

Updating the U.S. Census Bureau's Geospatial Database with Information from State, Local, and Tribal Government Partners

Leveraging ArcGIS and SDE-on-Oracle
to process local data files for the
GSS-I Partnership Program

Jay E. Spurlin

Software Architect and Development Manager for the
GSS-I Feature Source Evaluation software system

GSSI Partnership Program

Program Agenda

- § Overview
- § MAF/TIGER Database
- § Geographic Partnership Program

Overview

US Census Bureau – Geography Division

- § The Geography Division handles geographic and cartographic activities in support of the Census Bureau's statistical programs.
- § We continuously update features, boundaries, geographic entities, and address information
 - § TIGER (Topologically Integrated Geographic Encoding and Reference)
 - § MAF (Master Address File).
- § We research geographic concepts, methods, and standards needed to facilitate data collection and dissemination.

Overview

Geographic Support System Initiative (GSS-I)

- § In support of the 2020 Decennial Census, the Census Bureau is evaluating areas on which to focus for a traditional, on-the-ground address canvassing operation – and areas in which it is not necessary.
- § How will we decide which areas should be considered?
 - § GEO has evaluated the MAF/TIGER database and assigned quality indicators to each of the census tracts
 - § The MAF/TIGER is being updated through the Geographic Partnership Program
 - § A Targeted Address Canvassing strategy has been developed that contains an inventory of criteria for evaluation

Overview

Geographic Support System Initiative (GSS-I)

- § The Geographic Partnership Program is now underway.
 - § GEO is receiving both address and spatial data from invited partners.
 - § The data are being evaluated and integrated with the MAF/TIGER database.
 - § GEO is working to determine what feedback to give to the partners about their data.
- § To identify areas on which to focus for Address Canvassing, we will combine information from:
 - § the evaluation of the current MAF/TIGER database
 - § the partner data
 - § predictive modeling

MAF/TIGER Database

- § Implemented on Oracle Spatial Topology Data Model
 - § Custom topological primitive features
 - § Over 50 feature “layers” implemented in dozens of feature tables
 - § Hierarchical features (features built on other features)
- § Many software systems in MAF/TIGER system – all using Oracle
 - § GWCS (GSSI Workflow Control System)
 - § School District Review Program (SDRP) Crowd Source application
 - § PCPTR (Problem Capture, Prioritization, Tracking, and Reporting)
 - § QIs (Quality Indicators)

MAF/TIGER Database

- § Physical features
 - § Roads
 - § Hydrography
- § Geographic entities (both legal and statistical)
 - § States
 - § Counties
 - § County Subdivisions
 - § Places
 - § School Districts
 - § State Legislative Districts
 - § Census Tracts
 - § Tabulation Blocks
- § Addresses and structure points

Geographic Partnership Program

Acquisition of data

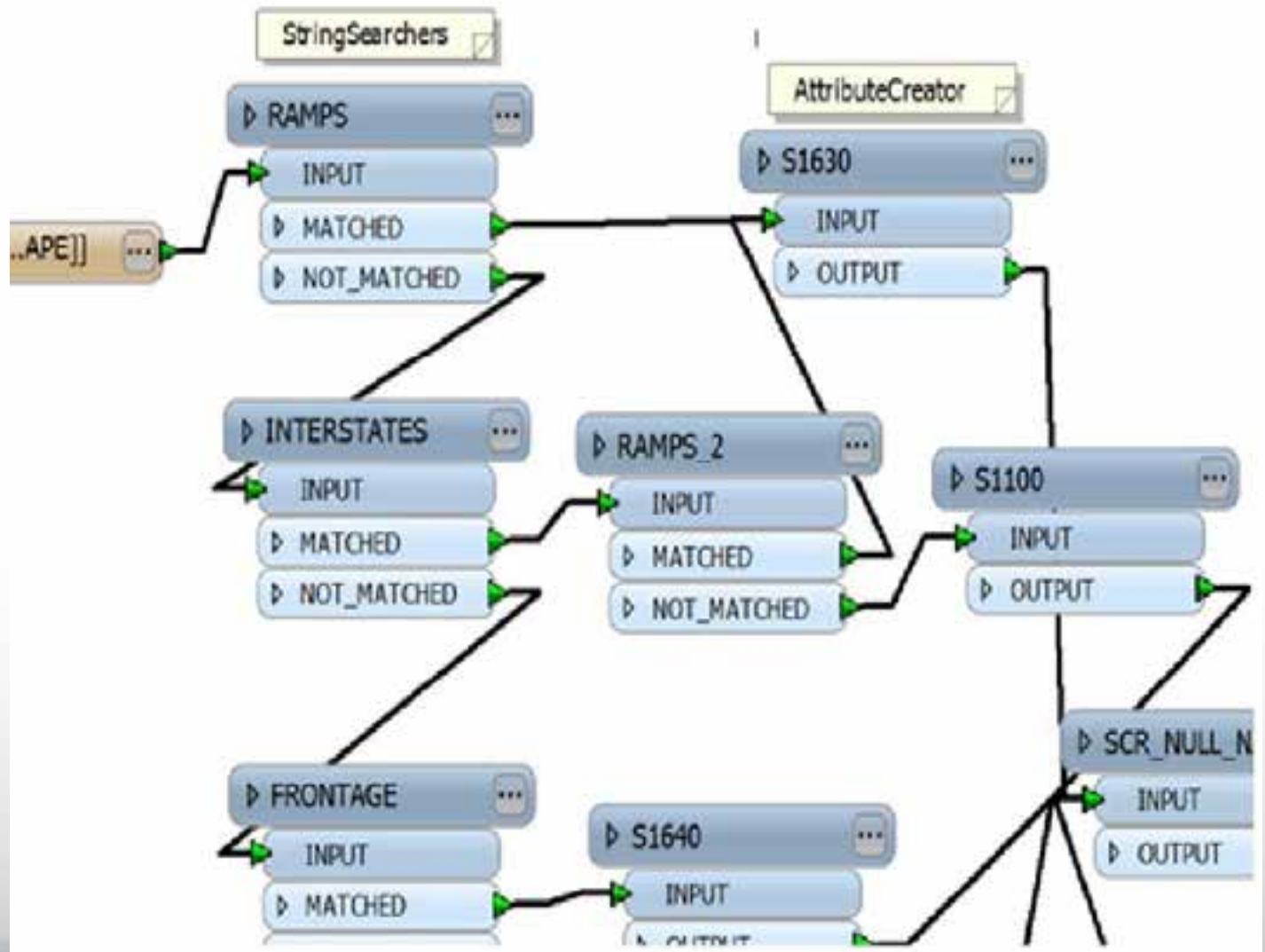
- § Data are provided at the state, county, and local level.
 - § A partner provides a set of source files
 - § The source files are moved inside the Census firewall via a secure web-exchange module
 - § The content inventory of the files undergoes initial verification
 - § The files are preserved, as supplied, for later reference

