GIST 7128
ArcGIS 1: Introduction

Lecture 8
**Spatial Analysis**

Module 8 Topics

- **Lecture**
  1. Geoprocessing
     - Proximity Analysis
     - Extracting Data
     - Generalization
     - Overlay
  2. Database Analysis
  3. GIS Problem Solving
  4. Geoprocessing Automation
- **Lab**
  - Chapter **18**. Preparing Data for Analysis
  - Chapter **19**. Geoprocessing Vector Data
- **Project**
  - **Vancouver** – Part **1**: Assemble Map (2%)
  - **Vancouver** – Part **2**: Spatial Analysis (10%)
- **Test**
  - **Final Exam** – all lectures and labs except 5 & 20 (20%)
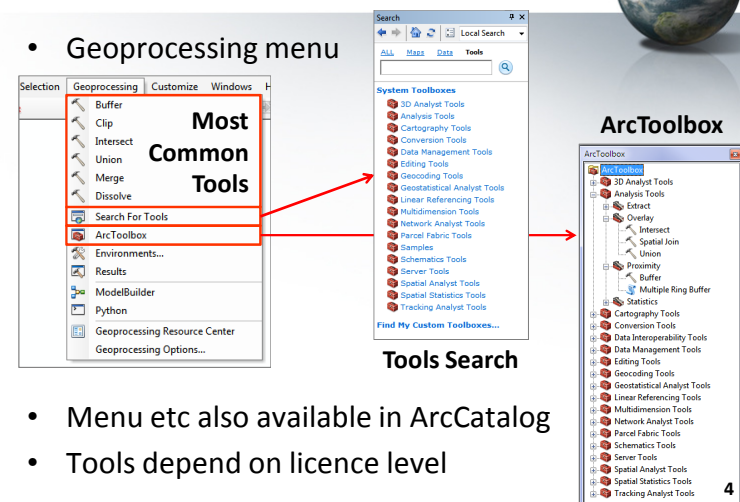
2

---

1. Geoprocessing

- Geoprocessing is the methodical execution of a sequence of computational operations on geographic data to create new data

- Geoprocessing tools can be used for data preparation, data conversion, or spatial analysis

- The fundamental purposes of geoprocessing is to help you complete GIS tasks and to analyze spatial data

Input Data → Geoprocessing Operation(s) → Output Data
Build and edit data ↔ Examine results
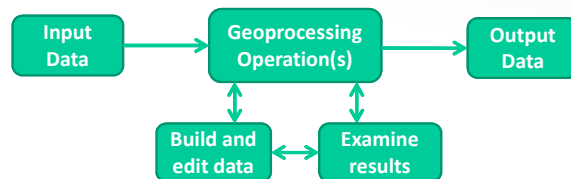
3

Geoprocessing: Access to Tools

- Geoprocessing menu

Most Common Tools

ArcToolbox

Tools Search
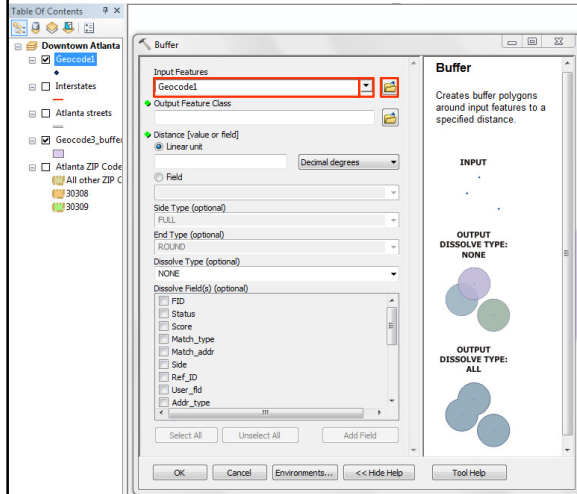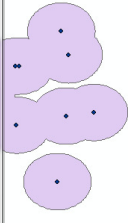
- Menu etc also available in ArcCatalog
- Tools depend on licence level

4

Geoprocessing: Using Tools

1. Start tool
2. Fill in parameters
   a. Enter/drag layer
   b. Menu of current layers
   c. Browse for layer
3. Get Help if needed
4. Click OK to run
5. Check Results

5

Multi-Part Polygons

- Geoprocessing operations on poly-gons can create multi-part polygons
- Several polygons are grouped as one feature
  - can be not adjacent, examples:
  - BC with all islands as one feature
  - State of Hawaii with all islands as one feature

The state of Hawaii is often represented as a multipart feature.

