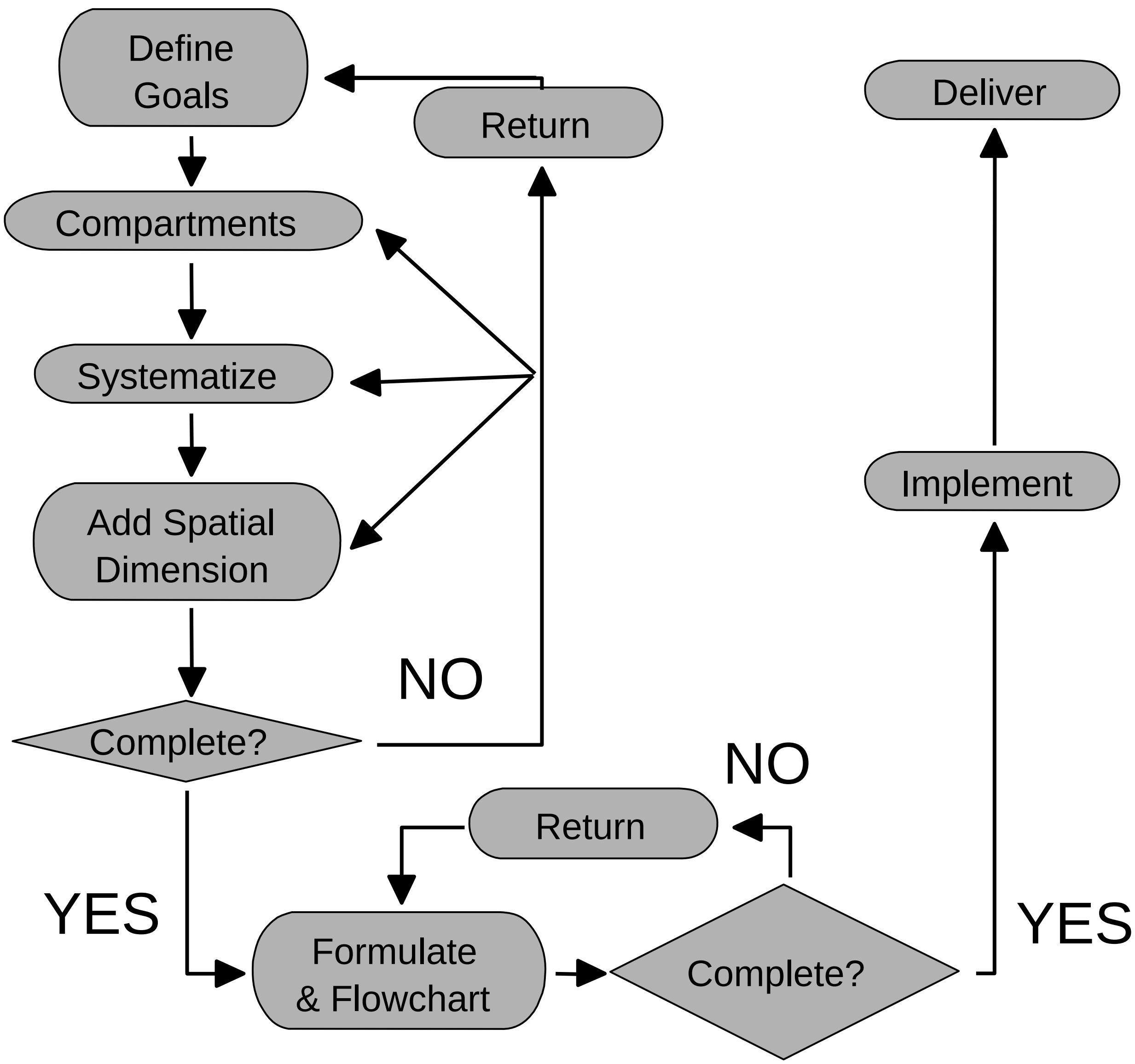
Model Design and Evaluation

The General Modeling Process



Some Difficult Problems

- **Spatial conflicts**
  - Ecosystem function vs. development factors
    - More (conflicting) demands on the **same resources**
    - Non-preference (or objection) but **still need to allocate**
      - LULU (locally unacceptable land uses)
      - NIMBY (not in my back yard)
  - Multiple development preferences
    - One use often **decreases suitability of another** (siting power line: preempts, and land used)
  - Suitable to multi-purposes
    - **Ranking** and restriction(s)

Spatial Conflict - Solutions

- **Solutions** are messy, difficult, often iterative
  - **GIS is of assistance** to solution
    - **Visualization as basis** for descriptive modeling
    - **Analytical methods** for prescriptive
    - Prescriptive often requires **direct interaction with decision makers** during the modeling process

Spatial Conflicts

- Two types of spatial criteria (e.g. **LESA**)
  - **Site criteria**
    - Deal with the **direct impact on the actual site**
      - e.g. size of parcel, current zoning, within a flood zone are all site factors related to viability for agricultural use
  - **Situation criteria**
    - Deal with the **impact on the surrounding area**
      - Neighborhood criteria
        - » Off site
      - These require us to know what **particular land use** is to occupy the site and its **potential impact on the surrounding area**
        - » The situation
  - LESA: Aims to include zoning, aesthetic, compatibility and infrastructure in **decisions**

Situation Criteria

- **Other situation criteria** that need to be addressed:
  - How far from the site is the **impact zone**
    - (e.g. the **distances used** in the LESA factors)
      - Often designed to be **very conservative**
      - Also chosen based on **rather arbitrary** decisions
  - **Explicit definition** of what is being impacted
    - e.g. for **each potential land use conversion** are that we concerned about… off-site hydrology, wildlife, land values, etc.
    - Often **more difficult to define** for situation than for site
    - Often **more spatially extensive** in outcome
  - Requires us to **limit** what is to be considered

Generating Alternatives

- GIS is capable of **generating alternative uses**
  - **Rating parcels** by capability or its suitability
  - The idea of **generating constraints** (multi-criteria)
    - Logic and ranking
  - Results in maps that show where certain land uses **cannot be supported**
    - These maps provide:
      - A set of **possible land use allocation solutions**
      - A set of **alternatives**
  - Overlay of these shows **many possible land uses** in certain areas of any region
    - May require us to **tighten our constraints**

The Orpheus Method

- The method to **facilitate the allocation process** (spatial conflicts)
- **Descriptive component**
  - Illustrate the **potential** of each possible land use
    - Examines site criteria **for each proposed use**
      - Conservation, agriculture, forestry, ……
      - Constraints
    - Then examines the situation criteria
- Results in a **suitability map for each possible use**
  - Each a descriptive model
  - Each has two parts:
    - Site suitability map
    - Situation suitability map

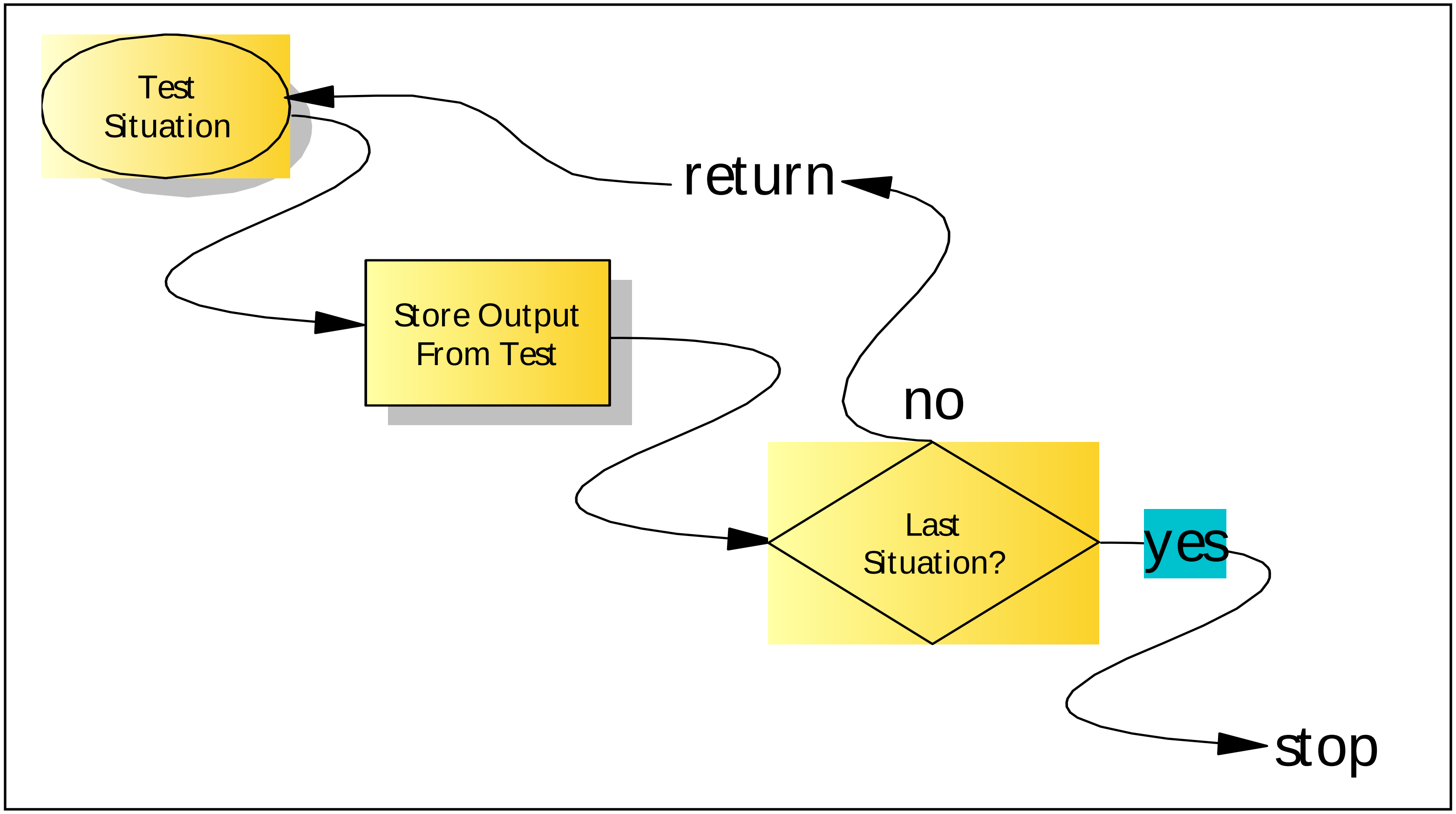
The Orpheus Method Cont.