Geographic Partnership Program

Evaluation of data

- § Data are evaluated
 - § More detailed content assessment is done
 - § The data and metadata provided are verified against published standards, it ensure the files meet minimum guidelines for content and metadata
 - § The files are prepared for automated processing
 - § ETL, including re-projection and mapping to a standardized schema
 - § Feature name standardization, using both feature and address standardizers
 - § A series of (mostly) automated checks is run, which provides metrics about the data in the files
 - § Geometric change detection, based on a buffered overlay
 - § Topology assessment, looking for any topological issues
 - § Feature naming gap check
 - § Address matching and geocoding, against the MAF

Evaluation of data – ETL (Extract, Transform, and Load)



Geographic Partnership Program

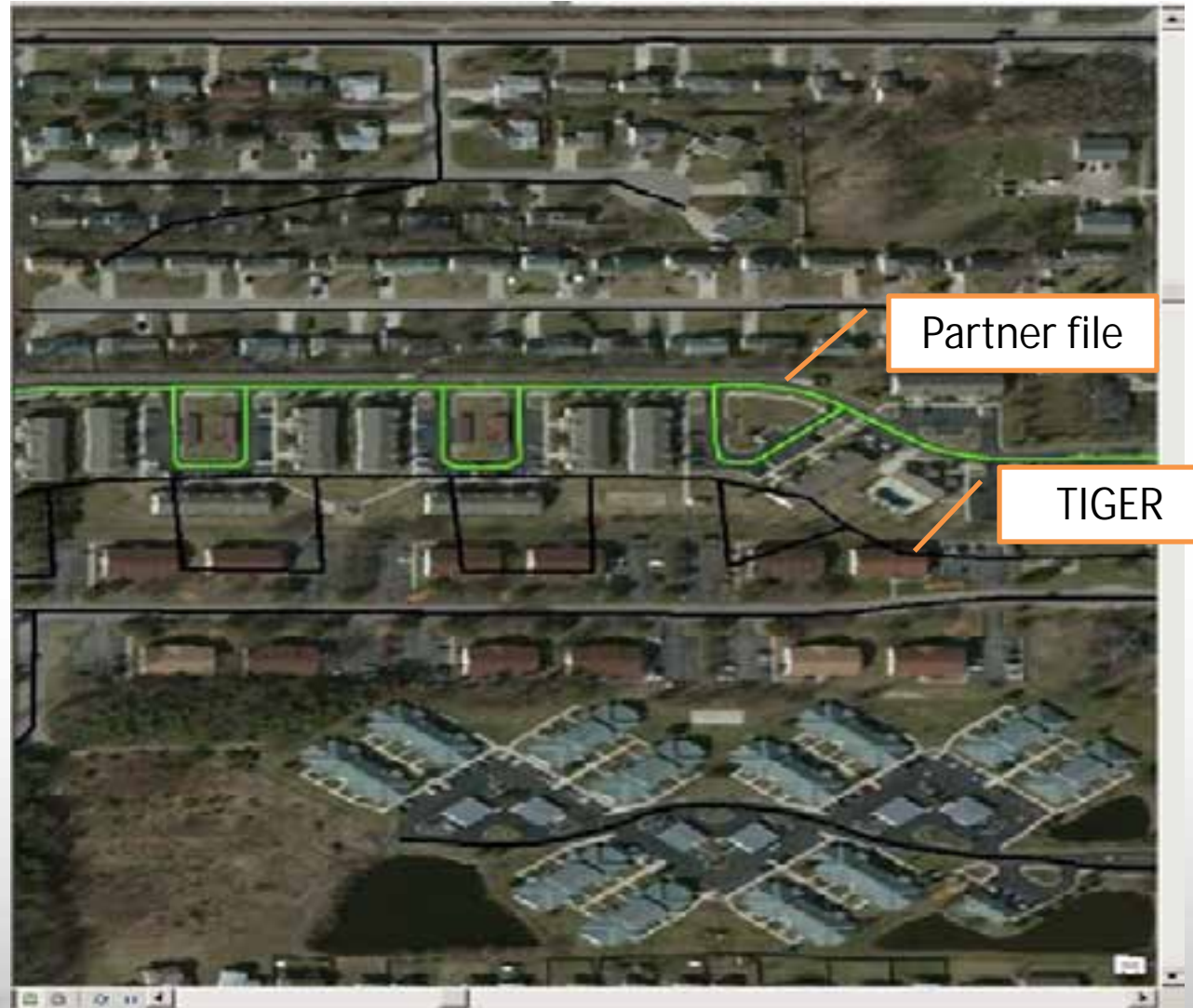
Evaluation and Integration of data

§ Update method is chosen

- § An interactive review is conducted, in which the files and their associated metrics are reviewed and a decision is made how to capture any new data
- § Some feature files are processed by an interactive editor, using the partner file, imagery, and other sources
- § Some feature files are processed through a semi-automated conflation process
- § Address files are processed through an automated update system

