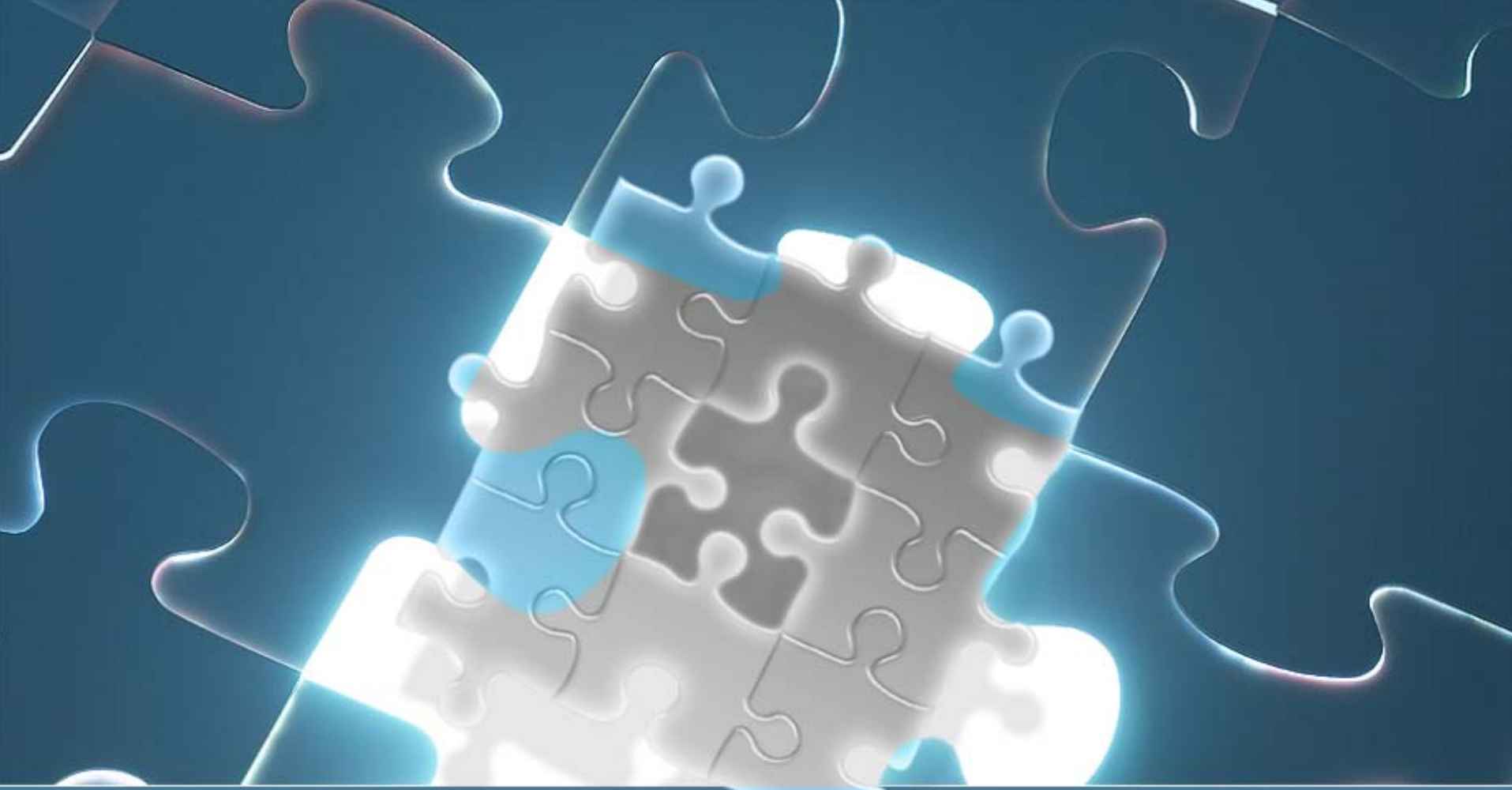


Socioeconomic Forecasting with the Transportation Economic Land Use Model (TELUM)