- To **facilitate** the allocation process (continued)
- **Prescriptive component**
  - An **iterative** process
  - Used to generate the **allocation** of land uses
  - Essentially, a **search for stability**
    - What achieves **most constraints and most uses**
    - Requires methods of **defining stopping rules**

Conflict Resolution

- **Consensus building**
  - **General agreement** on terms, conditions, and limits
  - Best started during **descriptive phase**
- **Hierarchical methods**
  - Weighting and re-weighting
    - Interactive, involving **decisions about importance**
- **Displaced Fuzzy Ideal**
  - Take advantage of **fuzzy logic** to resolve multiple considerations

The Iterative Prescriptive Component



An Example using LESA

- **Land Evaluation and Site Assessment** Model
  - Designed to evaluate land-use at the **county level**
    - Focus is the proper **allocation** of agr. land for non-agr.
  - The goal is to **preserve good farmland**
  - Review some **relationships** that form the LESA model (slides in the **Building Models** slideshow)
    - Infrastructure, socioeconomic factors, and zoning regulation
  - Now prescribe, rate **every land site**

Hierarchical Compartmentalization

LESA Model Components (Compartments)

1. Land use / agriculture (local and adjacent)
2. Agricultural economic viability (investment, size)
3. Land use regulations (% zoned agr., adjacent to zone)
4. Alternative locations (availability, productivity)
5. Compatibility of proposed use (surrounding hydrology, for example)
6. Compatibility with master plans
7. Infrastructure (city, transportation, utilities)

Adding the Spatial Dimension

- **Land use / agriculture**
  - Existence of **agricultural land** in and around the proposed land conversion sites
    - **Three components** (that are basically spatial in nature):
      - Land on site
      - Land adjacent to site
      - Land within a specified distance of site
  - **Size of farm** (which is fundamentally spatial)
    - Restricts use
  - **Agrivestment** (aspatial)
    - Do we abandon this factor?
    - Or do we find spatial surrogates?

Finding Spatial Surrogates

- **Agrivestment**
  - Obtain **financial records** (machinery, roads, buildings, improvements…)
    - Divide investments **by farm size** (a ratio)
    - Or we could **incorporate a $/ft^2** for individual items
  - Alternatively, we could use agrivestment as a **non-spatial multiplier or operator**
    - e.g. average annual investment **for each farm**

Accuracy Assessment and Measurement

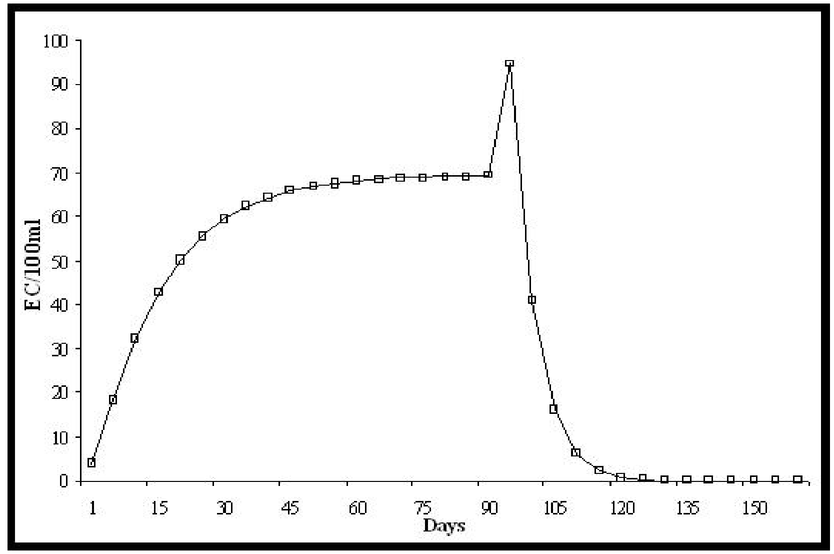
- The **final step** (in remote sensing and GIS and modeling)
  - Explanation and acceptance (**correct and useful**)
  - Describing environmental processes and **the utility of results** in doing so
  - Not only the production of results, but ensuring that **they have meaning** (i.e. not just pretty pictures)

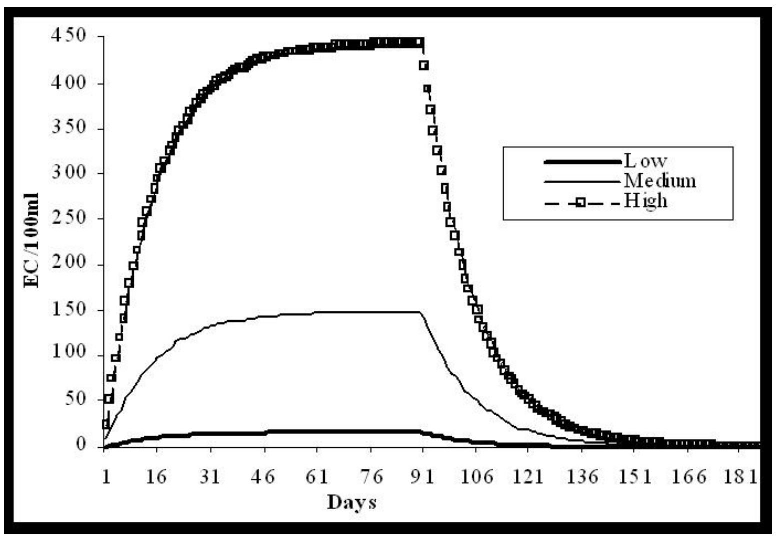
Defining the Terms

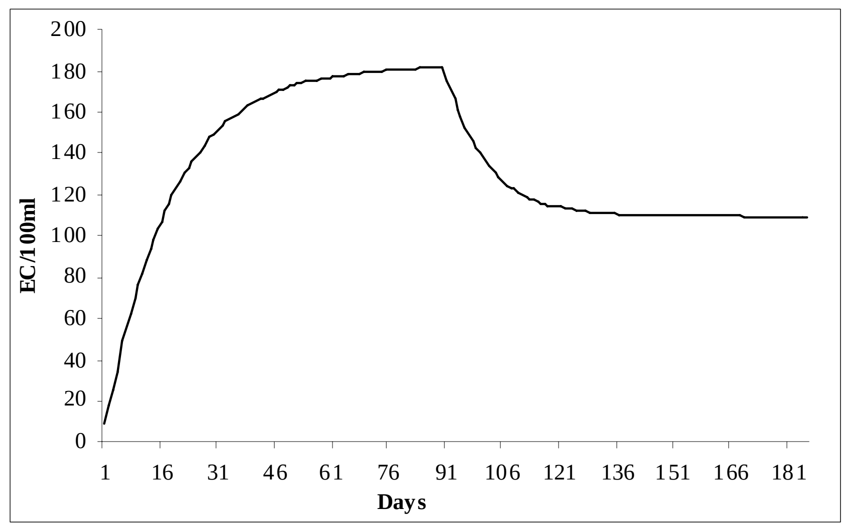
- **Verification**
  - Computational code and algorithms
    - **Correctness**: numerical values (not always the case?), **repeatability** of equation performance
    - **Consistency**: Desired results are **consistent** (application to application)
    - Conceptual level (ground truth **unimportant** here?)
    - Ideally **both** at component level and entire system
    - Can be **cartographic or non-cartographic**
- **Validation**
  - Does the model **correctly represent the real world**? (are the abstractions adequate / well applied?)
- **Acceptability**
  - **Does it work** as a decision-making tool?
  - Both Verification and Validation **support** Acceptability

Non-Cartographic Verification

- Verify the responses by **varying one variable** (and often by examining time series output)