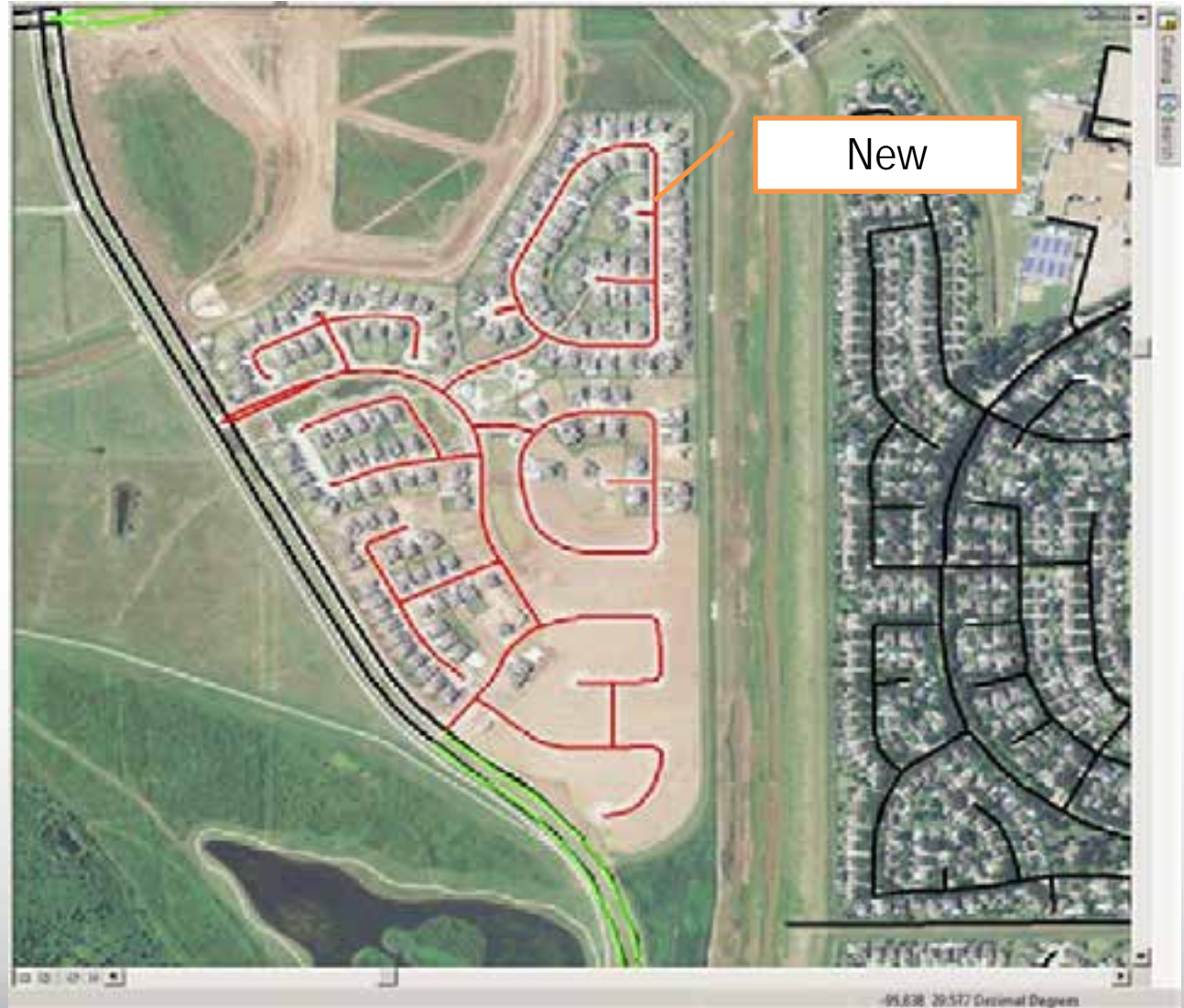
Evaluation of data – Interactive Review

Partner – green
MAF/TIGER – black



Evaluation of data – Interactive Review

Partner – red
MAF/TIGER – black



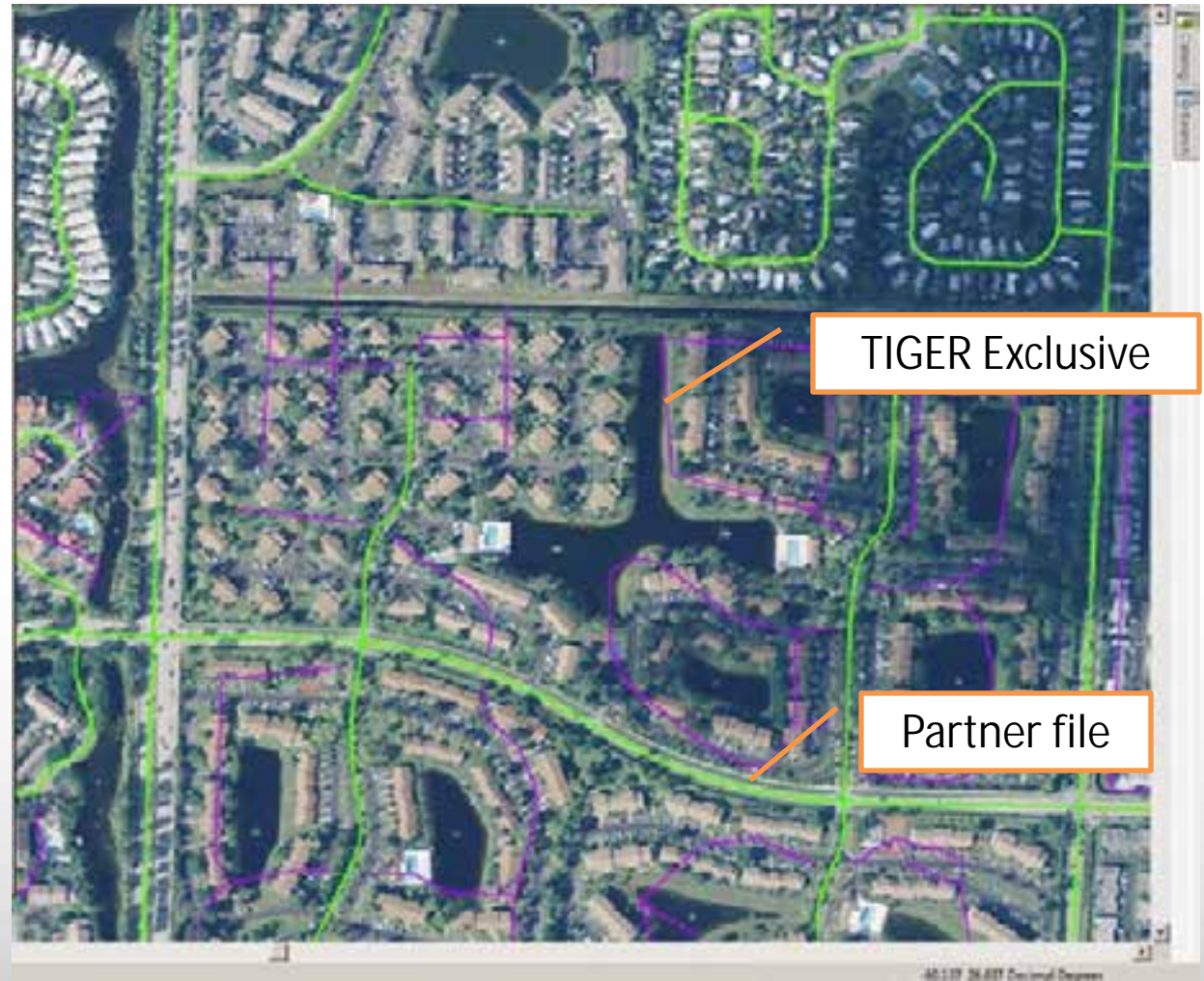
Evaluation of data – Interactive Review