TELUM

Defined: A user-friendly graphical user interface for implementing the DRAM/EMPAL model

DRAM/EMPAL: a complex set of algorithms for predicting future locations of households and employment based on market relationships

- Market based
- Rigorous
- The most widely used forecasting model in the United States since the 1970s

DRAM/EMPAL

- Developed by Dr. Putman during the 1970s
- Designed to work with widely available data using well-tested relationships
- Evolved over time – ITLUP, IPLUM, METROPILUS, DRAM/EMPAL, TELUM
- Latest evolution is TELUM version 4.0, completed in 2005

Intent

FHWA - $4 million under TEA-21

- Bring to small- and medium-sized MPOs rigorous methods that are normally only available to much larger agencies
  - Increase transparency
  - Reliability
  - Repeatability

Advantages

- Easy to use – substantial embedded instructions and explanations
- Flexible – user defines employment and income categories; constraints can be placed on results
- Makes the connection between land use and transportation
- Established track record throughout the U.S.
- Ongoing support
- Allows stakeholders to weigh in during the input phase, not just the results phase

Pikes Peak Area Council of Governments (PPACG)

- Medium-sized MPO, includes City of Colorado Springs and several smaller municipalities across two counties
- 6 Transportation staff; 1 Forecasting staff; 1 GIS specialist; 0 full-time modelers
- Minimal in-house data development

Forecasting

- During the 1990s, PPACG was a test calibration site for DRAM/EMPAL
  - Census-tract level; disaggregation still needed
  - Promising results
  - More of an academic exercise
- 2003 – forecast completed by local economist/consultant
  - Spreadsheet based
  - Building permits; auto sales; standard multiple linear regression

Changes We Wanted

Previous forecasts…

- Not enough employment categories
  - "Service" – criminal defense attourneys and cashiers
  - Additional categories easy to implement in TELUM
- Inadequate documentation
  - Not explainable
  - Not Repeatable
  - TELUM comes with documentation and bibliography – learn as much or as little as you want
- Need to reinvent the wheel for each forecast

TELUM Modeler

Background:

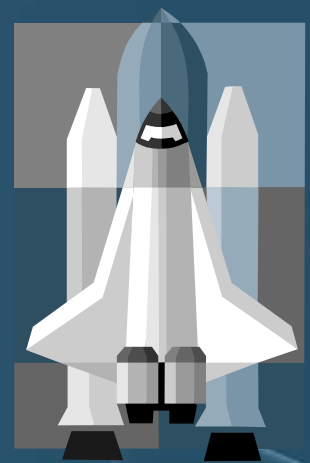
- Mathematical methods, statistics
- Database
- GIS

Needed:

- Understand variables and basic operations
  - Mathematician NOT required
- Familiarity with MS Excel, Access
  - Expert NOT required
- Experience with ArcGIS software
  - Developer/programmer NOT required
- Be able to interpret statistics generated by TELUM

Preparing to Launch

- Present methodology to committees
  - Precedent, users, costs, flexibility
  - Control totals
- Geography – establish analysis zone structure w/member entities
- Data acquisition
- Data cleaning and classification



Geography

- Regular unit of analysis – TAZ
- TELUM has a limit of 499 zones

Statistical Analysis Note

- TELUM is a *statistical* model that relies on sample size for accuracy (3,000 to 10,000 population)
  - Discrete choice models predict individual behavior rather than zone behavior and do not depend on sample size
- More zones does not necessarily mean more accurate
  - Precision vs. Accuracy (uncertainty principle)

Data Needs

Three Types

1) Zonal data
    1) Employment by category [calibrated]
    2) Households by income category [calibrated]
    3) Group quarters (Census defintion)
    4) Land use
2) Regional ratios (vs. national averages)
    1) Employees per household by income category
    2) Income category : job category matrix (PUMS)
3) Travel Times [calibrated]