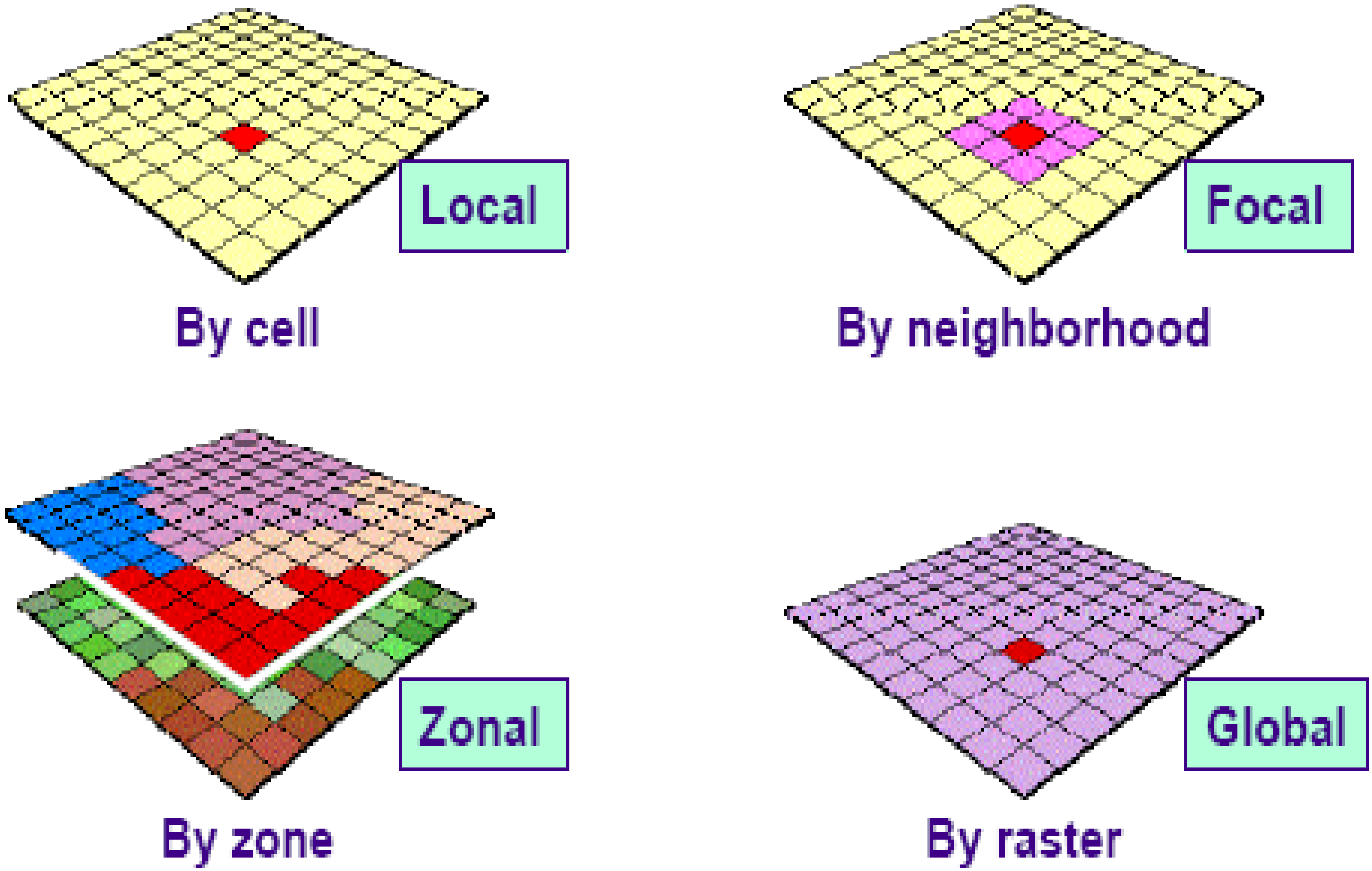




Verifying GIS Map Data

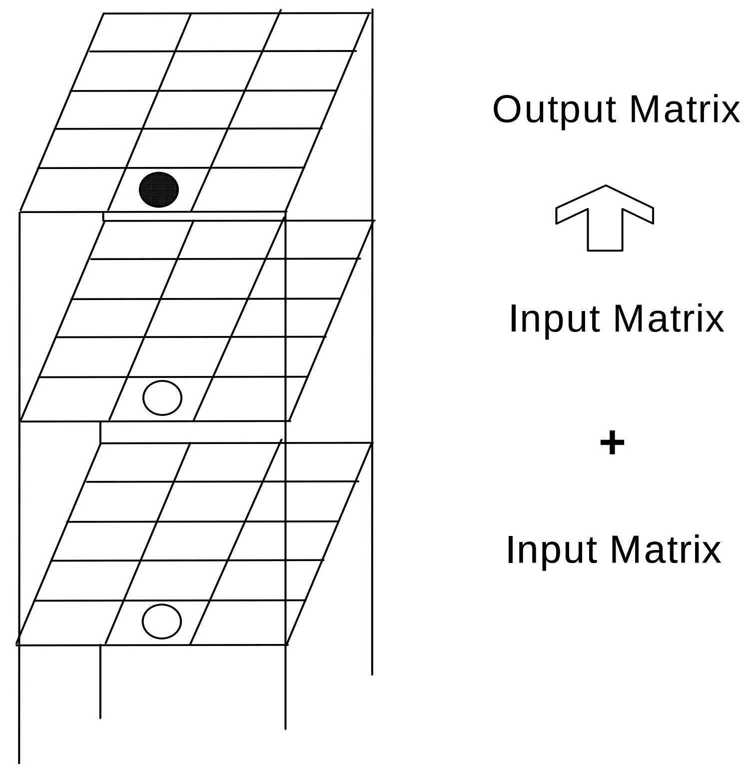
- Several **decisions** to be made in performing **cartographic verification**:
  - Selecting a **process** for testing the algorithm
  - Selecting a **useful portion** of data set to use for comparison
  - Selecting the **size of dataset** to be used for testing
- Necessary to have an **expectation** about the output beforehand (what should it look like?)

Verification and Function Types



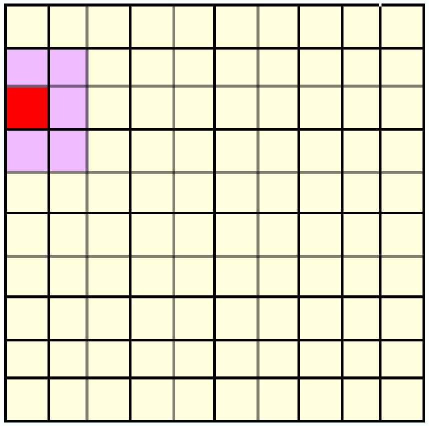
Verifying Local Functions

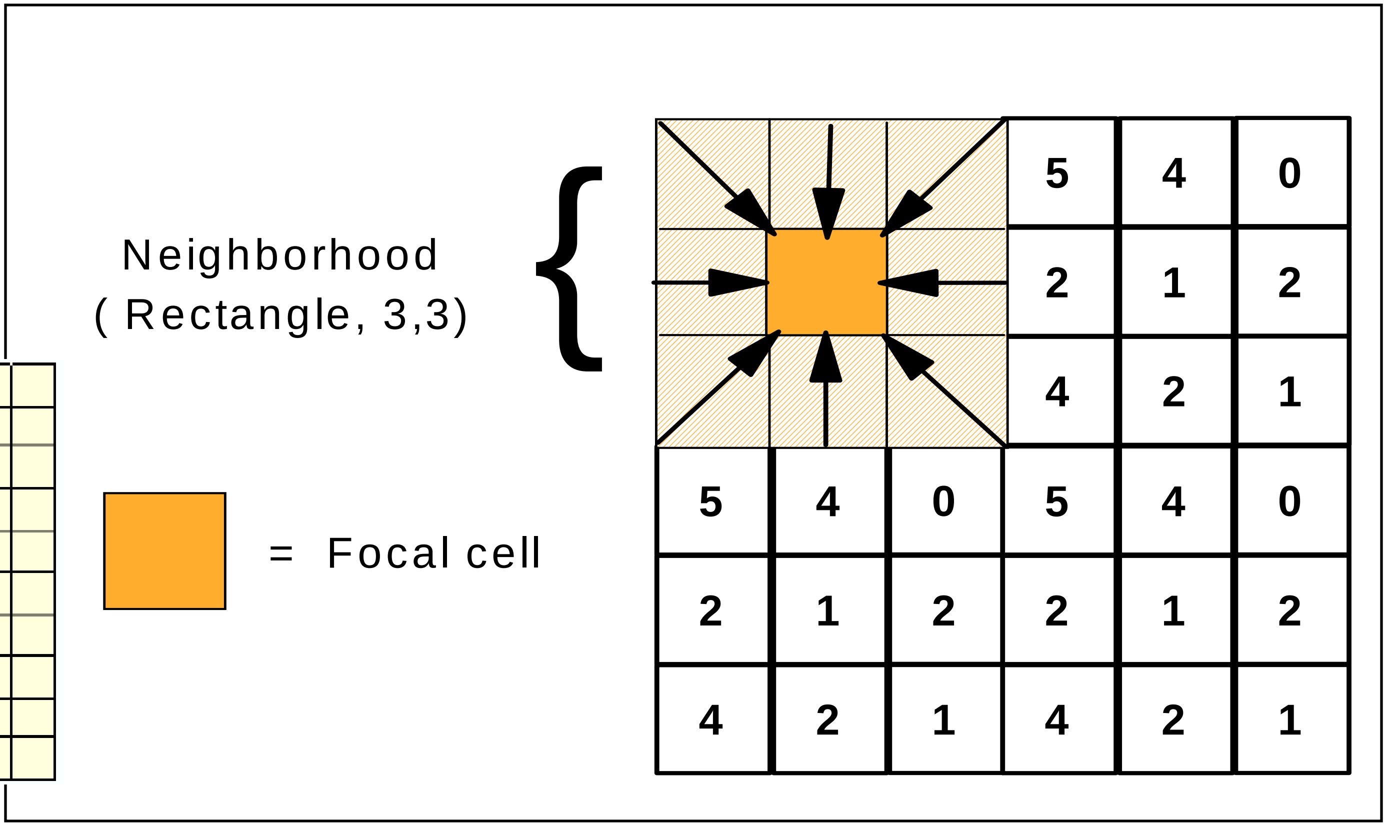
- Recall that local function use **by-cell** operations
  - Select cells for verification that **represent important categories or values** in the model
  - If we can verify adequate function in these locations, we have **confidence** that it is adequate



Focal Functions

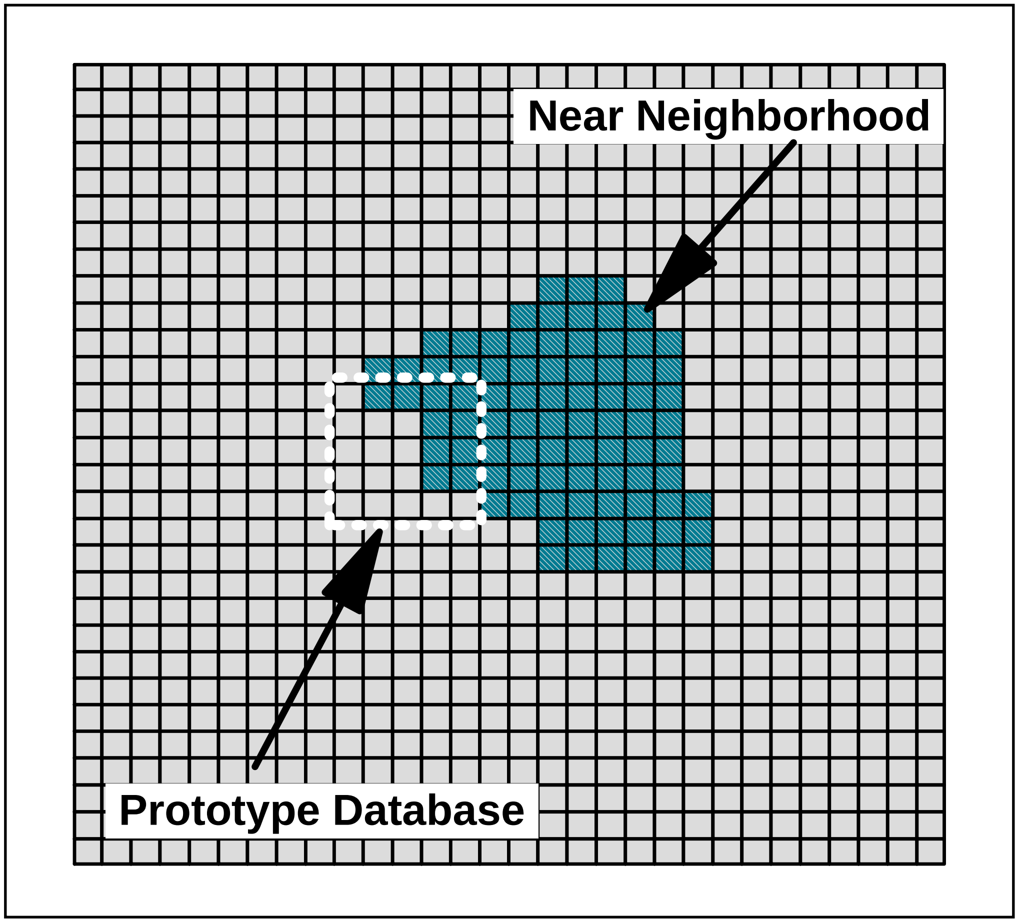
- Evaluate **individual cells** (focal cell) based on the conditions of a **neighborhood of surrounding cells**
  - **Output** consists of a **single cell** at a time
  - **Neighborhood** is a moving window of **input**