- Must be separate polygons for individual area calculations
  - Use *Multipart to Singlepart* tool (in Data Management)
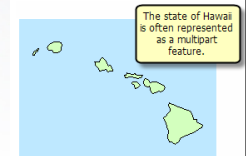
6

1.1 Proximity

Buffer - our first geoprocessing tool

- Buffering is the most commonly used feature-based proximity operation
- A zone of a specified distance around **input** feature(s)
- *e.g.* prohibit logging activities within 100m of lakes
- **Output**: polygons that are outside (or inside) each input feature by the specified buffer distance
  - almost always used as input to other geoprocessing operations
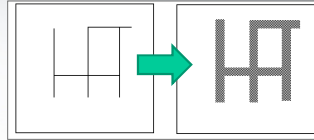
7

Buffer Options

- Geometry
  - Input: point, line, polygon
  - Output: always polygon
- Positive or negative distance (and output)
- Buffer respects *selected* features and *definition query* features
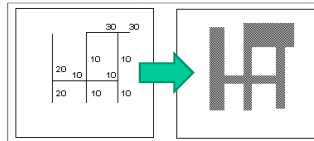- Overlapping buffers can be dissolved or not

Dissolve borders   Do not dissolve borders
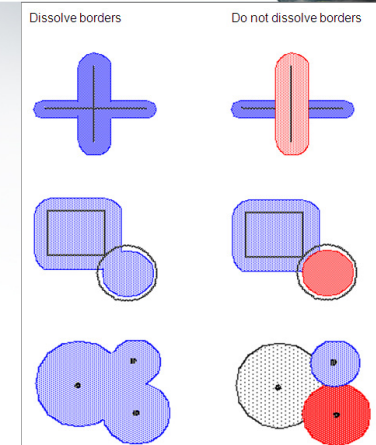
8

Buffer Options

- **Constant**: all features buffered same distance
  - *e.g.* buffer roads by 20 m

- **Variable:** buffered based on numeric attribute value
  - *e.g.* buffer roads based on Class
    - Primary = 30m
   - Secondary = 20m
    - Tertiary = 10m

9

1.2 Extraction

- GIS layers often contain more data than needed
- Need to work with a subset by extracting from larger/ more complex datasets
- Attributes of feature classes involved do not change
- Requires two **input** feature classes
- Geoprocessing tools for extracting data:
  - **Clip:** *keep* features that overlap **clip** features
  - **Erase:** *remove* features that overlap **erase** features
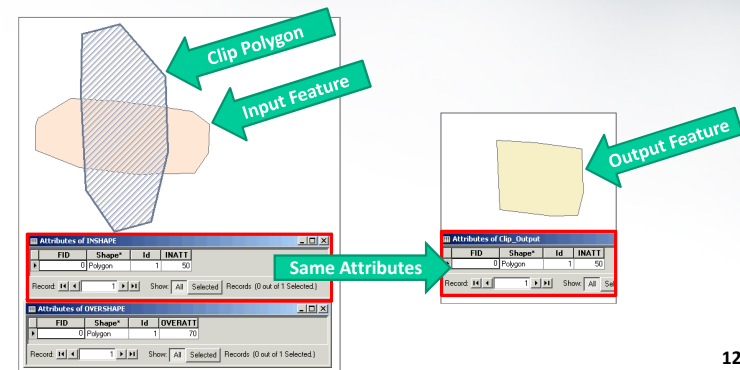
10

Extraction: Clip

- Extracts features or portions of features from **input** layer based on overlay with **clip** features
- **Input** features may be any geometry type
- **Clip** features must be polygon(s)
- **Output** layer: "geographic subset", input feature shape or quantity changes, but not attributes
  - Whole features excluded
  - Split features partially excluded
  - Input feature attributes retained
- *Example:* extracting water mains in specified neighbourhood as a new dataset for further analysis

