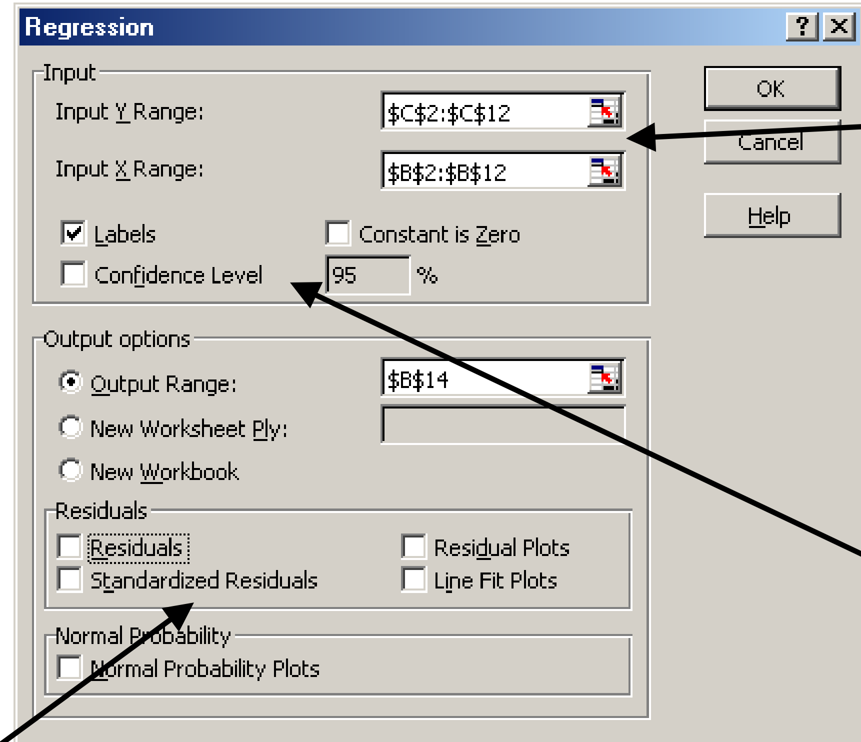
- In the Data Analysis window, **select** the appropriate tool:



- After clicking OK, you'll be presented with the **tool** window:



Regression Analysis Tool



Of course, when specifying the input ranges, you must **distinguish** between the dependent variable (y) and the independent variable(s) (x); this tool can also be used for **multiple linear regression**, so more than one x variable can be used

The tool will **automatically** test the significance of the parameters at the 95% confidence level, but if you check the checkbox and **specify another confidence level**, it will test the significance of the regression parameters at that level of confidence **as well**

Checking boxes in the **Residuals** portion of the tool will produce other output including calculating the **residuals** for each value, calculating standardized residuals, and **plotting residuals** versus independent variables, and line fit plots as well

Regression Analysis Tool

- The **basic output** the tool produces includes:

The coefficient of determination (r^2)

The standard error of the estimate (e.g. the standard deviation of the residuals), s_e

An ANOVA table, including the minimum α where F would be **significant**

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.87163053
R Square	0.75973978
Adjusted R Square	0.72970725
Standard Error	0.05996834
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.090973945	0.09097394	25.2972303	0.001014626
Residual	8	0.028769614	0.0035962		
Total	9	0.119743559			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0.60320076	0.061926011	9.74066875	1.0324E-05	0.46039903	0.74600249
TVDI (x)	-0.5923931	0.117780521	-5.0296352	0.00101463	-0.863995597	-0.3207905

The regression coefficients produced by the least squares optimization (in the simple case, like this one, the intercept and the slope)

The standard error associated with each parameter (e.g. for the regression slope parameter, this is s_b, the standard deviation of the slope)

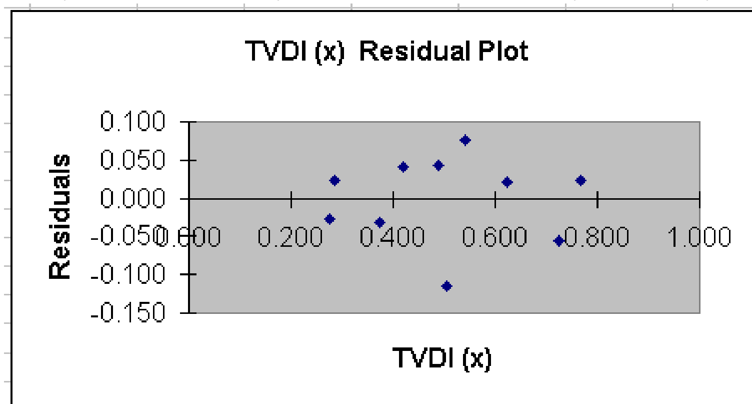
The t-statistic and the minimum α where each parameter would be significant

Regression Analysis Tool

•Checking the Residuals checkbox will produce a **table of the regression estimates** (the $\hat{y}_i$ values) and **residuals**:

RESIDUAL OUTPUT		
Observation	*Predicted Soil Moisture (y)*	*Residuals*
1	0.441	-0.027
2	0.282	0.077
3	0.355	0.041
4	0.434	0.024
5	0.382	-0.031
6	0.313	0.043
7	0.234	0.020
8	0.304	-0.115
9	0.148	0.023
10	0.173	-0.055

•Residual Plots creates a **scatter plot of the residuals versus x** (this is useful for checking assumptions about the residuals):



•Line Fit Plots creates a scatter plot of the **actual and predicted values versus x** (this is useful for getting a visual sense of the accuracy of the estimates):