Verifying Focal Functions

- Select some cells for examination that are **within** the near neighborhood, and some **outside** of it
- This way we can check to make sure the function is **working correctly**

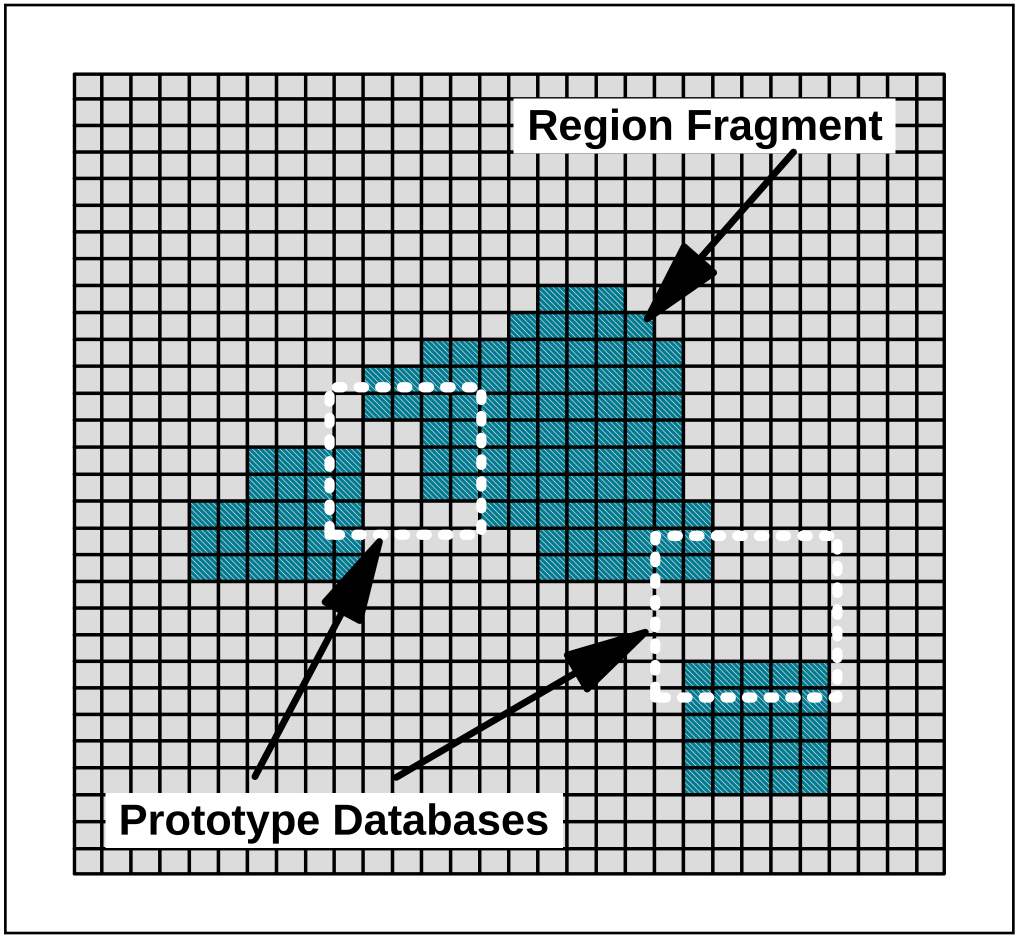


Zonal Functions

- By **zone** (formal regions)
  - Contiguous, fragmented, perforated (whatever!)
  - Zones **defined in a separate grid** (i.e. 2 grids required)
  - **Statistical operands** used **per zone**
    - Min, max, majority, mean, median, std dev, variety, range, sum, etc.

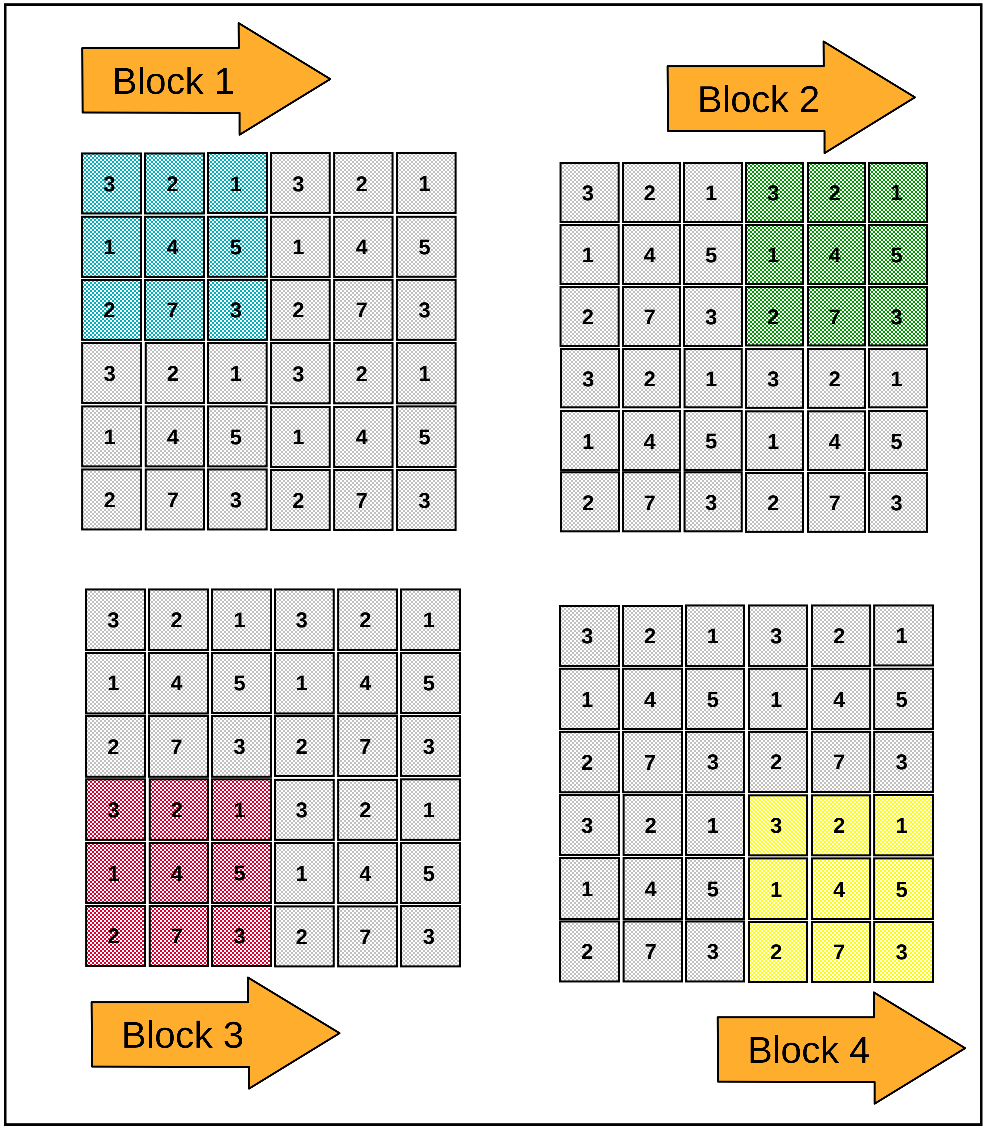
Verifying Zonal Functions

- Select two or three locations that **contain both regional and extraregional portions**, and also test if there is any effect of **fragmentation**
- Basically, trying to **include all possible scenarios**, rule out problems



