Note on Applications of Geographical Information in Air Quality Modeling

Byeong-Uk Kim, Ph.D.
Georgia Environmental Protection Division

8th Annual CMAS Conference
Chapel Hill, NC
October 19-21, 2009

Objective

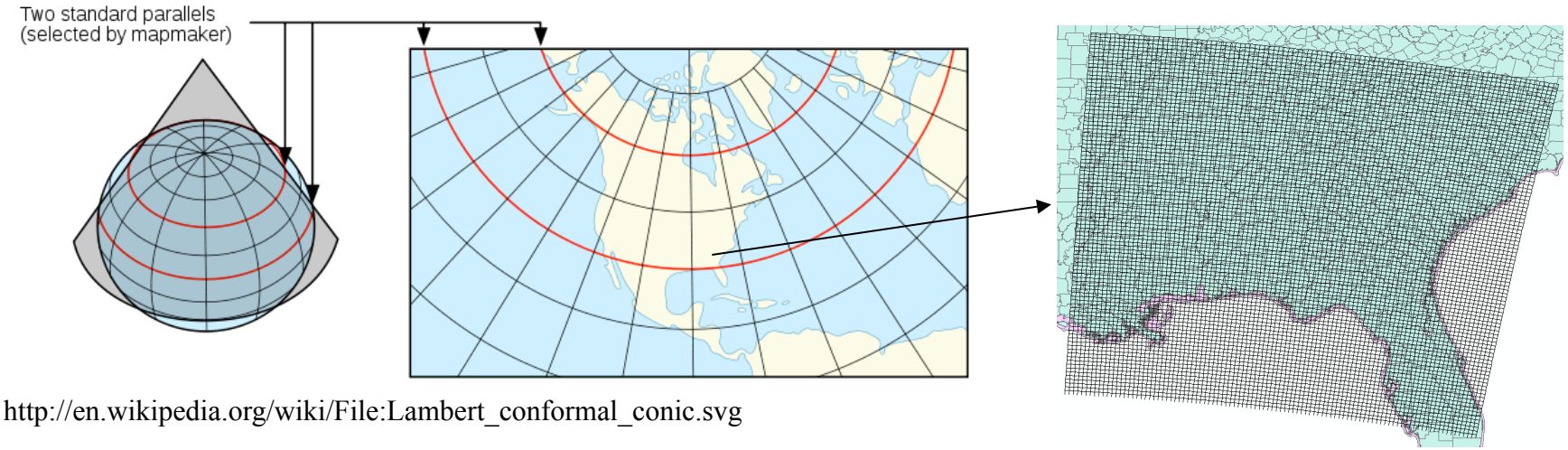
- Investigate the effects of using different models of the earth (i.e. geodetic datum) for air quality modeling
- Propose the “best practice” to utilize air quality modeling related information generated by different models of earth

Approach

- Review the most commonly-used geodetic datum (hereafter, datum) in air quality modeling exercise
- Examine how choice of datum may affect air quality modeling related information
- Develop an approach to minimize the impact of information distortion due to differences in datum on air quality modeling exercise and its applications

Projected Coordinate Systems: Lambert Conformal Conic (LCC)

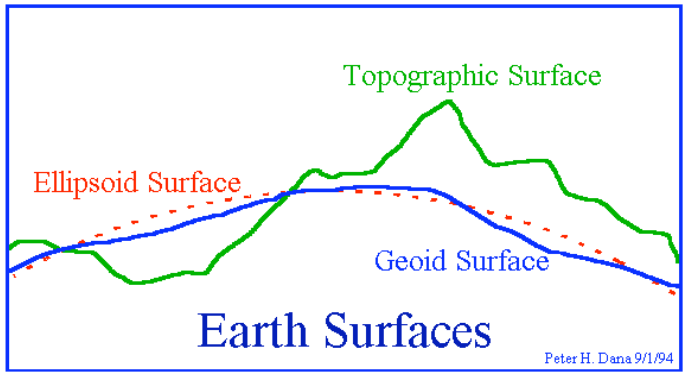
- Lambert Conformal Conic (LCC) with a sphere is one of the most frequently used projected coordinate systems (PCS) for regional-scale photochemical air quality modeling in US.
- A PCS is defined on a flat, two-dimensional surface and built on a geographical coordinate system (GCS).