INPUT   CLIP FEATURE   OUTPUT

INPUT   CLIP FEATURE   OUTPUT

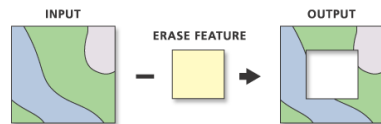INPUT   CLIP FEATURE   OUTPUT
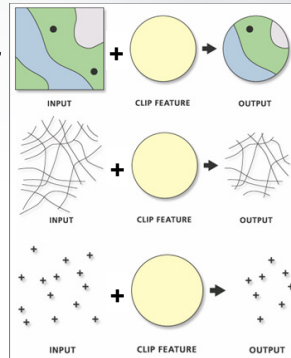
11

Extraction: Clip Example

- **Input** feature: Tree species polygon
- **Clip** feature: Soil type polygon
- **Output** feature: Tree subset (on Soil type) with input attributes

Clip Polygon

Input Feature

Output Feature

Same Attributes

12

Extraction: Erase

- Removes features or portions of features from **input** layer based on overlay with **erase** features
- **Input** features may be <u>any geometry type</u>
- **Erase** features must be <u>polygon(s)</u>
- **Output** layer: "geographic subset", input feature shape or quantity changes, but not attribute structure
  - Whole features excluded
  - Split features partially excluded
  - Input feature attributes retained

13

Extraction: Erase Example

- **Input** feature: Tree species polygon
- **Erase** feature: Soil type polygon
- **Output** feature: Trees (*off* Soil type) with input attributes (x 2)
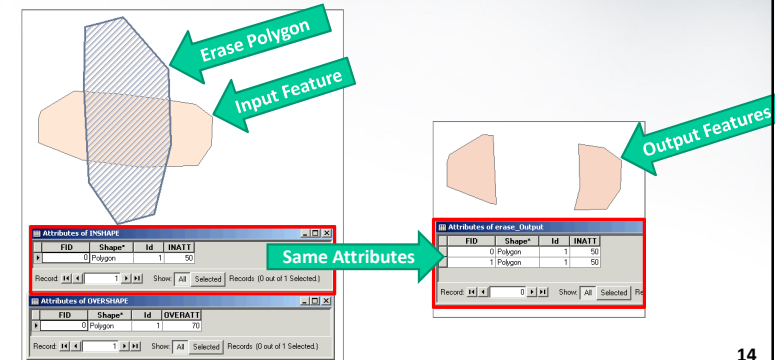
14

1.3 Generalization

- GIS layers often contain more data than needed
- Can simplify complex datasets by removing redundant boundaries and removing unneeded features
- Attributes of feature classes involved do not change
- Requires two **input** feature classes
- Geoprocessing tools for extracting data:
  - **Dissolve:** *remove* redundant boundaries
  - **Eliminate:** *remove* unneeded features

15

Generalization: Dissolve

- Features with same values for specified "Dissolve" fields will be aggregated (dissolved) into single feature
- *Example:* county attributes include sales manager name which can be used to combine counties into "sales regions"
- Dissolve fields included in output layer
  - can also include summaries of any other input attributes
  - e.g. revenue from all member counties summed for each sales region
- May result in multipart polygons
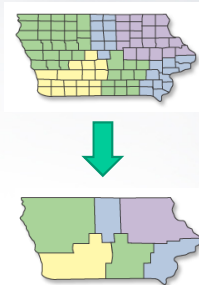
9　　　　　　　　　　　　　　　　　　　　　**16**

Generalization: Eliminate

- Merges polygons with neighboring polygons that have largest area or longest shared border
- often used to remove small "sliver polygons" resulting from overlay operations (Intersect)
- Polygons to be eliminated must be selected first
  - e.g. with Select by Attribute or Location, etc.
  - Eliminate will fail if Input Layer has no selection
- Selected polygons will be merged with adjacent unselected polygon (by removing shared borde)
- A selected polygon will *not* be merged with a neighbouring selected polygon.
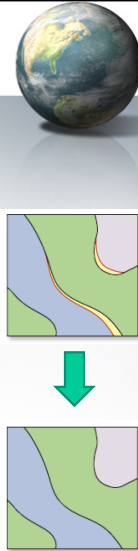
17

1.4 Overlay

- Common question: "What's at the same location?"
- **Input**: two (or more) geographic datasets
- **Overlay**: combine spatial and attribute data from input
- **Output**: spatial and attribute data of <u>all</u> input features
  - New dataset, can be saved
- *Examples*: input → output
  - Point + Polygon → Point (with Point & Polygon attributes)
  - Line + Polygon → Line (with Line & Polygon attributes)
  - Polygon + Polygon → Polygon (with both Polygon attributes)
- **Basic Tools**
  - **Intersect** – *common* area from input layers
  - **Union** – *combined* area of input layers

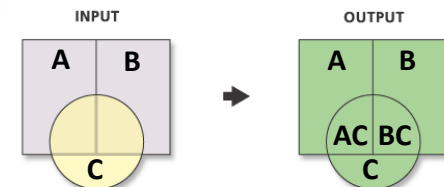18

Overlay: Intersect

- Integrates input layers but only keeps those areas that are <u>common</u> to both
- **Input** layer(s): point, line, or polygon
- **Output** layer: common features (or portions)
  - same as input feature with lowest dimension (or lower)
  - attribute table is a combination of input attributes
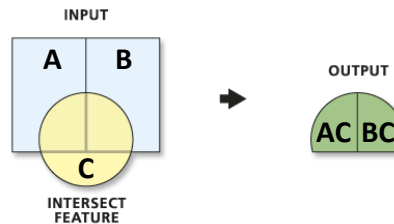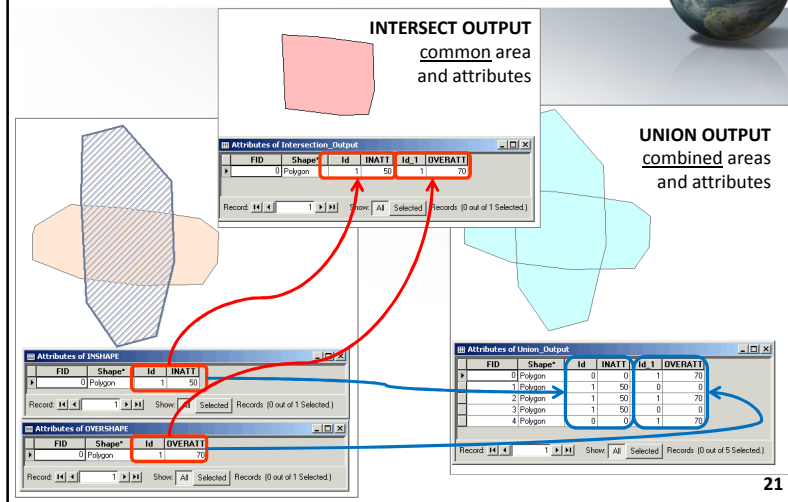

19

Overlay: Union

- Combines all the input features
- **Input** features must have polygon geometry
- **Output** layer: all features superimposed
  - attributes from each of the input features that overlap

20

Overlay: Union vs. Intersect

INTERSECT OUTPUT
common area
and attributes

UNION OUTPUT
combined areas
and attributes

Attributes of Intersection_Output

FID	Shape*	Id	INATT	Id_1	OVERATT
0	Polygon	1	50	1	70

Record: 1 Show: All Selected Records (0 out of 1 Selected.)

Attributes of INSHAPE

FID	Shape*	Id	INATT
0	Polygon	1	50

Record: 1 Show: All Selected Records (0 out of 1 Selected.)

Attributes of OVERSHAPE

FID	Shape*	Id	OVERATT
0	Polygon	1	70

Record: 1 Show: All Selected Records (0 out of 1 Selected.)

Attributes of Union_Output

FID	Shape*	Id	INATT	Id_1	OVERATT
0	Polygon	0	0	1	70
1	Polygon	1	50	1	70
2	Polygon	1	50	1	70
3	Polygon	1	50	0	0
4	Polygon	0	0	1	70

Record: 1 Show: All Selected Records (0 out of 5 Selected.)

21

Overlay: More Tools

- **Intersect** – *common* area of input layers
- **Union** – *combined* area of input layers
- **Symmetrical Difference** – common to either input layer *but not both* input layers
- **Identity** – input features are *split* and attributes *supplemented* by overlapping identity features
- **Update** – input feature geometry and attributes are *replaced* by overlapping update features
- "Degree"
  - Intersect & Union can have any number (2+) of input layers
 - The others must always have exactly 2 input layers

22

Overlay: Symmetrical Difference

- **Output** layer: Features, or portions of features, that exist in only one of the input layers
- This is the complement of the Intersect operation
- The input and difference feature classes must be the same geometry type

INPUT   OUTPUT

23
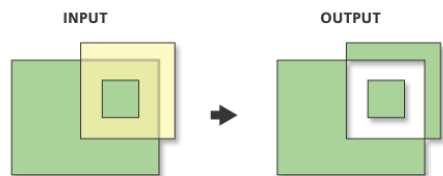
Overlay: Identity

- **Output** layer:
  - input features are split by overlapping identity features
  - supplemented by attributes of overlapping identity features
- **Input** features can be any geometry
- **Identity** features must be polygon

INPUT   OUTPUT

| A | B | | A | B |
| | C | | AC | BC |

IDENTITY FEATURE

24
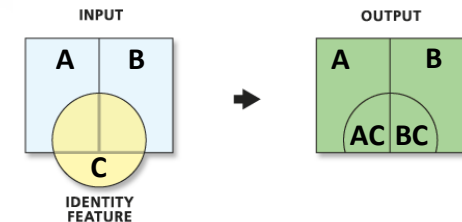
Overlay: Update

- **Output** layer:
  - input features (or portions) are <u>replaced</u> by update features
  - attributes are <u>replaced</u> by update feature attributes
- **Input** and **Update** features must both be polygon
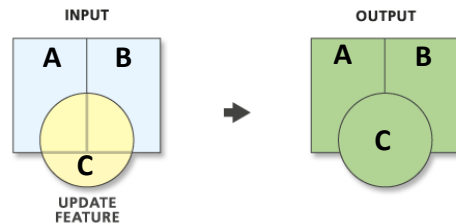

25

2. Database Analysis

- In cases where you may want features that have a particular characteristic
  - *e.g.* Zoning = 'PS School'
- Use Tools covered in Lecture 7
  - Select By Attributes
  - Definition Query
- Many Geoprocessing tools will work on subset of features based on selection or definition query
  - *e.g.* Buffer on School zones only

26

Database Analysis: Calculate Field

- Assign Values to Attributes
- Assign constant value
  - e.g. Year = 2012
- Use expression with existing fields
  - e.g. Density = Pop / Area
- Use Field Calculator
  - right-click on field heading in data table to open
- *Also available (next)*
  - Calculate Geometry
  - Summarize
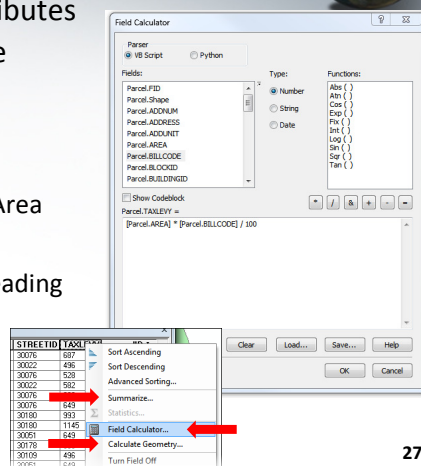
27

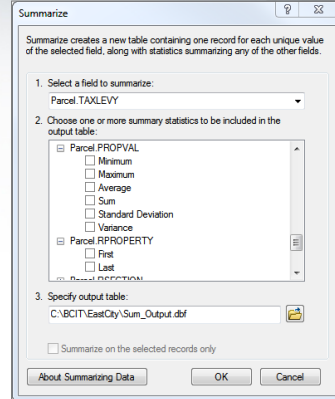## Database Analysis: Calculate Geometry

- Calculate geometry values to load into database fields
- Geometry properties available:
  - Area, Perimeter
  - Centroid X coordinate
  - Centroid Y coordinate
- Uses Coordinate System of layer or data frame
- Units based on spatial CS but can adjust
- Essential for Shapefile revision
  - since geometry fields are not automatically updated after editing
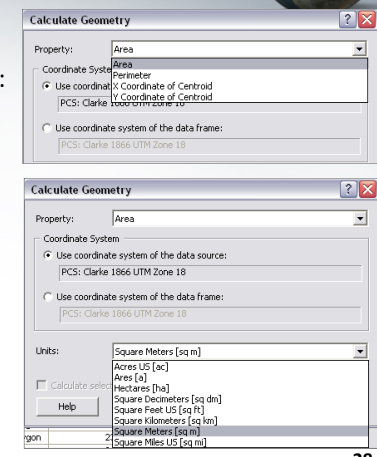
28

Database Analysis: Summarize

- Calculate statistics based on one or more fields
- Statistics types:
  - Numeric
    - Sum, Mean
    - Min, Max
    - Standard Deviation
    - Variance
    - Count
  - Text
    - First, Last
    - Min, Max
    - Count

Summarize creates a new table containing one record for each unique value of the selected field, along with statistics summarizing any of the other fields.