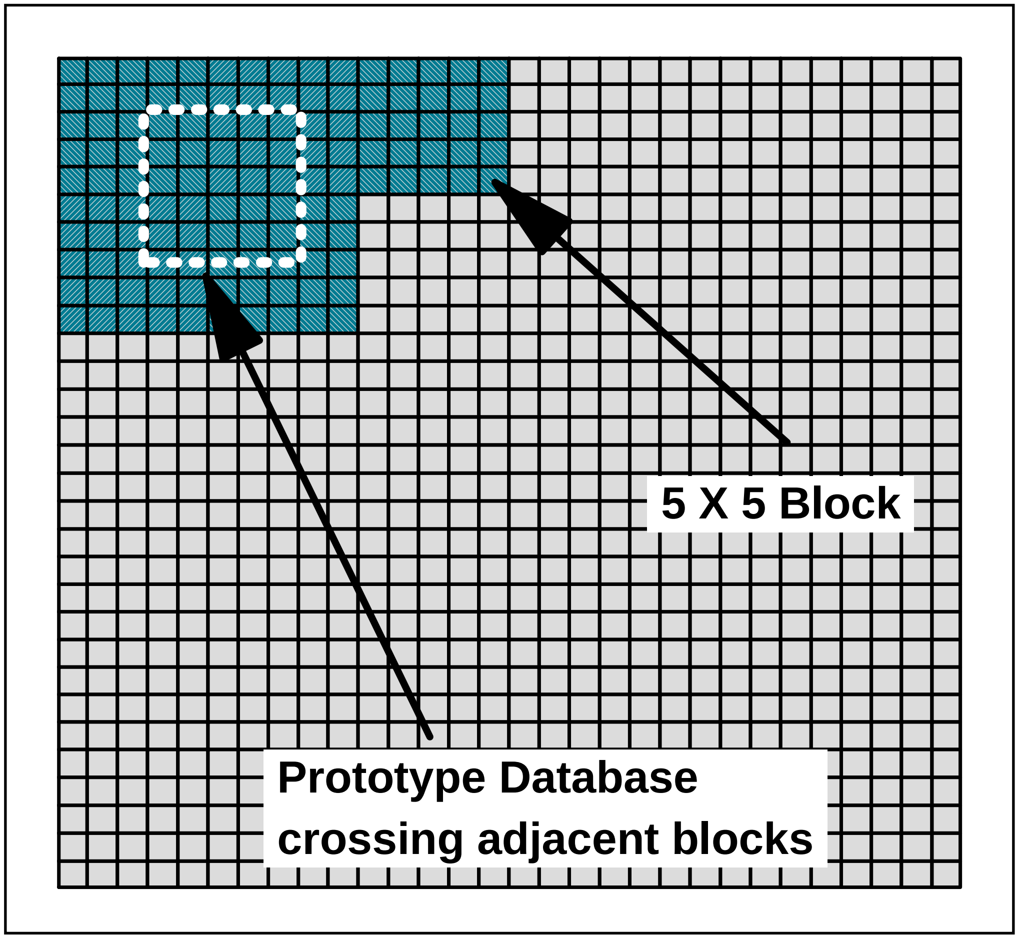
Block Functions

- **Modified versions** of focal functions
  - Use **moving window**
  - **Unique block** at a time
  - Uses **typical** operands
    - Min, max, majority, mean, median, std dev, variety, range, sum, etc.



Verifying Block Functions

- Select a prototype database that **crosses multiple blocks** to make sure the function is operating properly from one adjacent block to another
- That is, use one sample to look at 2 or 4 blocks at once; an **efficient** way to check

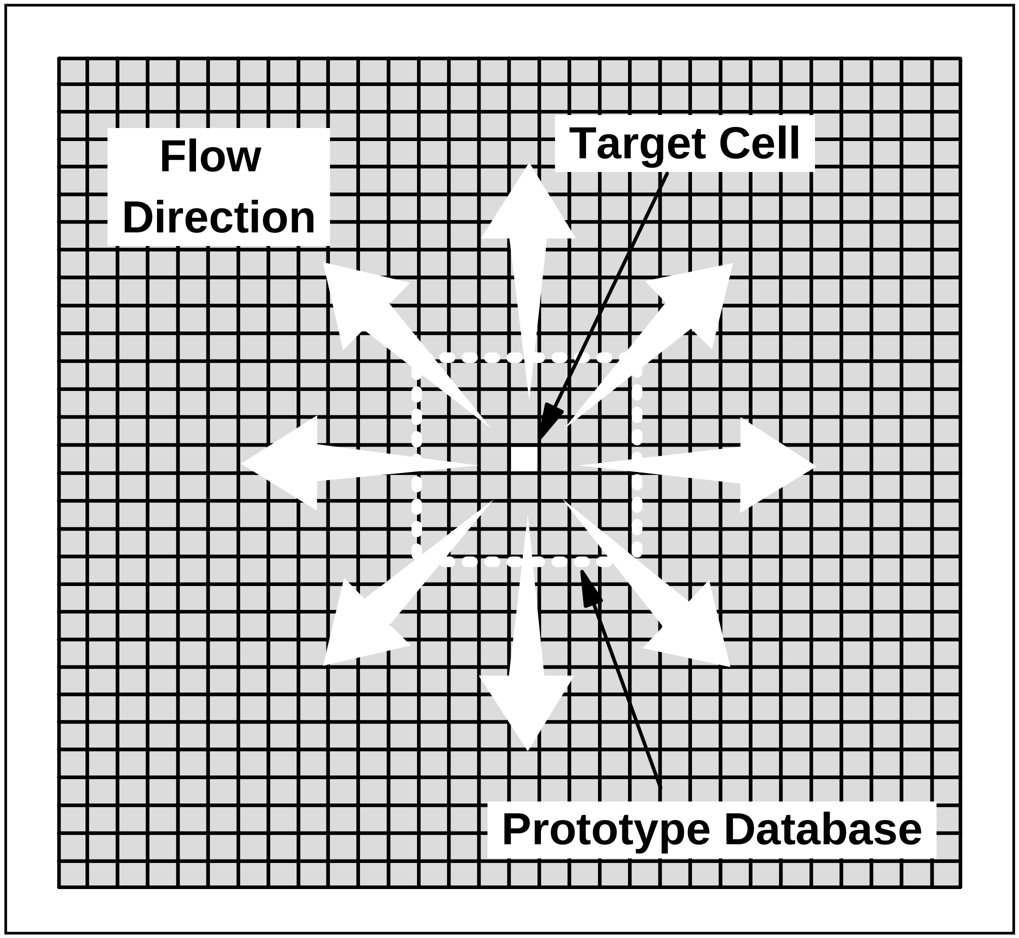


Global Functions

- Truly a **bird's eye view**
  - Considers the **entire grid at once**
  - Output may be **functionally related to every grid cell** in one or more grids at any given time
    - Software must have access to **all grid cells**
- Groups of global functions **radically different:**
  - Euclidean distance global functions
  - Weighted distance global functions
  - Surface global functions
  - Hydrologic global functions
  - Groundwater global functions
  - Multivariate global functions

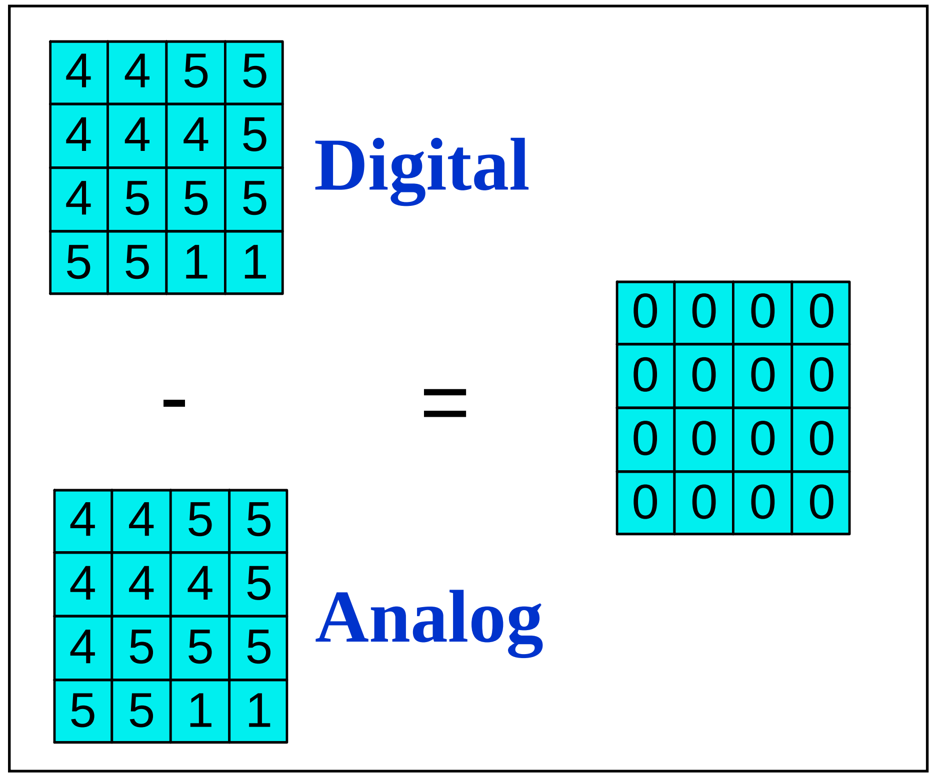
Verifying Global Functions

- Perhaps the **hardest** to verify, because of global nature of input
- A safe method is to **start at a target cell**, and then check successive **groups of cells that surround that cell** to see if the results make sense
- **Replicating algorithm manually** …