Data Sources

- Households by income category
  - 2000
    - U.S. Census - TAZs conflated to Census blocks, making aggregation easy
  - 2005 (off-census year)
    - 2000 numbers times 1990-2000 growth rate
    - Field checks in newer high-growth areas
- Employment by category
  - Purchased Claritas (marketing solutions firm) business DB for 2000 AND 2005
    - Relatively cheap
    - Purchased Qwest (telephone company) business DB previously

Data Sources

- Land Use - PPACG asked member entities to provide zoning, comp plans, flood plains, etc.
- Group quarters
  - Dorms – college web sites
  - Barracks – correspondence and economic impact reports
  - Retirement homes and assisted living – Area Agency on Aging database
  - Jails – county web sites
- Regional Ratios – U.S. Census Public Use Microdata Sample (PUMS)
  - Answers the question: In this particular region, for a change in type x employment, what will be the corresponding change in type y households?

Data Cleaning - Households

- Pikes Peak is a rapidly-growing region
  - Undercounts in high-growth area (edge of Colorado Springs)
- Quality control performed using County Assessor's Parcel Search and Google Maps aerial photographs, both free on the web
  - About 20 percent of zones manually checked, only where rapid growth had recently begun

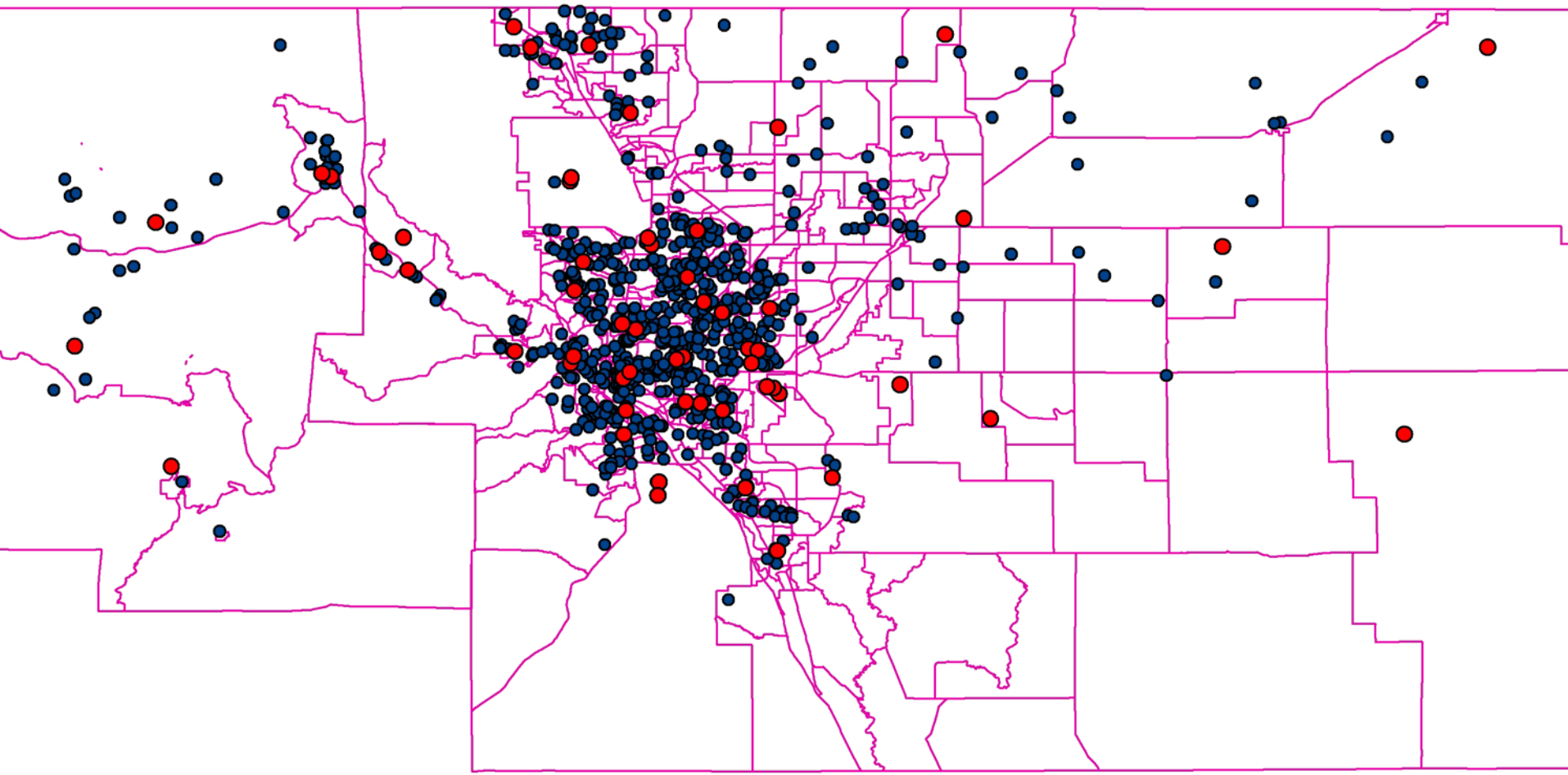
Data Cleaning - Employment

24,000 business points included in purchased database

- Good picture of the regional distribution
- Not very accurate at the analysis zone level
- Cannot check every data point, so innovative quality control was required
- What approach yields the highest rate of return for time spent checking the data?

Quality Control Innovations

- All data points (employers) listing over 80 employees were checked – this represented 50 percent of all regional employees
- Geocoded all points in a GIS, labeled them with city attribute to find incorrect geocoding
- No address or P.O. Box needed to be researched



A Note on Geocoding

- Business databases require refinement before they can be used
- Geocoded by street address without field verification
  - This process is imperfect and is often done against different street layers (where the addresses are stored) in different years
  - PPACG discarded the X,Y coordinates that came with our data and re-geocoded the 2000 and 2005 business points against a common street layer
    - This ensured consistency for calibration
    - Otherwise, calibration might pick up trends that result from different geocoding rather than changes "on the ground"

Data Cleaning Recap

Data will never be perfect, but at minimum…

- Household numbers should be verified for zones on a rapidly-growing urban fringe
- The locations of large employers should be checked
- Land use classifications should be spot-checked for consistency

Simplified Timeline

1. Setting up zone structure, shapefiles, acquiring data – 1 month
2. Data cleaning – 3 months (without dedicated modeling staff)
3. Data prep and Calibration – 1 week
4. **TELUM reveals errata during calibration**
5. More data cleaning – 1 month
6. Data prep and Re-calibration – 1 week

--Periodic technical committee meetings throughout--

Calibration

Interpreting calibration

- If you are bringing the forecast in house for the first time, you are now responsible for *explaining* calibration
- A reasonable definition: calibration is a mathematical description of what happened between two or more known points; this becomes the behavior that the model prescribes to the region in the future
- TELUM is self-calibrating
- Accurate and consistent base year data is critical to proper calibration; otherwise, the model may appear to behave irrationally

Organizational Changes

Models such as TELUM may require new communication pathways within the agency

- Extensive data tasks may require working groups, teams, etc.
- Organizational barriers must be overcome
  - Autonomous departments may need to collaborate
  - New types of accountability may need to be established

Case: Aligning the Planets

Transportation Model

1. Consultant conducts Travel Survey to establish trip rates (2002)
2. Forecasting Department generates pop and employment figures for trip generation (2003)
3. Transportation Department assigns trips to the network (2003)

Case: Aligning the Planets

TELUM – Travel times from travel demand model are forecast input

Income categories from Travel Survey

vs.

Income categories from socioeconomic Forecast

vs.