Partner – red
MAF/TIGER – black



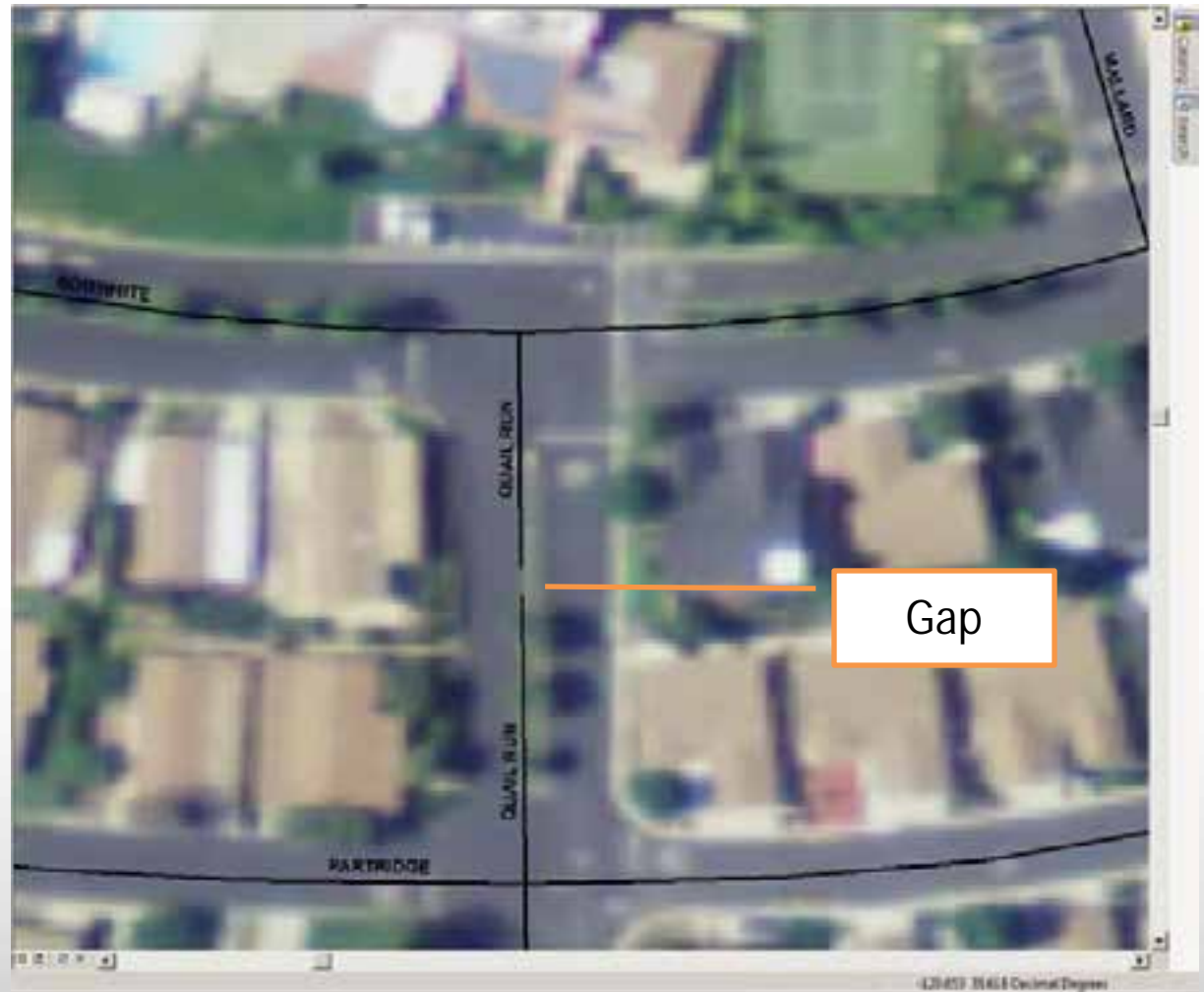
Evaluation of data – Interactive Review

Partner – green
MAF/TIGER – purple



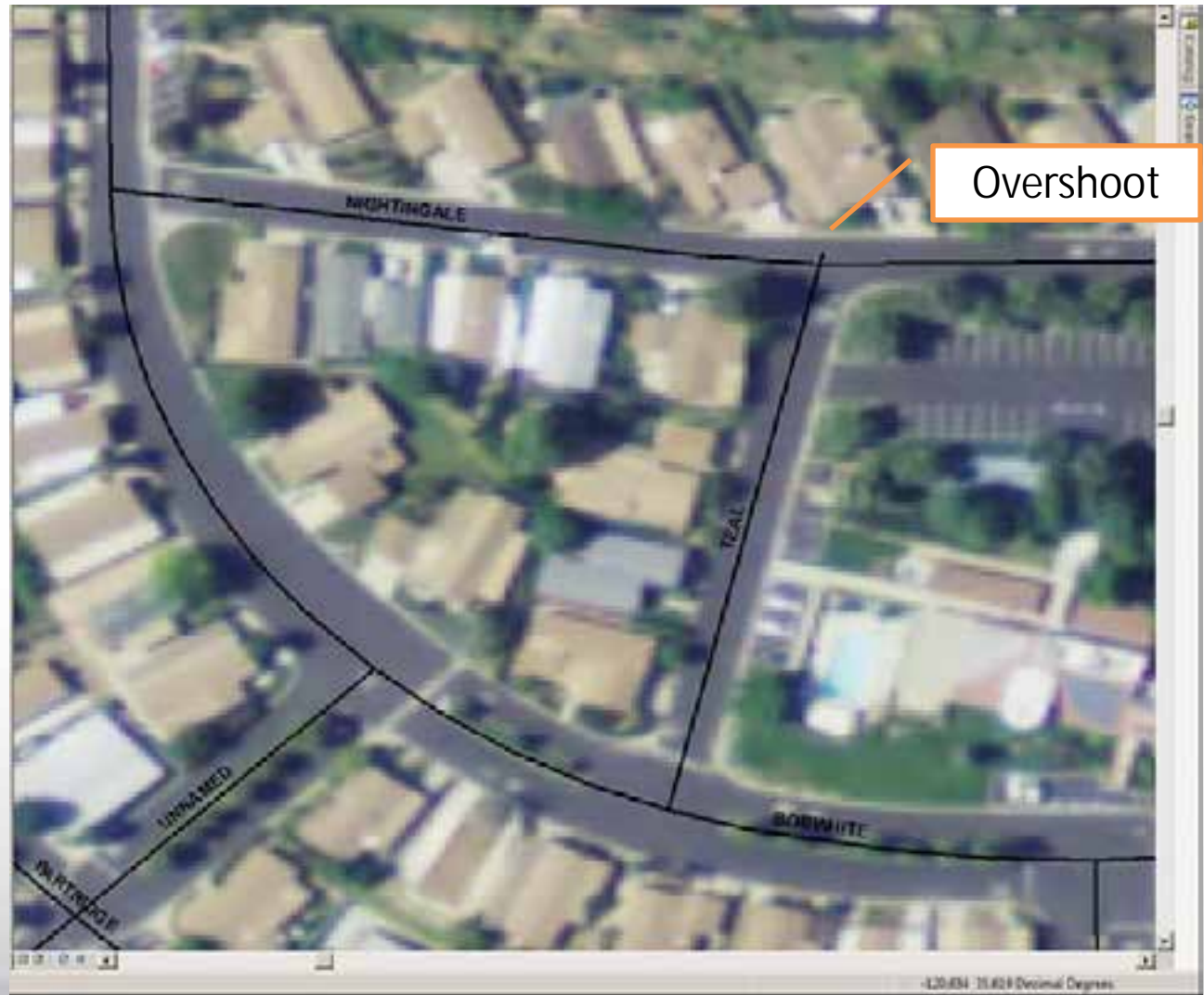
Evaluation of data – Interactive Review