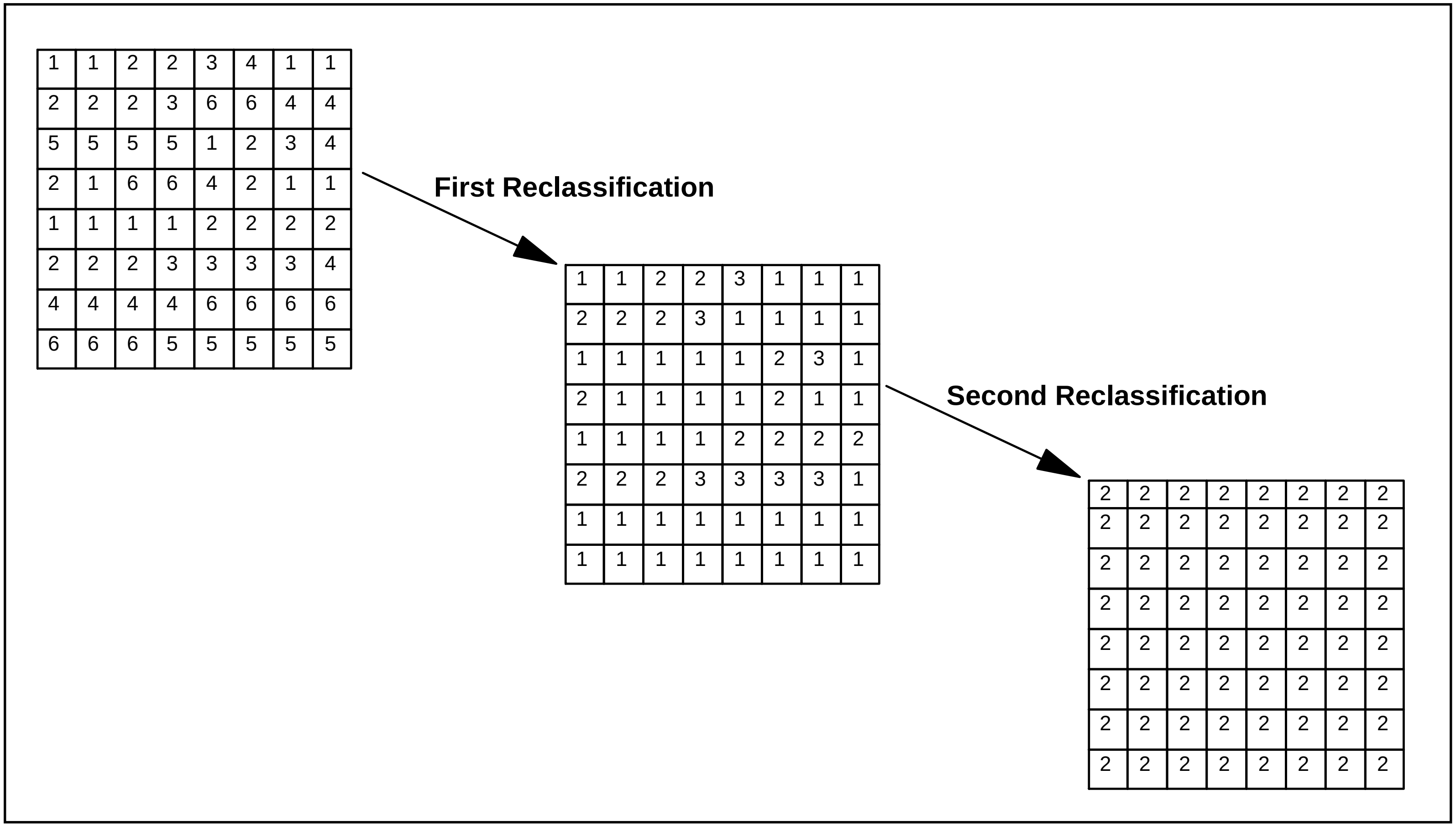
Verification Analog vs. Digital

- In all cases, what we are doing is trying to **reproduce the output** of the function **manually**, and **comparing** it to the computer-generated version
  - i.e. an analog vs. digital comparison of the output … their **difference should be 0**, and thus be identical if all is well and working properly



Impact of Successive Reclassification

- Are our operations in the **correct order** / possessing the **distributive** property?



Problems in Logic

- An **easy way to make mistakes** in raster GIS modeling is to **use numerical scales improperly**
  - This is reasonably **easy to do** because:
    - Regardless of the scale of measurement, all are **encoded in rasters as values** → you have to either have the metadata or remember what the values mean!
  - e.g. Nominal values **multiplied** by ratio value
    - 5, 10, and 15 for urban, agriculture and vacant land
    - Multiplied by 10, 20, and 30 feet produces … gibberish?

Model Validity

- How well does the model **mimic / represent / approximate** reality?
  - **Tomlin**'s approach: "Assuming my logic is correct, and the algorithms correctly implement that logic, then the model is valid."
  - A **better approach**: Ask the following - "Is the model actually modeling what I think it is?"
    - **Visit** the site (field checking)
    - **Compare** to validation dataset
    - Small area **prototyping** (has time constraints)
      - Might require **sub-setting the dataset** (frowned on by many modelers as invalid)
    - **Statistical analysis** (regression or similar predictive models as a check)

Parsimony and its Importance

- How **elegant** is your model (and why is that important)?
  - Easier to explain
  - Easier to check for correctness
  - Easier to understand complex situations
  - Easier to refine and expand
- Ways of **measuring** parsimony:
  1. Number of steps
  2. Simplicity of steps
  3. Amount of computation time
  4. Ease of comprehension
  5. Number of iterations
  6. Ratio of parsimony to model thoroughness

Model Acceptance

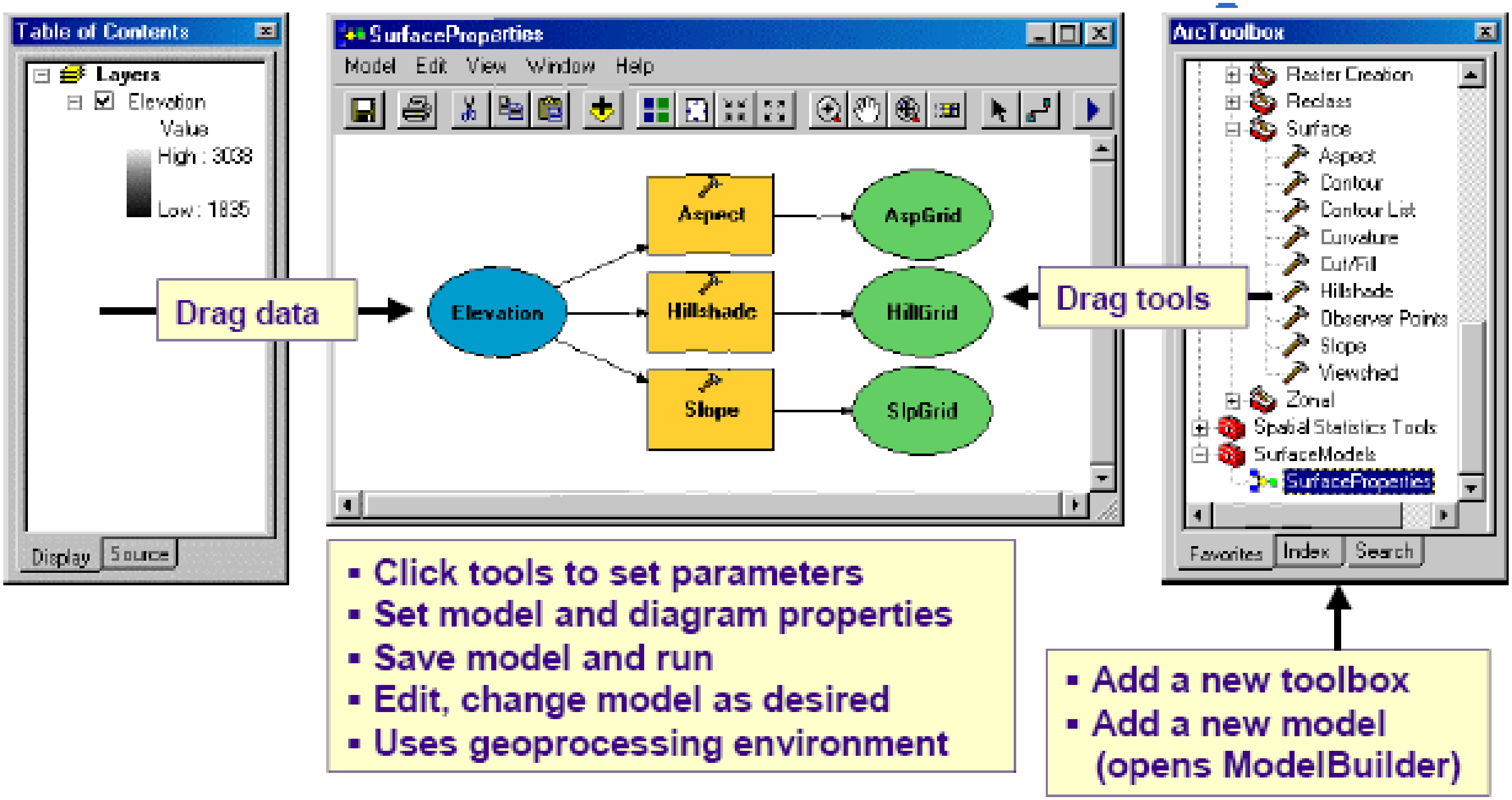
- This may be the **most important** step
- Does the **client accept it** as a decision tool?
  - May even determine **whether or not you get paid**
- Provide **tests** of model results
- **Demonstrate** the GUI for applications
- Is the model **provided in time** to be used?
- Are there **any parts missing**?
- The key is to keep the **client involved throughout** the process → Ownership of process & product

Exercise 10: Building Models

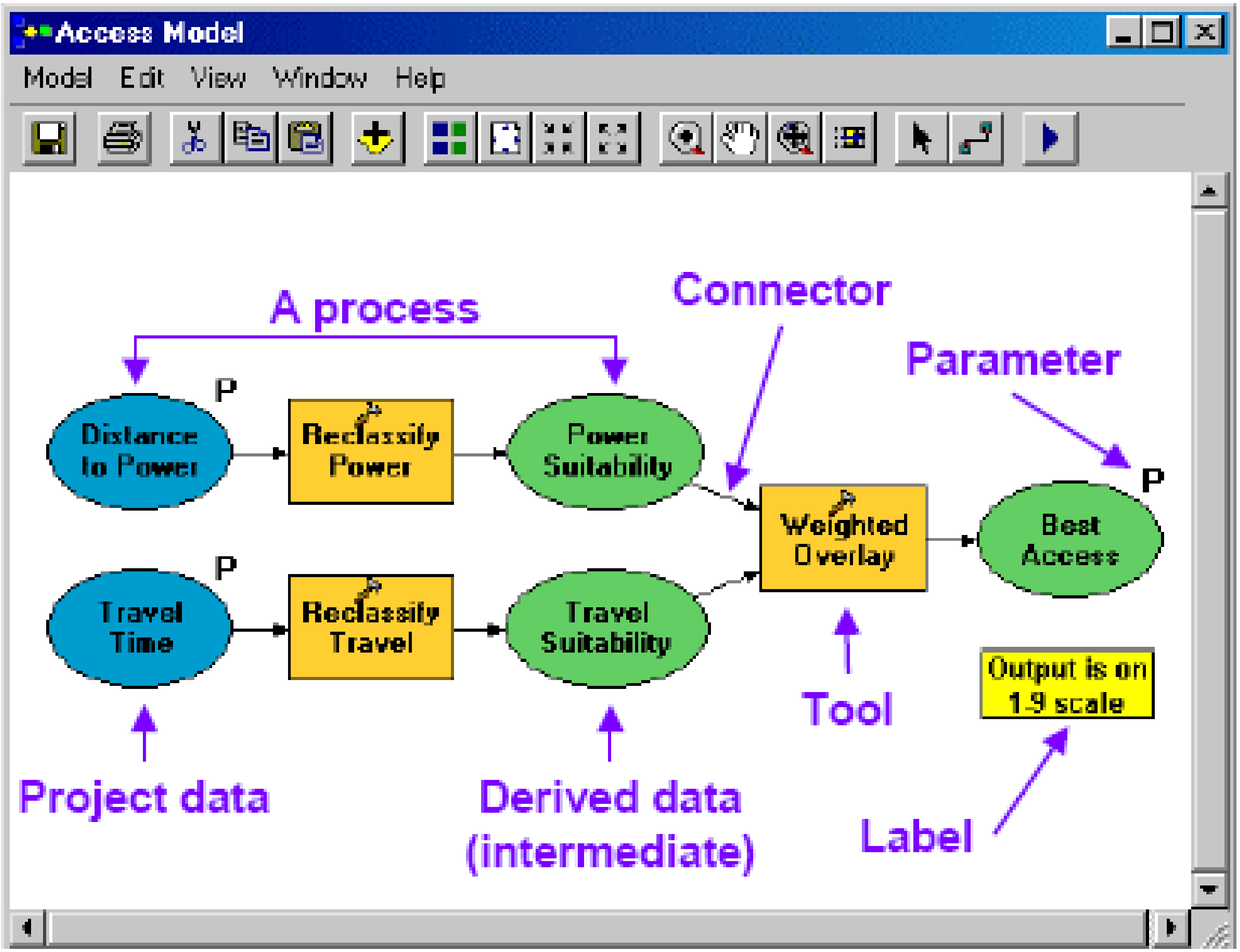
- EXERCISE 10A: MODELING TECHNIQUES AND TOOLS
- EXERCISE 10B: MODELBUILDER AND WEIGHTED SUITABILITY