http://en.wikipedia.org/wiki/File:Lambert_conformal_conic.svg

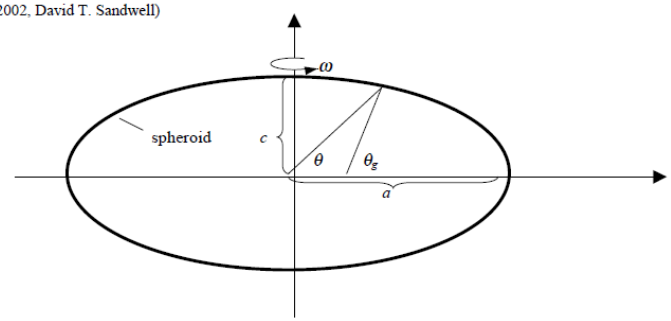
Geographical Coordinate Systems: WGS84, NAD83, NAD27, and Spheres

- All of these datums are Geographical Coordinate System (GCS) frequently used throughout air quality modeling processes.
- Geographical Coordinate System (GCS) refers a location on the earth with three coordinates (longitude, latitude, and altitude) with respect to its datum.
  - Datum defines a reference point on a model of the earth (i.e. a spheroid or an ellipsoid).
  - A spheroid or an ellipsoid defines a shape of the earth and is often derived from satellite observation (e.g. GRS80 ellipsoid for WGS84 and NAD83).
- Often, a GCS uses only one datum that is interchangeable with ellipsoid or spheroid.



http://www.colorado.edu/geography/gcraft/notes/datum/gif/surfaces.gif

Reference Earth Model - WGS84
(Copyright 2002, David T. Sandwell)



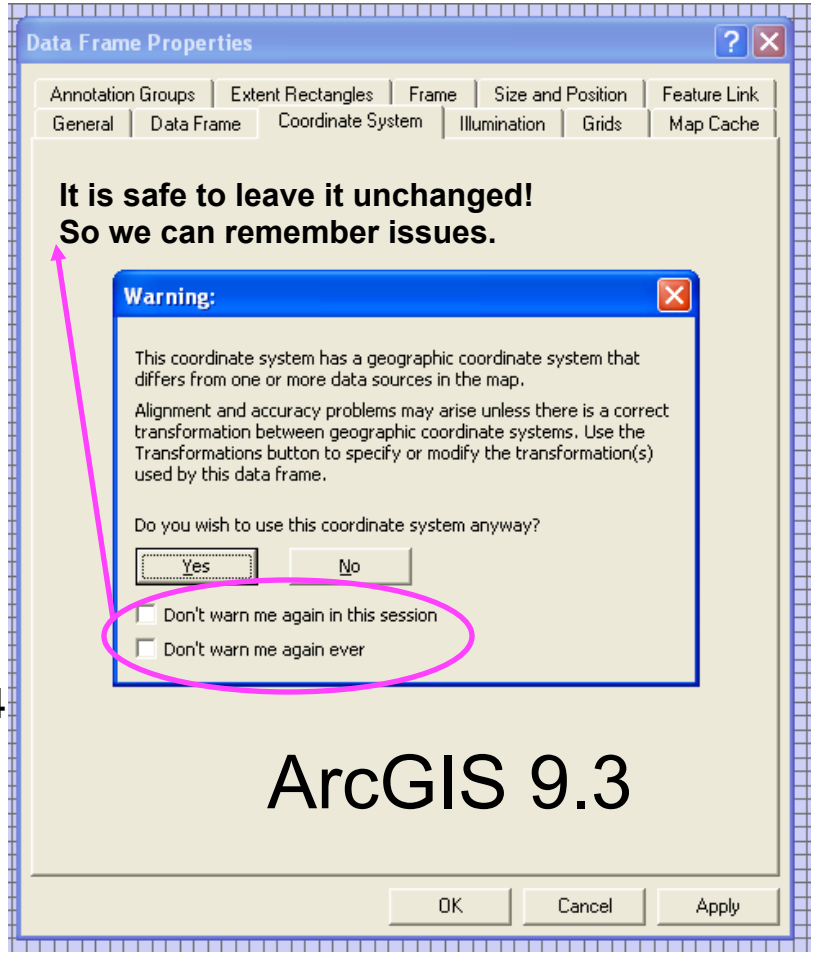
Datum	Ellipsoid	Semimajor axis (a)	Semiminor axis (c)
NAD27	Clarke 1866	6,378,206.4 m	6,356,583.8 m
NAD83	GRS 80	6,378,137 m	6,356,752.3141 m

Use of WGS84, NAD83, NAD27, and Spheres

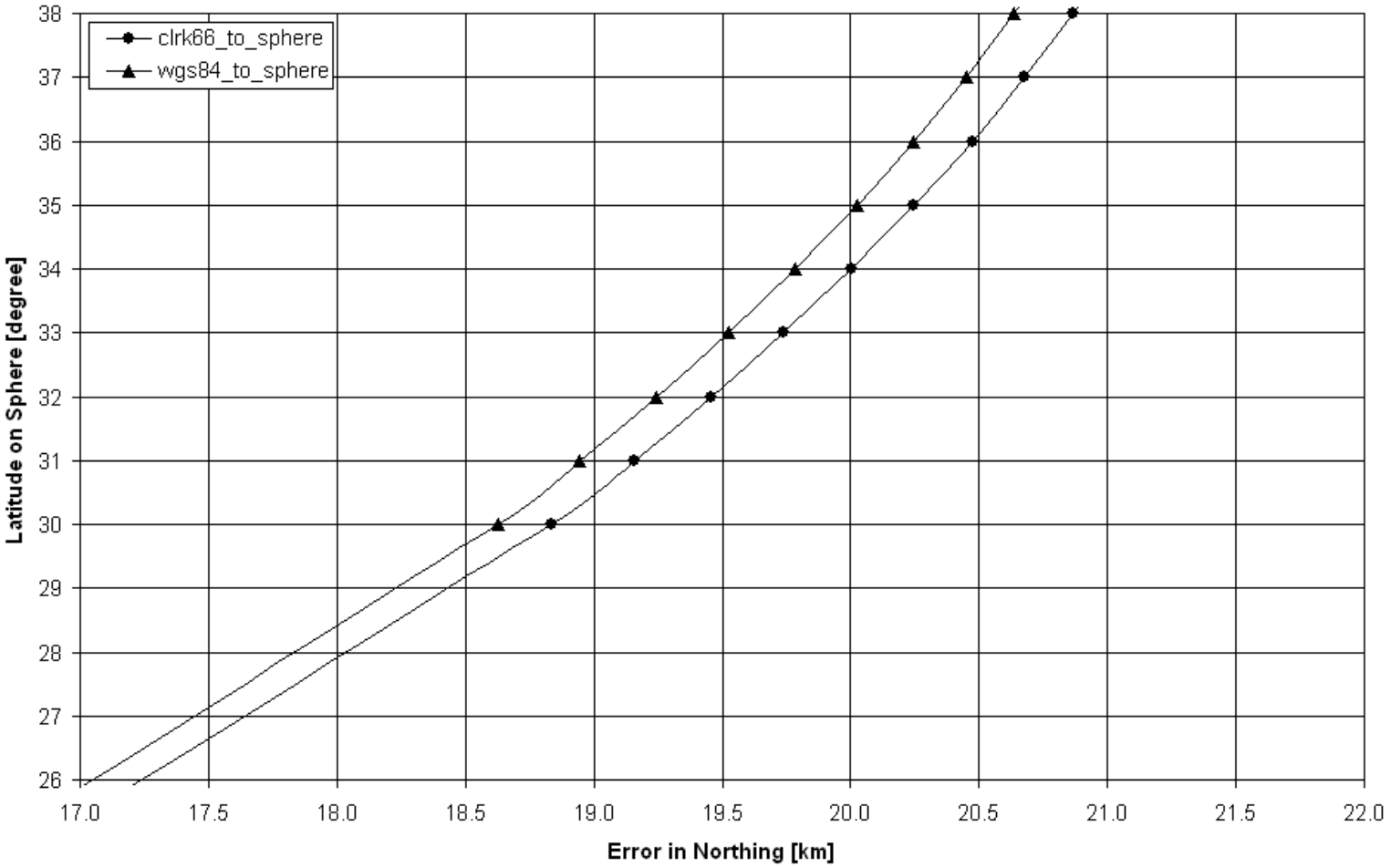
- WGS84
  - Ellipsoid: GRS80
  - Most of GPS application
  - Google Earth
  - USGS Global Land Survey Products
    - meteorological model inputs
- NAD83
  - Ellipsoid: GRS80
  - Recent US Census Data, e.g. 2000 Census Tracts
    - emission related information
- NAD27
  - Ellipsoid: Clarke1866
  - Old US Census Data, e.g. 1990 Census Tracts
- Sphere
  - Meteorological Models and Air Quality Models
    - CMAQ : R=6370.997 km
    - WRF/ARW, CAMx: R=6370.000 km
    - NCEP: R=6371.200 km

Datum Transformation

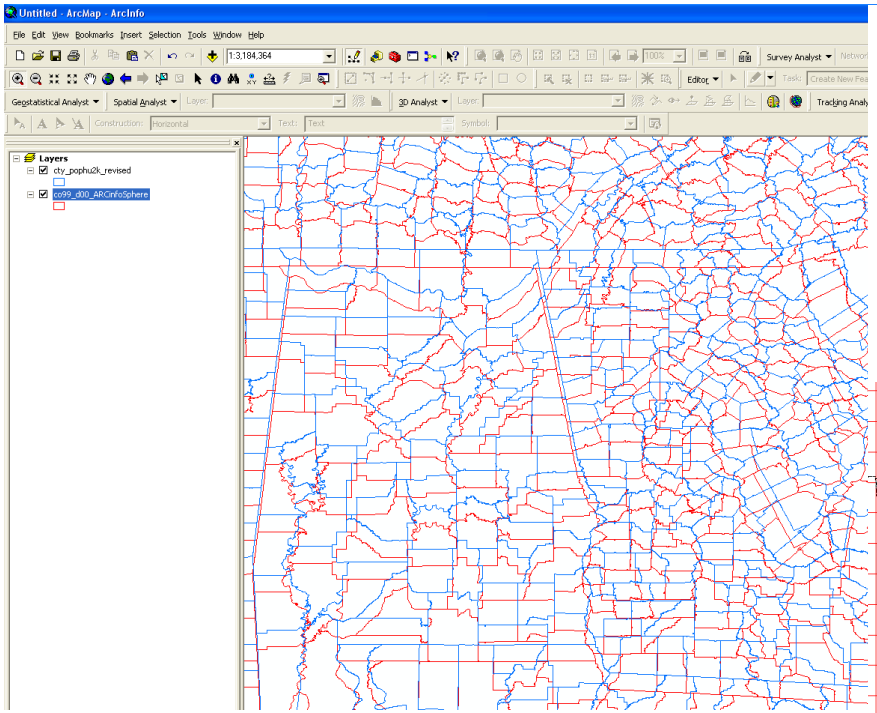
- In general, data sets based on different datums (or GCS) should undergo GCS transformation (i.e. datum transformation) even though their PCS parameters are identical except for PCS reference datum.
- Often datum transformation is not straightforward nor a single step process.
  - For example, it is essentially two-step process to convert data sets from LCC/NAD83 to LCC/WGS84 although both of NAD83 and WGS84 are using GRS80.
    - LCC/NAD83->NAD83->WGS84->LCC/WGS84



Difference in Positional Error by Datum

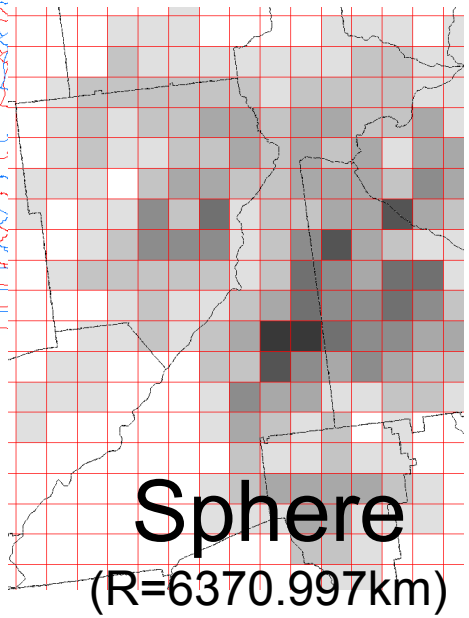


Example of Incorrect Datum Issue: Emission Processing

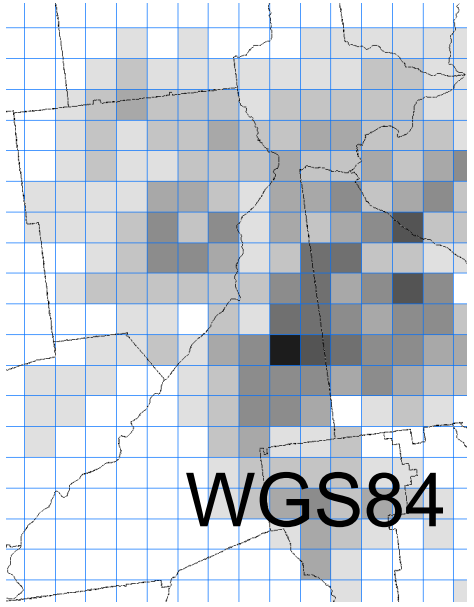


EPA's document for the county boundary file in NEI FTP site reports it is on sphere (R=6370.997 km) while the base file from US Census used to create the EPA file is based on WGS84.

The issue is, however, that proper datum transformation (i.e. from WGS84 to sphere, vice versa) will lead to misalignment between two datasets. It indicates some discrepancies in meta-data.