1. Select a field to summarize:
   Parcel.TAXLEVY
2. Choose one or more summary statistics to be included in the output table:
   - Parcel.PROPVAL
     - Minimum
     - Maximum
     - Average
     - Sum
     - Standard Deviation
     - Variance
   - Parcel.RPROPERTY
     - First
     - Last
3. Specify output table:
   C:\BCIT\EastCity\Sum_Output.dbf

   Summarize on the selected records only

   About Summarizing Data   OK   Cancel

29

Database Analysis: Summarize

- Summary written to specified output table
- Can generate a frequency table
  - listing a value and count for each distinct value
  - *e.g.* frequency of street suffixes in a layer

Contents | Preview | Description

OID	FREQUENCY	MEAN_Shape
0	77	75730.9160895

Contents | Preview | Description

OID	FREQUENCY	STTYPE
0	211	
1	8	AL
2	503	AV
3	25	BLVD
4	27	BRDG
5	9	CRES
6	9	CSWY
7	4	CT
8	341	DR
9	7	HWY
10	1	LA
11	1	MEWS
12	10	PL
13	14	RAMP
14	114	RD
15	4	SQ
16	1873	ST
17	14	WALK
18	10	WY

30

3. GIS Problem Solving

- Check criteria and input data, ask questions like…
- *Are multiple tools required for this analysis?* (typically)
- *Is the required data available in existing features?*
  - Select by Attributes *or* Select by Location
  - returns subset of existing features
  - does not modify features or attributes
- *Does data extent or complexity need to be modified?*
  - Does my dataset have too much data?
    - Clip – reduce extent to match study area
  - Are there areas that should be excluded?
    - Erase – remove areas to match study area
  - Do multiple polygons provide same attributes for analysis?
   - Dissolve – simplify polygons

31

## GIS Problem Solving

- *Are new or modified features needed?*
  - *Need info about area surrounding existing features?*
    - Buffer – determine the area impacted
  - *Need spatial & attribute data from multiple layers?*
    - Union – e.g. areas that are *either* brushland *or* wooded
    - Intersection – e.g. *both* wooded *and* suitable soils?
- *Need to update attributes?* → Database Analysis
  - *e.g.* Calculate Field (density), Calculate Geometry (area)
- *Example*: Get total length of streams within 50m of road
  - Isolate portions of streams within 50m of road
    1. Buffer 50m area around road
    2. Clip Streams on Buffer
  - Measure length of isolated streams
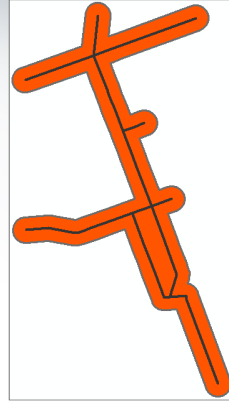   3. Summarize Length of all clipped streams

32

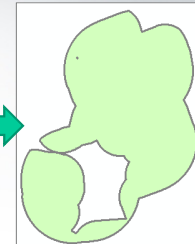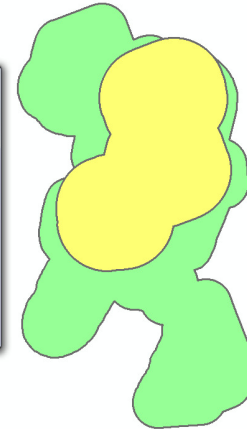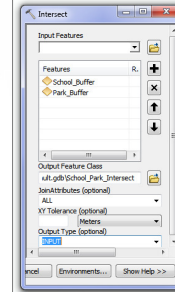# Project: Spatial Analysis – 1. Buffer
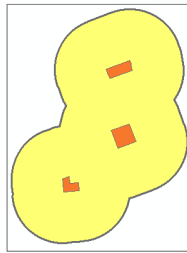
School Buffer
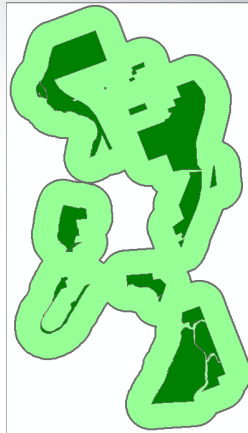
Park Buffer

Busy Street Buffer 33

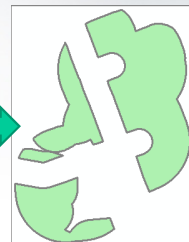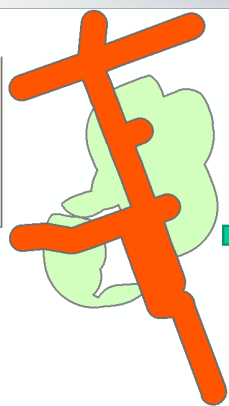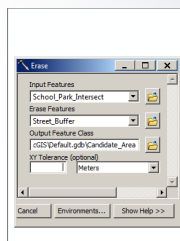# Project: Spatial Analysis – 2. Intersect

School Park Intersect

Intersect School and Park Buffers 34
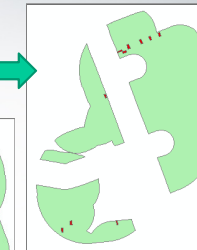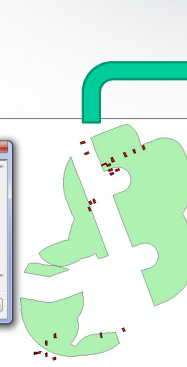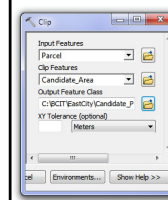
Project: Spatial Analysis – 3. Erase

Candidate Area

Erase Busy Streets from School Park Intersect 35

Project: Spatial Analysis – 4. Clip

OUTPUT:
Candidate Parcels satisfying ALL criteria
- save as shape file
- verify results

Clip Parcels satisfying attribute criteria to Candidate Area 36

4. Geoprocessing Automation

Why Automate?

1. Complex Analysis
  • Reduces tedium and errors
  • Provides documented workflow for reference & refinement

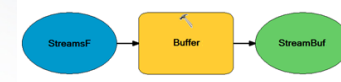**2. "What If?" Scenarios**
  • Adjust parameters and rerun

3. Repeated Processes
  • Temporally: e.g. same process every month or on data revision
  • Geographically: e.g. same process on various regions

37

Model Processes

• Each process has Input, Tool (with parameters), Output

  • e.g. Generate Buffer *(tool)* by 50 meters *(tool parameter)* around Stream Layer *(input)* to produce Stream Buffer *(output)*
• Processes are connected together in a model
  • output of one process becomes the input for next process
  • the processes must be completed in a specific sequence
• Planning complex analysis takes time and effort
  • First run may not work right
 • Takes practice and experience

38

Process Symbols

• Flowchart symbols (shapes) & connectors (arrows)
• Basic symbols for our use (top row):
• Input
  • blue oval
• Tool
  • orange rectangle
• Output
  • green oval
• Connector
  • black/blue arrow
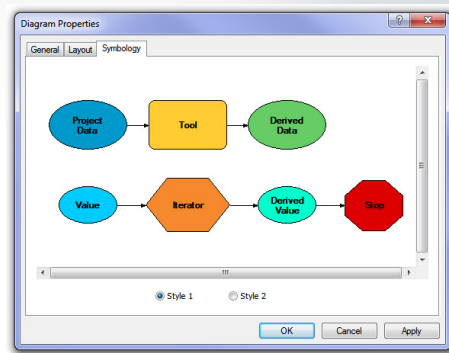
39