Income categories from travel demand model

Minimal direct communication/collaboration, no model-wide accountability, no modeling manager

Selling TELUM

- Getting away from the "black box"
  - Staff must demonstrate…
    - Understanding of model fundamentals
    - Ability to interpret results
    - Ability to incorporate local knowledge

$$Wjt\text{-}1 = \Pi_m(E_{jmt-1})^{e(m)}(L6_j/E_j)^{\delta}(P^*j)\theta(K_j)^n$$

Sidebar: Carolina Calling (east vs. west)

Unique Snowflakes

-"Seems like a good model, but it doesn't really apply *here*."

-Rural County

"This is a unique kind of area."

-Everybody I talked to

Water restrictions, retirees, second homes, other policies?...

How TELUM can *implicitly* account for some of the effects through calibration.

Local Knowledge and Constraints

- TELUM can be constrained in four different ways
  - Set the number of HH or jobs in a zone by income or employment category
  - Set a total number of HH or jobs without specifying income/category distribution
  - Set minimum total of HH or jobs
  - Set maximum total of HH or jobs

This feature created comfort among committee members and the public, made TELUM a collaborative process

National Averages!@#*?

TELUM uses region-specific ratios to determine settlement patterns, NOT national averages

U.S. Census: Public Use Microdata Sample (PUMS), $70

Pikes Peak	LI	UI
MNFG	80	20
PROF	20	80

Land Use and Transportation

- Abundance of research shows causal connection
- Many simple models use gravity model alone to show effect of travel cost on settlement pattern
- TELUM can incorporate congested travel times from a transportation demand model
  - Much more realistic than gravity model
  - Calibrated against HH and employment
  - Can be updated as the TELUM is run to reflect road projects completed during the forecast period (every five-year period, if desired)

Density and Adjacency

Western communities very sensitive to high densities

- Settlement patterns predicted to be more dispersed
- Radically different densities will not be found in close proximity
- TELUM
  - existing densities influence forecast-year densities (not strictly deterministic)
  - Conditions in adjacent zones will also influence density – excess demand for development can "spill over"

Death of a StrawMan

Preliminary results

- Calibration complete, statistics look OK
- TELUM shows that base data still needs work
- No constraints set yet

Small Area Forecast Subcommittee Review……

Issues

- Dramatic decreases in households and employment in several zones
- Too much basic employment in rural areas
- Not enough research into unique zones
  - Regional parks
  - Military installations
  - Revisit base data to look for incorrect input

Schools of thought and the two-party system

Model liberals: “Free the model!”

- The model knows best
- Use constraints when absolutely necessary

Model Conservatives: (The model goes to jail)

- Do not allow zones to decrease
- Do not allow densification

Bounding the Model

- "Floors"
  - Number of households shall not decrease
  - Number of jobs shall not decrease by more than 1 percent per year
- "Caps"
  - Once land is consumed, redevelopment shall not occur
- Implementing bounds required many spreadsheets – acreages, densities, updating floors ea. 5 years

Stakeholder Feedback

1. Nobody meets deadlines
    1. Checkpoints needed
    2. Proactive communication
2. Everything is relative
    1. A model only does whatever we think
    2. Example: "Household construction"
3. "Thanks for nothing"
    1. Only certain information can/should be used to constrain a model
        1. Desired outcomes, zoning, astrology, not usable
        2. Need specific plans, platted developments, policies
    2. Be very specific about what you are asking for

The Importance of Story

Prejudice: Agency wants to replace stakeholder input with a mathematical model, and they probably don't even know how it works.

1. Present the model and demonstrate understanding before beginning
2. Inform stakeholders that they will be asked for input - what type and when
3. Re-convene at decision points to discuss difficulties and successes and ask advice (this is the story)
4. Tell stakeholders how their specific input was used; if not, why not

Retrospective

- Leave enough time to perform several iterations of the model
- Model area is a critical decision
  - Calibration – where averages fail
  - Example: Teller County water policy
  - Rural vs. urban and the attraction of vacant land
- Listen to TELUM – calibration statistics reveal early mistakes

Information

http://www.telus-national.org/telum/

TELUM is free

Tech support is free