Sphere
(R=6370.997km)



WGS84

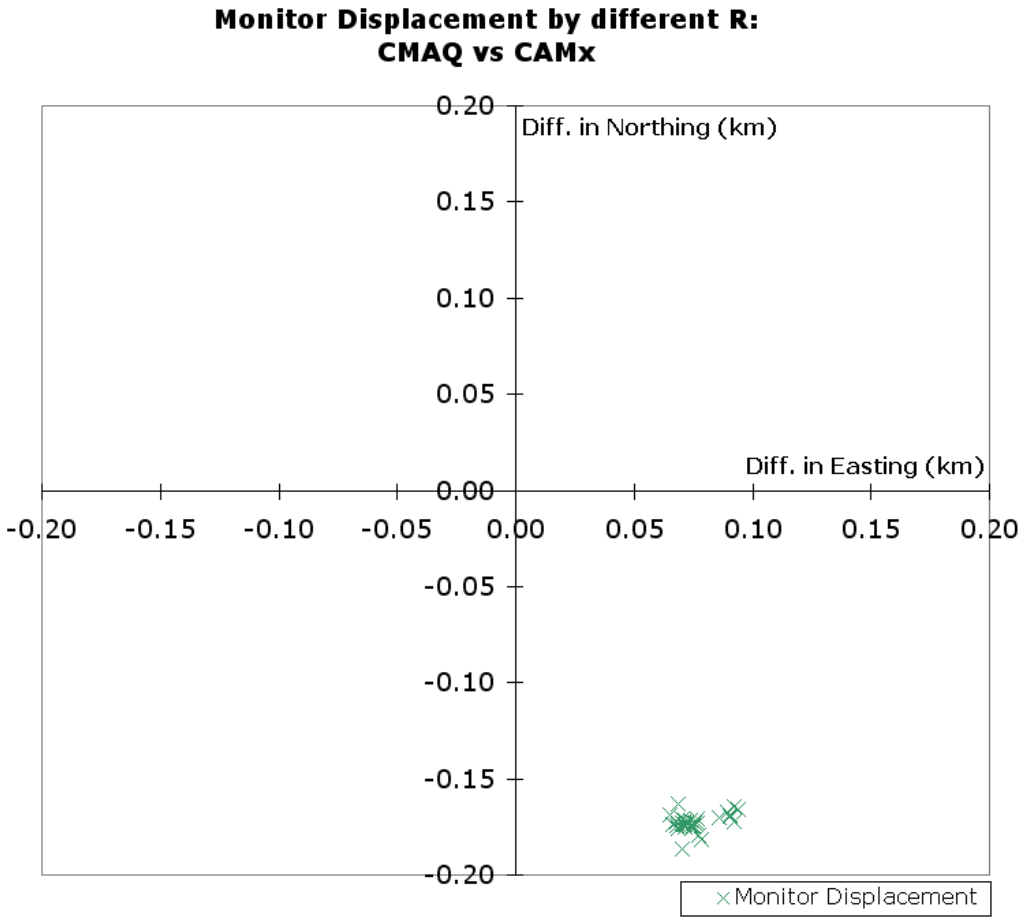
It also means using EPA files with US Census files will create incorrect gridded surrogate data for emission modeling (See left figures).

Example of Mixed Datums Issue: Monitor locations Displacement

R_CMAQ = 6370.997 km
R_CAMx = 6370.000 km

0.997 km difference in earth radius results in ~0.2 km different in this particular PCS.

Sometimes, 0.2 km is big enough to place monitors in different modeling grid cells!



Air Quality Models and Map Scaling Factors

•CMAQ and CAMx solve PDEs with map scaling factor adjustment.

$$\frac{\partial \hat{\mathbf{V}}_s}{\partial t} + (\hat{\mathbf{V}}_s \bullet \nabla_s)\hat{\mathbf{V}}_s + \hat{v}^3 \frac{\partial \hat{\mathbf{V}}_s}{\partial s} + f\hat{\tau}_3 \times \hat{\mathbf{V}}_s$$

$$+m^2\left(\frac{1}{\rho}\nabla_s p + \nabla_s \Phi\right) - m^2\left(\frac{\partial s}{\partial z}\right)\left(\frac{1}{\rho}\frac{\partial p}{\partial s} + \frac{\partial \Phi}{\partial s}\right)\nabla_s h = \hat{\mathbf{F}}_s$$

$$\frac{\partial \hat{v}^3}{\partial t} + (\hat{\mathbf{V}}_s \bullet \nabla_s)\hat{v}^3 + \hat{v}^3 \frac{\partial \hat{v}^3}{\partial s} + \hat{v}^k \hat{v}^j \frac{\partial^2 h}{\partial \hat{x}^k \partial \hat{x}^j}\left(\frac{\partial s}{\partial z}\right)$$

$$-m^2 \frac{\partial s}{\partial z}\left(\frac{1}{\rho}\nabla_s p + \nabla_s \Phi\right) \bullet \nabla_s h + \left|\nabla_s s\right|^2 \left(\frac{1}{\rho}\frac{\partial p}{\partial s} + \frac{\partial \Phi}{\partial s}\right) = \hat{F}_3$$

Map scaling factor

$$\left.\frac{\partial c_l}{\partial t}\right|_{Emission} = m^2 \frac{E_l}{\partial x \partial y \partial z}$$

$$\left.\frac{\partial c_l}{\partial t}\right|_{X\ advection} = -\frac{m^2}{A_{yz}}\frac{\partial}{\partial x}\left(\frac{uA_{yz}c_l}{m}\right)$$

$$\left.\frac{\partial c_l}{\partial t}\right|_{Y\ advection} = -\frac{m^2}{A_{xz}}\frac{\partial}{\partial y}\left(\frac{vA_{xz}c_l}{m}\right)$$

$$\left.\frac{\partial c_l}{\partial t}\right|_{Z\ transport} = \frac{\partial(c_l \eta)}{\partial z} - c_l \frac{\partial}{\partial z}\left(\frac{\partial h}{\partial t}\right)$$

$$\left.\frac{\partial c_l}{\partial t}\right|_{Z\ diffusion} = \frac{\partial}{\partial z}\left[\rho K_v \frac{\partial (c_l/\rho)}{\partial z}\right]$$

$$\left.\frac{\partial c_l}{\partial t}\right|_{XY\ diffusion} = m\left\{\frac{\partial}{\partial x}\left[m\rho K_X \frac{\partial (c_l/\rho)}{\partial x}\right] + \frac{\partial}{\partial y}\left[m\rho K_Y \frac{\partial (c_l/\rho)}{\partial y}\right]\right\}$$

$$\left.\frac{\partial c_l}{\partial t}\right|_{Wet\ Scavenging} = -\Lambda_l c_l$$

$$\left.\frac{\partial c_l}{\partial t}\right|_{Chemistry} = Mechanism - specific\ R eaction\ Equations$$

CMAQ

Specified/Generated by RDMM5V3.F or SETUP_WRFEM.F in MCIP runs

- MSFX2 in GRID_CRO_2D
- MSFD2 in GRID_DOT_2D

CAMx

Calculated on-the-fly by in-line calculation code, GRDGREP.F, using LCPGEO.F (based on MM5's TERRAIN preprocessor) by assuming one degree in N-S direction is equal to 111.1338 km.

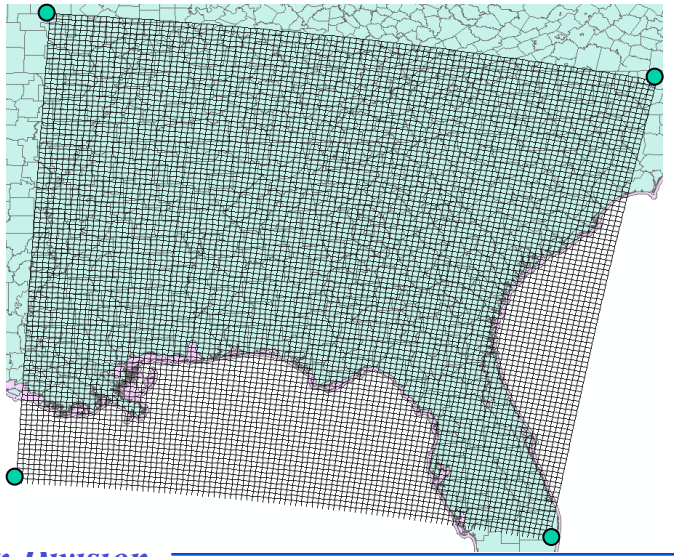
Air Quality Models and Map Scaling Factors

- Therefore, map scaling factor differences by various datums with same projection parameters can be good information to examine the impact of datum choices on air quality modeling simulations *indirectly*.
  - Meteorological models can be more sensitive to small perturbations than air quality models, especially for very small scale flows.

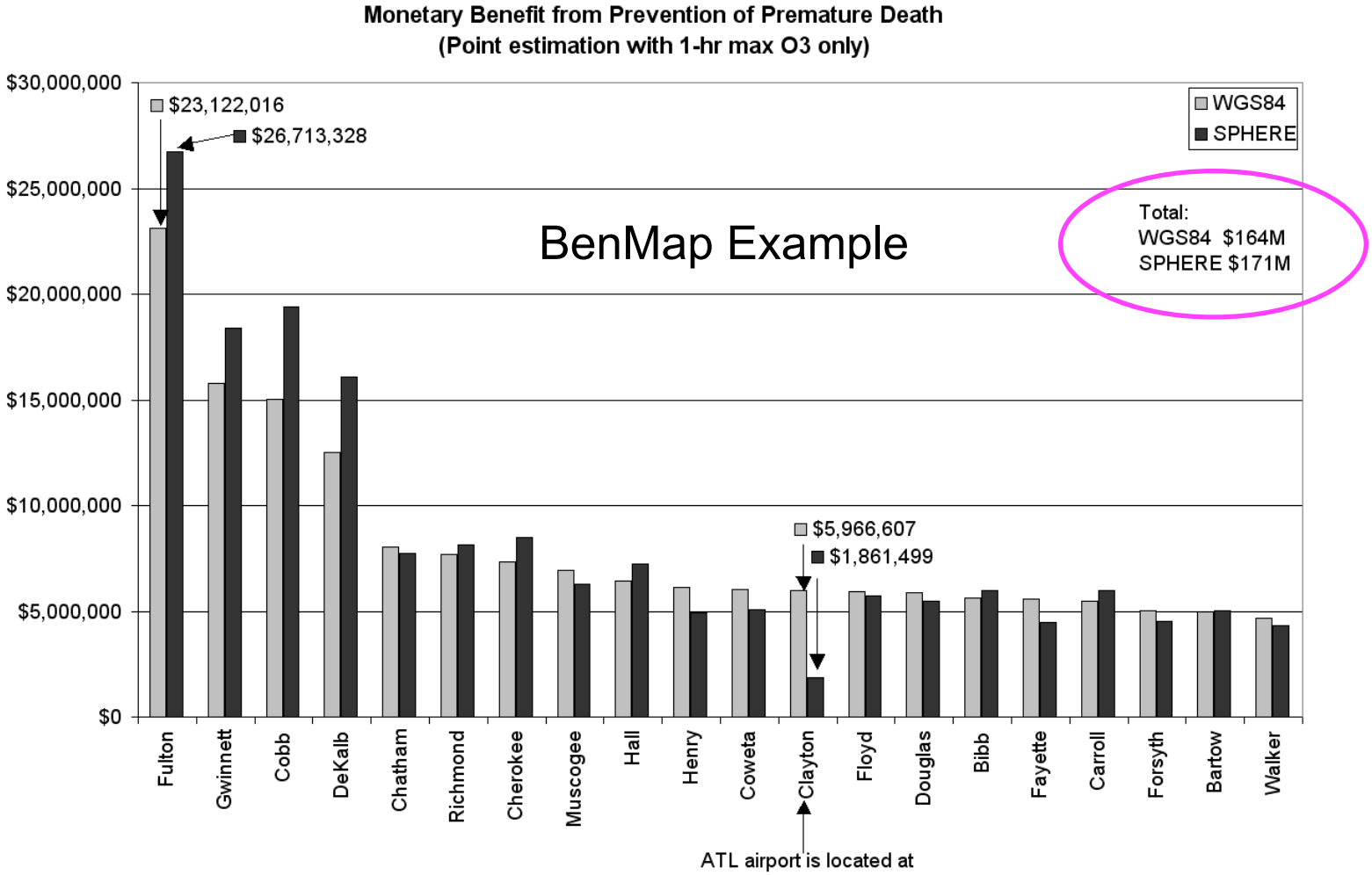
PROJ.4 output

		Areal scale (s)			Ratio To Sphere	
LONG	LAT	CMAQ Sphere	WGS84	CLARKE 1866	WGS84	CLARKE 1866
-91.83475766	27.57216433	1.02804311	1.02790442	1.02790287	0.9998651	0.9998636
-80.23506434	26.31672311	1.03704513	1.03685800	1.03685592	0.9998196	0.9998175
-77.89039337	36.32469693	0.99133558	0.99137234	0.99137274	1.0000371	1.0000375
-91.09313632	37.78669395	0.98957026	0.98961333	0.98961381	1.0000435	1.0000440

- Square of map scaling factors (or areal scales) of LCC using different datums are small although it is still not clear how much the impacts are quantitatively on air quality modeling outputs and/or meteorological modeling outputs.

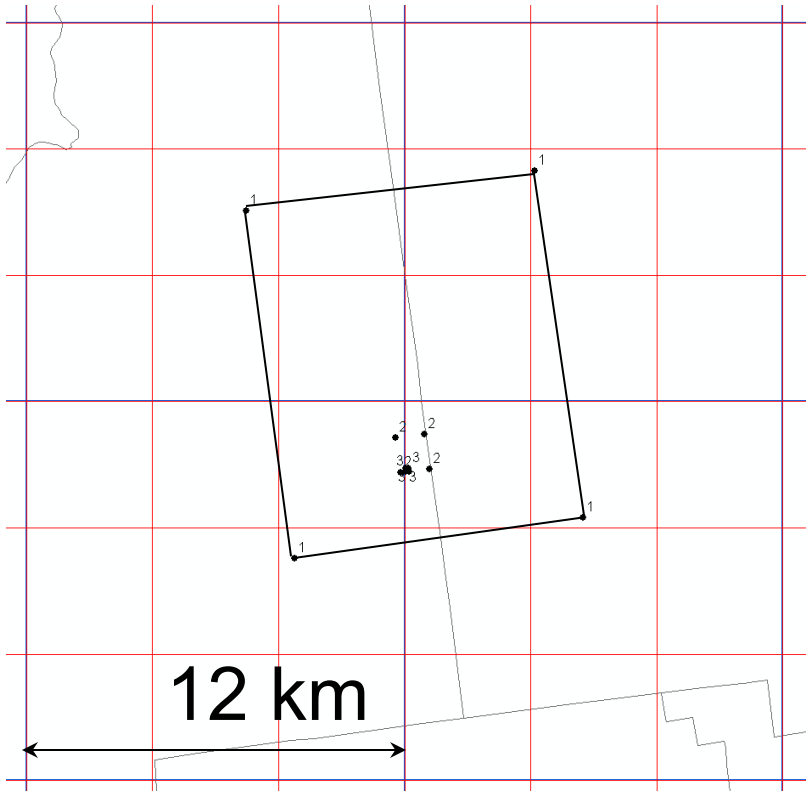


Example of Incorrect Datum Issue: Application of air quality model outputs



Note on Precision

- Numbers and dots associated with them represents the corners of boxes showing uncertainty bounds due to the number of decimal places in latitude and longitude used.
  - '1' indicates the case where (lat, lon) data is given with 1-digit decimal place. So, the point (lat, lon) can be anywhere in the box if we consider round-off from the 2nd decimal place. For example, 40.1 could be any number from 40.05 to 40.14.
- Red lines indicate 4-km modeling grid. It is highly recommended to ***use at least four or preferably five decimal places for location data used in air quality modeling***.



Conclusion (1)

- Metadata reporting
  - It is critical to provide the correct metadata for geospatial information, esp. the original datum on which datasets were created.
  - It is important to provide meta-data on the datum information (i.e. spherical earth) for air quality model outputs when it is transferred to data users and advise **the data users to treat air quality outputs as if they are on the same datum as their input datum.**
  - This issue is particularly important when air quality modelers communicates with air quality output consumers such as GIS users.

- Datum Transformation
  - Mathematically, air quality models solve its partial differential equations as if they are on Cartesian coordinate system (regardless of underlying LCPs) with map scaling factor adjustments.
  - Map scaling factors from a LCC built on most frequently used datums in US regional scale air quality modeling are very close to each other.
  - Therefore, it is recommended:
    - **Use identical datums for all inputs; at minimum, do not mix spheres and ellipsoids throughout modeling processes**
    - **Do not perform datum transformation from ellipsoids to spheres for air quality modeling outputs, vice versa; unless all inputs (including meteorological model inputs) can be prepared for spheres used in typical air quality modeling**

Conclusion (2)

- Data Precision
  - It is highly recommended to use at least four or preferably five decimal places for location data used in air quality modeling.

- Implication on Model Performance Evaluation
  - When ground monitoring data is compared with model outputs, it is highly recommended to perform map projection with the common ellipsoid such as WGS84 used in a specific air quality modeling process.
    - In the past, it is common to use sphere-based LCC projection for locating monitors in model grids. It might have been a reasonable approach because typical modeling grid resolution was 36-km or 12-km. It will be challenging in the future when air quality models run on finer scale grids.
  - When satellite data is compared with model outputs, it is less erroneous not to perform datum transformation from ellipsoids to sphere, vice versa.

Future Work

- Examination of datum impacts on meteorological modeling and the subsequent effects on emission modeling and air quality modeling, esp. local scale modeling
  - Meteorological models can be more sensitive to small perturbation of map scaling factors.
  - Changes in meteorological model outputs will propagate through the rest of air quality modeling processes.
    - Temperatures and other factors are used for some emission estimations such as biogenic emissions.
    - Changes in atmospheric dynamics and emissions will result in changes in air quality model's predictions.
    - At this stage, it is very uncertain how much changes we will see at the end of modeling processes including the application of air quality modeling outputs such as monetary benefit estimations.
  - It is clear that impacts will be more significant at local scale modeling and applications.

Contact Information

Byeong-Uk Kim, Ph.D.
Georgia Environmental Protection Division
4244 International Parkway, Suite 120
Atlanta, GA 30354

Byeong.Kim@dnr.state.ga.us
404-362-2526

Useful Information

- Geodetic Datum Overview
  - http://www.colorado.edu/geography/gcraft/notes/datum/datum.html
- PROJ.4
  - http://trac.osgeo.org/proj/
- Datums and grids: what you don't know can kill you
  - http://findarticles.com/p/articles/mi_m0IBS/is_4_28/ai_94538584/
- Spheres versus Spheroids
  - http://www.crwr.utexas.edu/gis/gishydro06/SpaceAndTime/SphereVsSperoid2006.htm