Symbol Details

• Symbol shadow *indicates:*
  • Tool was run
  • Output was generated
• Symbol without Fill *indicates:*
  • Component is incomplete
  • E.g. needs input data
• Automatic Connection
  • Drop Input on Tool symbol
  • Select Input on pop-up menu
• Menu & Toolbar
  • Auto Layout
  • Full Extent
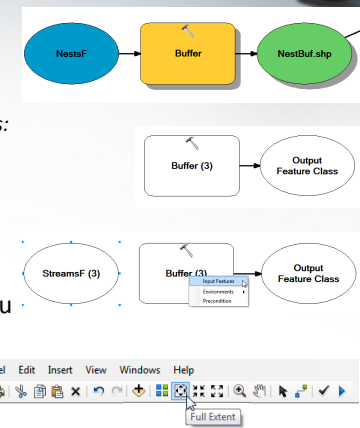

40

ModelBuilder & Models

- **ModelBuilder** is a design environment
  - Graphical interface to interactively build, edit, & view models
  - Bigger than a tool, smaller than a stand-alone application
- **Model** is a workflow diagram
  - Interconnected set of processes to be run together
  - Saved in a toolbox in a folder, a geodatabase, or in ArcToolbox
  - Can be shared with colleagues or transferred to other systems
  - Documented to describe process in detail
  - Run the entire model with a single click
- **Modeling** is the process of creating & using models, with:
  - ArcToolbox tools
  - Other models (one model can be input to another)
  - Custom Python scripts

41

Example Model

- Harvest Profitability Analysis
  - Forestry planning for tree harvesting
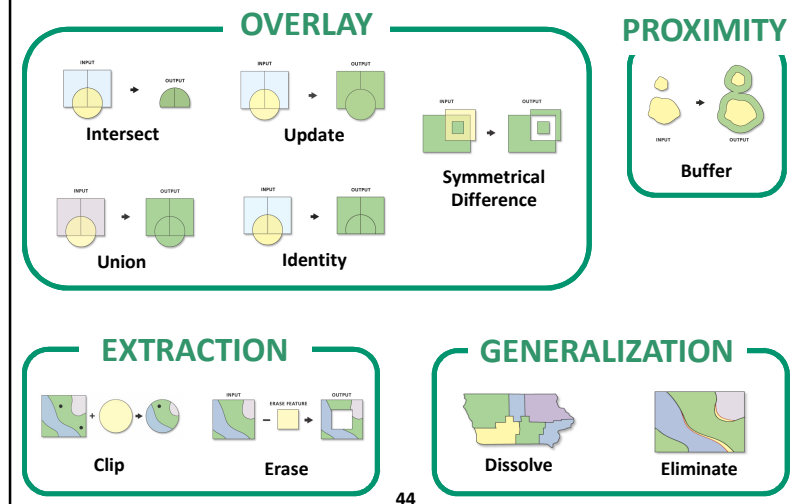 - Must avoid Hawk Nests and Streams
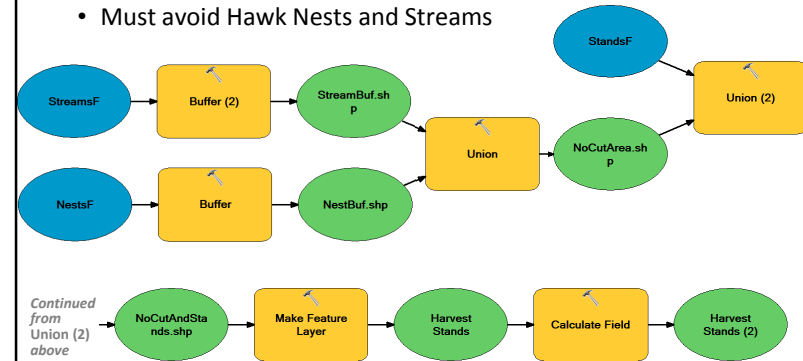
42

Using ModelBuilder

1. Create new toolbox in ArcCatalog (or use existing)
   - Right-click on ArcCatalog folder, select New > Toolbox
2. Add to ArcToolbox (optional)
   - Right-click in ArcToolbox window, select Add Toolbox…
3. Create model inside toolbox
   - Right-click on new toolbox, select New > Model…
4. Edit model to open model builder
   - Right-click on model, select Edit…
5. Drag & Drop components into the model
   - Tools from ArcToolbox, Data from ArcMap ToC or ArcCatalog
   - Set parameters by 'double-clicking' symbols
   - Add connections between data and tools
6. Save, Validate, and Run

43

Review: Geoprocessing Tools

44