Building models with ModelBuilder

- A **graphical modeling environment** for ArcGIS



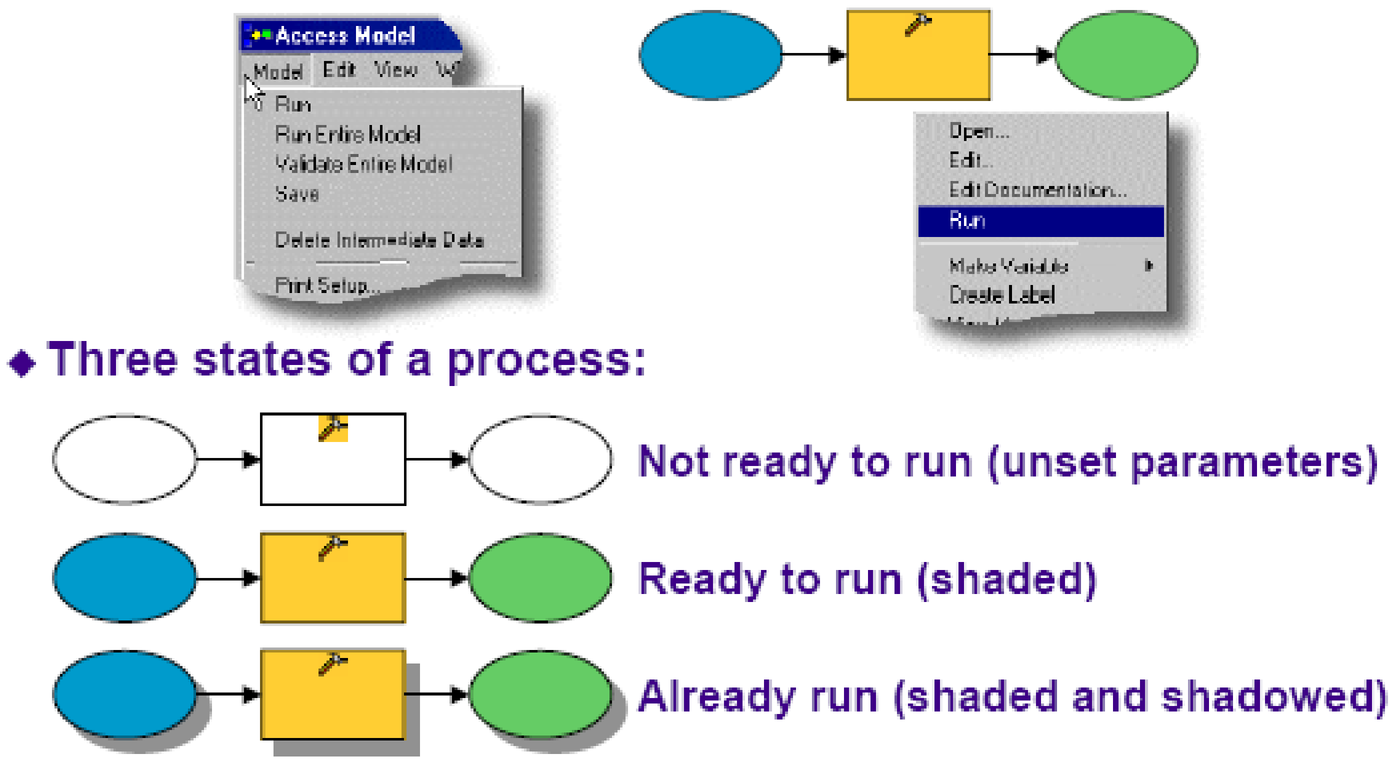
Model elements



- All elements have **context menus**

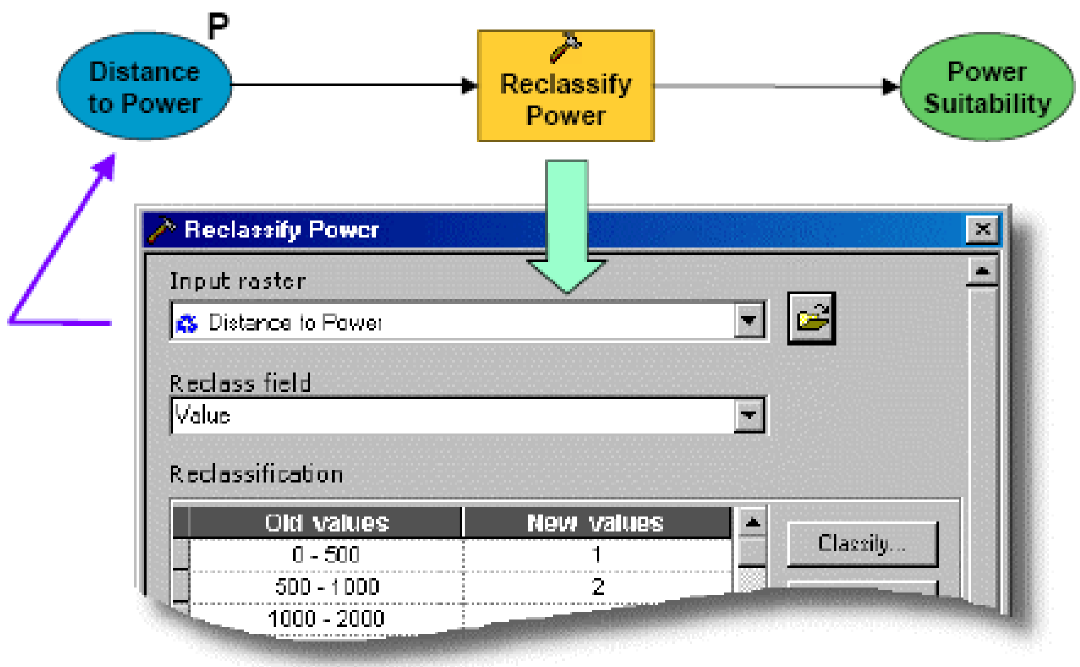
Running a Model

- Two **options**:
  - Run from ArcToolbox like other tools
  - Run in the ModelBuilder window—all or one process



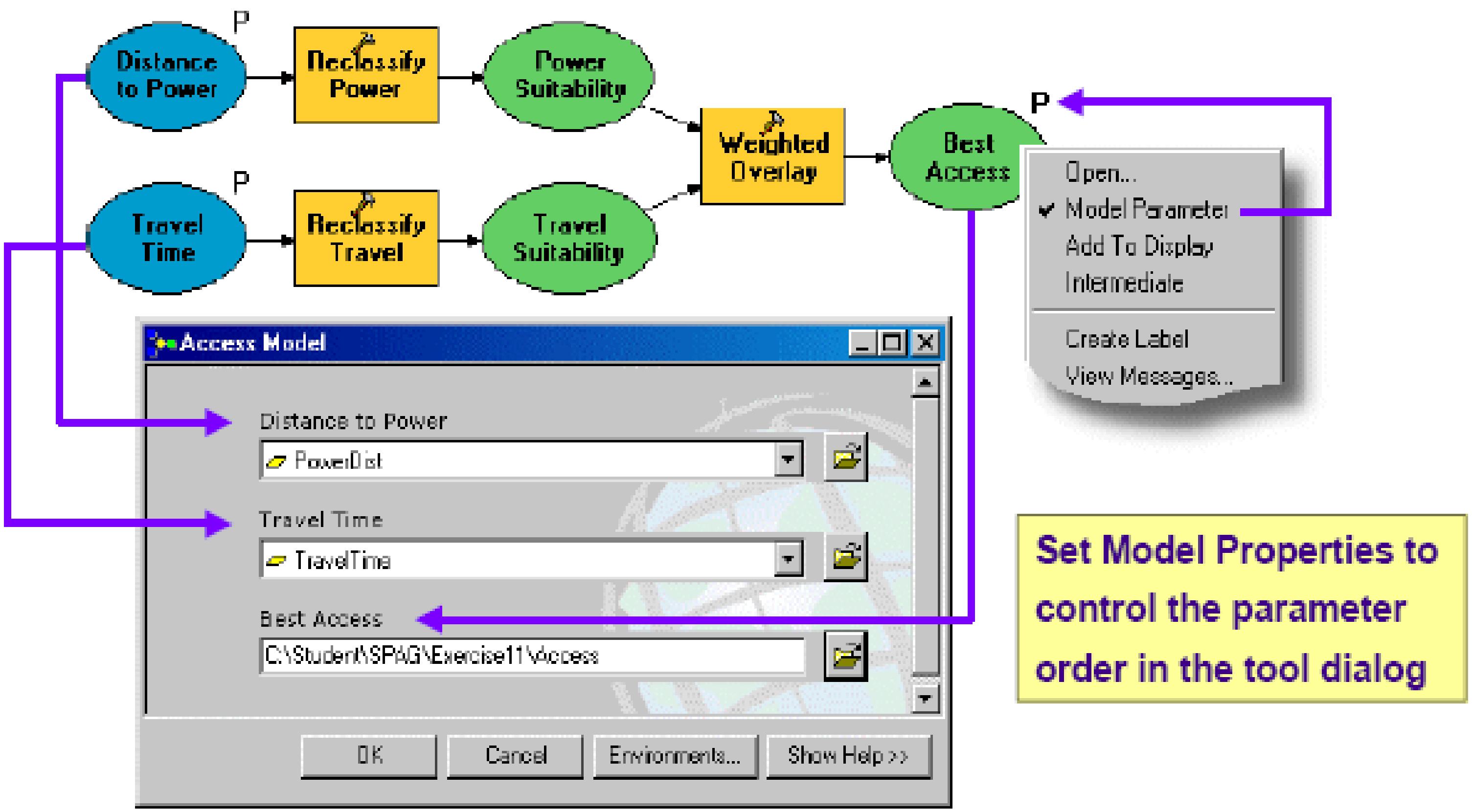
Setting Tool Parameters

- Open the standard **tool dialog:**
  - **Double-click** the tool or **choose Open** from the context menu



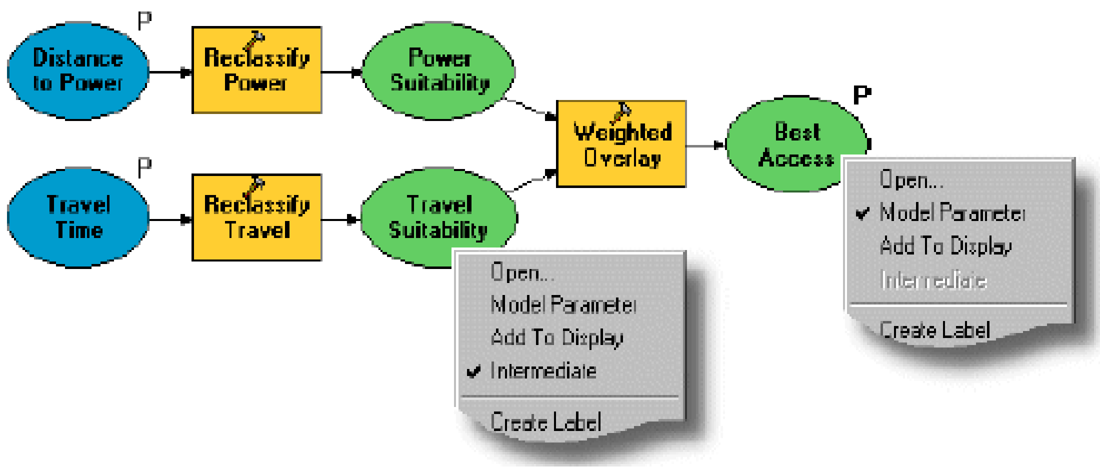
Setting Model Parameters

- **Mark data** as a parameter; **appears** in the model dialog

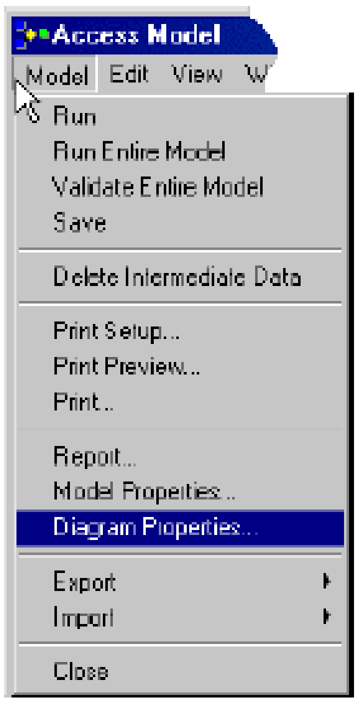


Setting Derived Data Properties

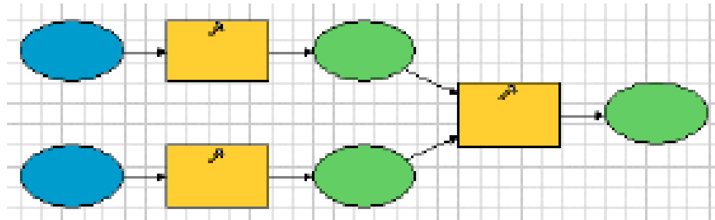
- Controls how derived data is handled:
  - **Intermediate**: Temporary (auto-delete … or not)
  - **Add to Display**: Add to ArcMap Table of Contents
  - **Model Parameter**: Add to ArcMap and permanent

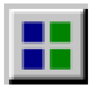


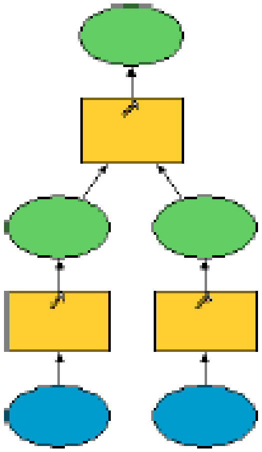
Setting Diagram Properties



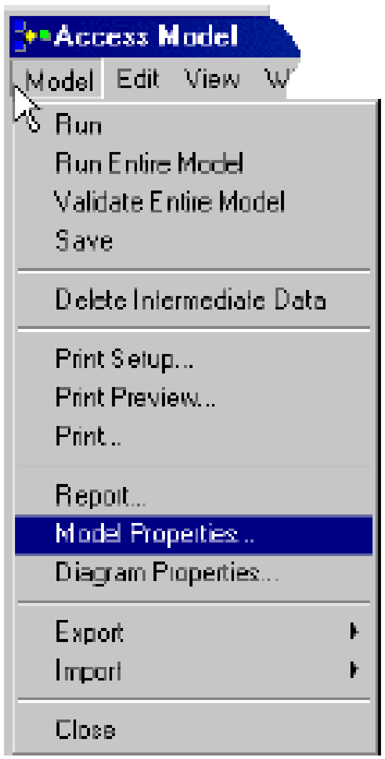
- Set Manual or Automatic layout mode
- Turn snapping grid on or off



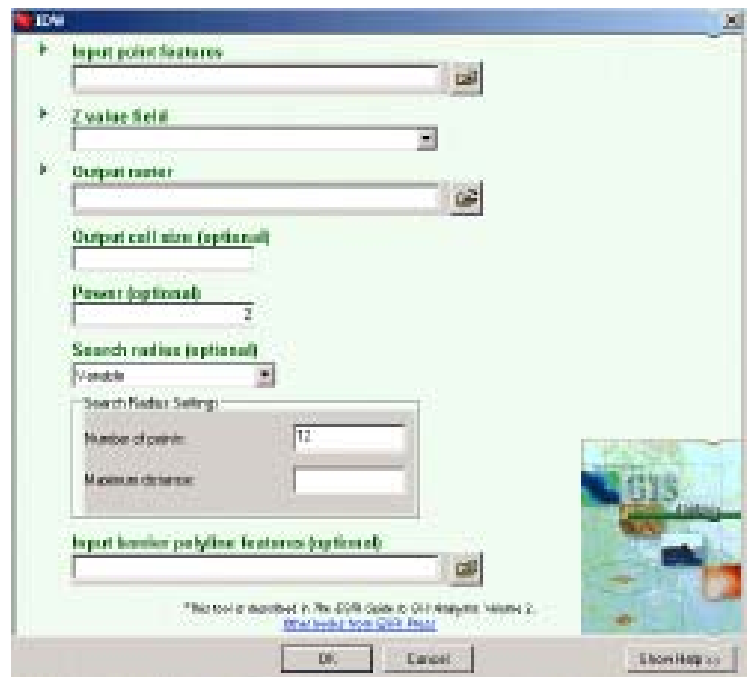
- Control Auto Layout tool settings 
  - Orientation
  - Element spacing
  - Layout quality
  - Connector routing
  - More …



Setting Model Properties

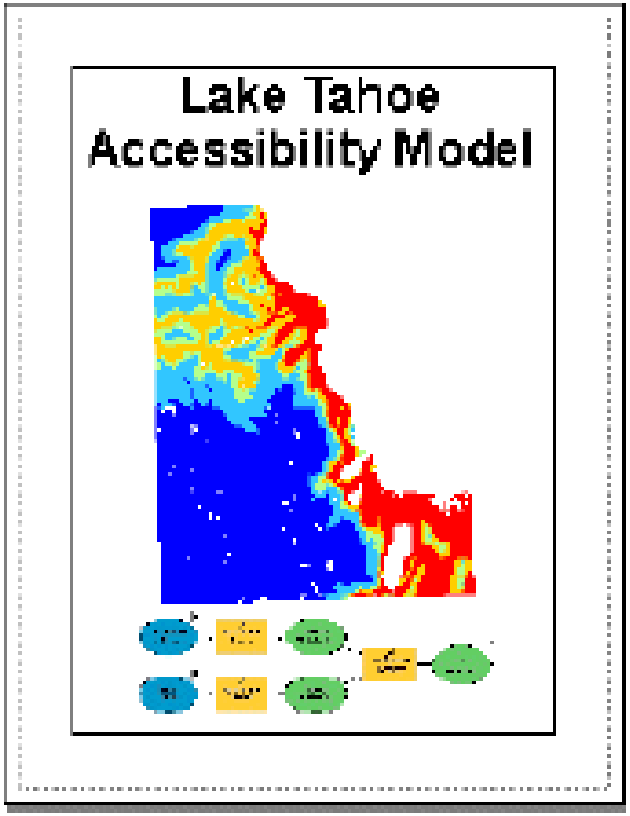


- Set Name, Label, Description
- Set Parameters order
- Set Environments
- (local to model)
- Set Help file
- (HTML)
- Set Stylesheet
- (XSL)



Saving, Exporting, and Printing a Model

- Model is **saved** to a TBX file or a geodatabase
  - **Share** TBX or geodatabase with others
- **Export models**
  - To a graphic: BMP, JPG, EMF
    - (may add to ArcMap layouts)
  - To a script: Python, JScript, VBScript
    - (quick way to learn scripting)
- **Print models**
  - With borders, captions, page numbers
- **Generate reports**
  - List data, tool parameters and so on



Next Topic:

March Vacation!