Partner – black

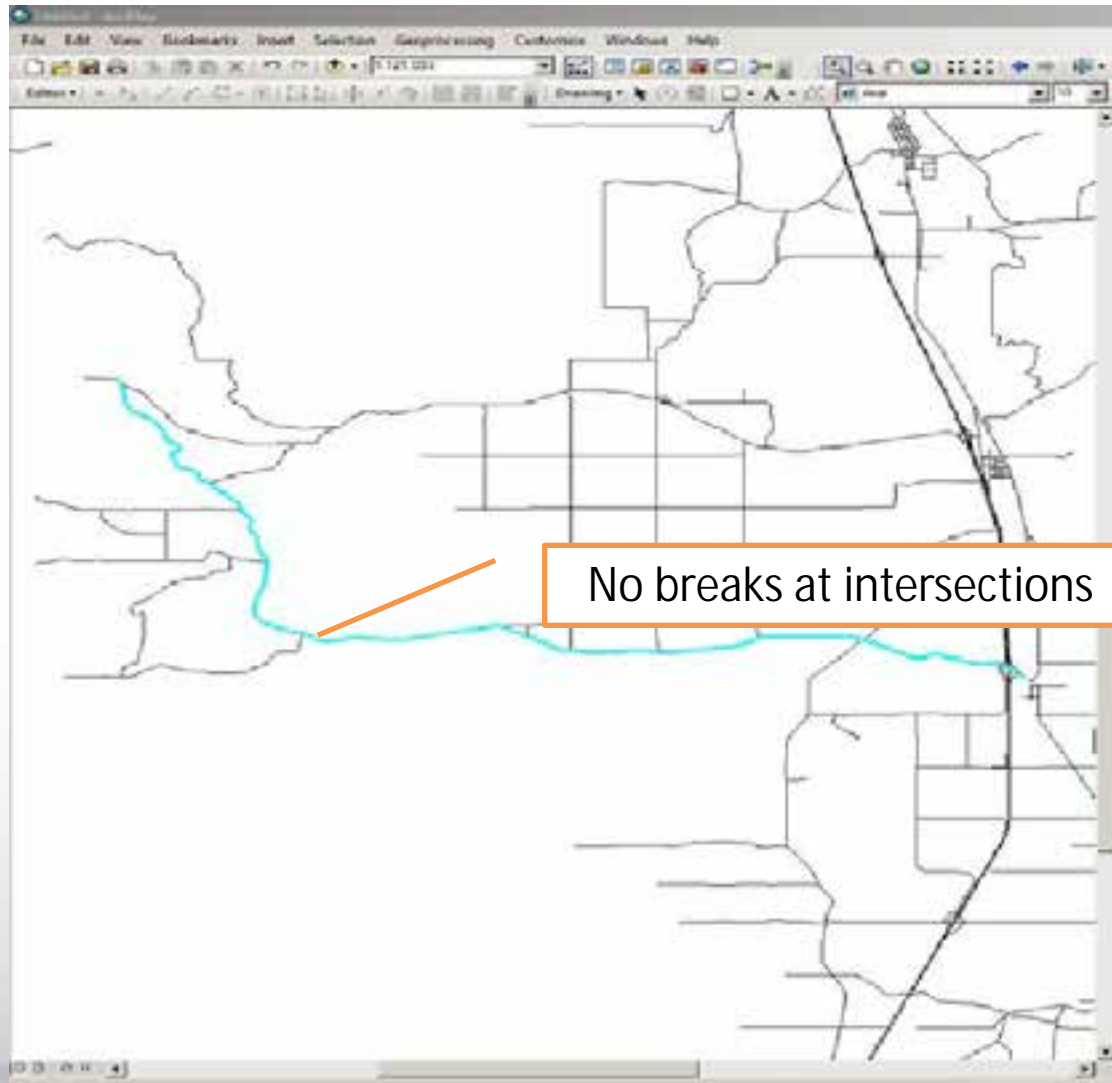


Evaluation of data – Interactive Review

Partner – black

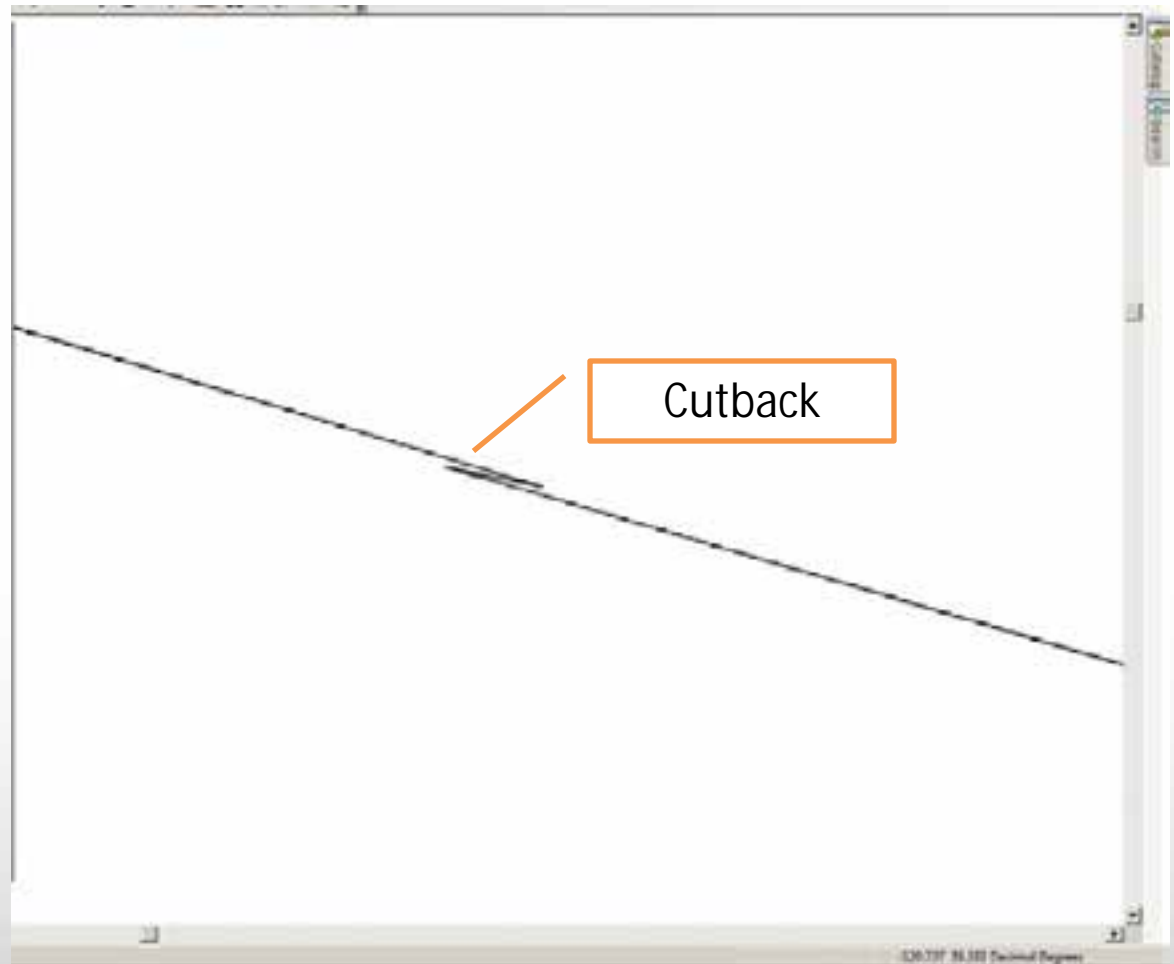


Evaluation of data – Interactive Review



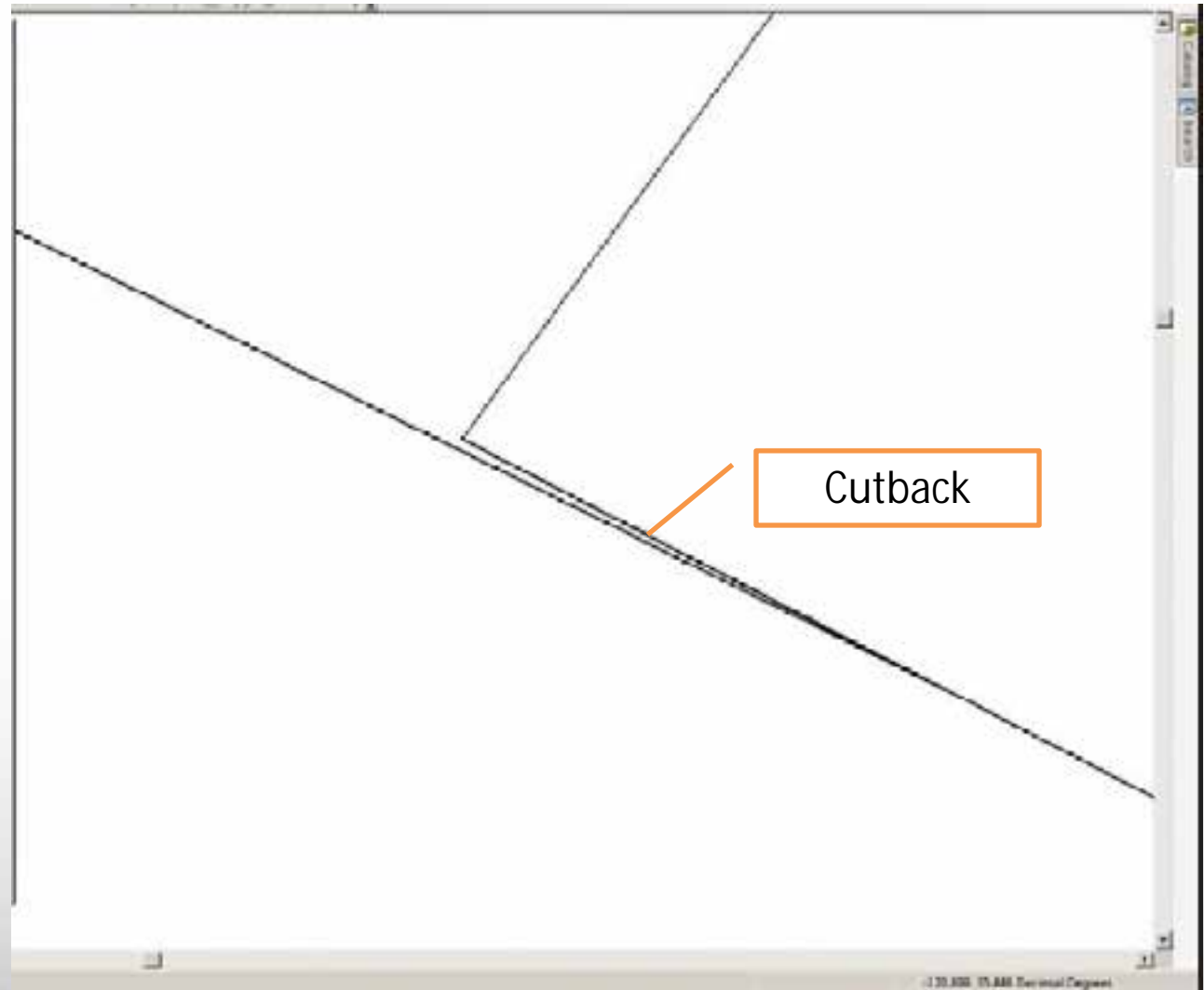
Evaluation of data – Interactive Review

Partner – black



Evaluation of data – Interactive Review

Partner – black



Geographic Partnership Program

Integration of data - Conflation

- § Data are integrated into the MAF/TIGER database
 - § Scrubbing: Any data that are not useful for updating the MAF/TIGER database get removed from the files
 - § Matching: Features or addresses are matched to the MAF/TIGER database, using an automated process.
 - § An interactive review (using a custom ArcMap Extension) identifies any issues and allows adjustments to be made.
 - § Update: All approved changes are applied to the MAF/TIGER database, using an automated process.
 - § Metadata is captured and stored in the database, with quality metrics about each feature

Integration of data – