Data Development

Using ArcGIS-based Analysis to Recalculate Data to Suit your Geography
• Data-related limiting factors
• Using GIS to analyze and recalculate data
• Data creation/recalculation workflow
• Examine some examples of data processing applications
• Data processing using GIS in support of decision-making concerning environmental issues related to noise & vibration and air quality
  • Example One: Residential dwelling count calculations within specific spatial geometries
  • Example Two: Calculated estimates of expected emissions within an air quality modelling grid (also known as Spatial Surrogates)
Data-Related Limiting Factors

• Data availability
  – “Best” sources of GIS data are government bodies – municipal, provincial/state, federal
  – Various governmental departments and agencies now employ and store information in GIS formats
  – Despite movement towards Open Data, access to data can still be quite limited
• Data suitability
  – Data suitability should be evaluated early on in a project / application
  – measures should be taken to correct data where possible
    • Is the data current?
    • Is the source reputable?
    • Does it cover my area of interest?
    • Are there better sources of data available?
Data-Related Limiting Factors

• Data limitations
  – Level of detail, spatial extent, spatial and temporal resolution, etc.
  – Understanding the limitations of your data will help to guide steps to data processing
    • Are the limitations going to effect end result?
    • If so, can the data be processed to help mitigate the limitations?
    • Is the effort of processing the data going to make a worthwhile impact in the end result?
• GIS software, such as ArcGIS Desktop offers many tools to aid spatial data manipulation
• Offers a visual interface so the user can see and check the data, and view changes made in each step as it is processed
• ArcGIS Tools
  – Large selection of out-of-the-box tools
• Projections, Transformations
  – Customizable, re-projection on the fly
• Basemaps
  – Adds quick context layers for cross-checks
• Model Builder, Python tools
  – Automation of repetitive tasks, build customized tools and workflows
Workflow

Input: Raw Sourced Data

Data Process Workflow

Output: Final Data Product
Workflow

Input: Raw Sourced Data

Data Process Workflow

Output: Final Data Product

Input: Raw Sourced Data

Data Process Workflow

Output: Final Data Product
Model Builder Screen Shot

- Step4_Points.Join_to_27km_MCIP
- Step3_ClipRaster_to_Points
- Step2_Raster_Clip_to_81km_WRIF
- Step1_NetCDF_to_Raster
Model Builder Screen Shot
Let's Look at Some Examples

• Data analysis and recalculation that were performed using GIS-based tools workflows

• Input data were manipulated to suit particular application, solve a problem, or fill a need
Example 1: Dwelling Count Calculations

- Noise and vibration assessment as part of environmental assessment of construction activity
- Involved quantifying potentially affected residential dwellings
- 600m, 5km and 10km buffer areas of the pipeline project, each broken into several segment areas
Example 1: Dwelling Count Calculations

- Due to size and timeline of project heads-up digitization was not an option
- Leveraged dwelling counts in the 2011 Census data from Statistics Canada, along with federal land use datasets
- Allowed us to isolate and modify census boundaries down to the areas that would have dwellings (eg: urban areas)
- Estimates of dwellings were then calculated based on study area vs. modified census area geometries
Example 2: Spatial Surrogates for Emissions Inventory Modelling

• Regional airshed modelling project for Region of Peel, located in southwestern Ontario, Canada

• Purpose of project as a whole is to assist in evaluating public policy decisions and their potential effect on regional and local air quality and human health
Example 2: Spatial Surrogates for Emissions Inventory Modelling

- GIS was used to calculate spatial surrogates
- Surrogates are used to spatially allocate provincially-aggregated emissions totals across a model area
- Surrogates for multiple emissions sources were generated for 4 different modelling grids – 36km, 12km, 4km and 1km
- Some surrogates were found to be “problematic” – these were improved during Year 2 of the project using GIS to create and process new datasets
MCIP 36km Model Domain & Cells
• In Year 1, a traditional surrogate method and dataset was used to generate surrogates for commercial marine vessel emissions
• The result for the 1km surrogate was problematic due to the coarseness of the input dataset
• To correct this, a different dataset with a refined geometry of vessel routes was sourced and modified in the 1km model area
• Modifications to the dataset were based on information collected from nautical charts and satellite imagery
SSC 945 – MARINE – Commercial Vessels

Year 1 Surrogate

Year 2 Surrogate

Legend

MARINE - Commercial Vessels

1 km Model Domain

1 km Model Cells

Peel Region

Expected % of Province-Wide Emissions
0.000001 - 0.000065
0.000066 - 0.000240
0.000241 - 0.000676
0.000677 - 0.001784
0.001785 - 0.002912
Airport Surrogate Improvement

- Geometry for a surrogate related to airport Ground Support Equipment (GSE) was digitized using imagery and airport diagrams to isolate areas where GSE activity on airport properties.
- Geometry was added in areas where aircraft loading/unloading occurs - in the vicinity of passenger terminal buildings and air cargo buildings.
- Additionally, the number of movements by airport (2011 data) was added into the GIS data attributes and used in the surrogate calculation.
- Resultant surrogate concentrated the GSE emissions to airport locations, with greatest emissions at Pearson International airport.
Example of geometry creation at YYZ (Pearson International) airport
• Surrogates for commercial aircraft were also traditionally based on census data as input data

• Changing the input dataset to a dataset that was a multi-ring buffer of the airports with LTO data as attributes improved the resultant surrogate dramatically, concentrating the aircraft emissions around airports
SSC 903 – AIRPORTS – Commercial Aircraft

Year 1 Surrogate

Year 2 Surrogate

Legend

AIRPORT - Commercial Aircraft

1 km Model Domain
1 km Model Cells

Expected % of Province-Wide Emissions

0.000000 - 0.000160
0.000161 - 0.001156
0.001157 - 0.003056
0.003057 - 0.005603
0.005604 - 0.008872

Service Layer Credits: National Geographic, Esri, DeLorme, HERE, UNEP-WCMC, USGS, NASA, ESA, METI, NRCAN, GEBCO, NOAA, increment P Corp.
• In Year 1, a traditional surrogate method and census dataset was used to generate surrogates for emissions related to rail yards
• Newly released federal rail network dataset included classifications of rail line geometries
• Classifications allowed different rail geometries to be isolated from one another
• Drastic improvements were realized as a result, especially to the rail yards surrogate
• GIS offers many tools that help to assess the suitability and limitations of data for use in different applications
• Tools and models can also be used to manipulate existing data and automate data processing workflows
• Data that has been processed with the particular application in mind can drastically improve the quality of results
• *Better quality input = better quality output*
• Better quality, more informed decisions can then be made on scientific model results
Thank You!

Carol McClellan, GISP
RWDI
Email: carol.mcclellan@rwdi.com
Visit us at: www.rwdi.com

Acknowledgements:
Louise Aubin, Kim McAdam
Regional Municipality of Peel Public Health Department

Michael D. Moran, Junhua Zhang, Qiong Zheng
Environment Canada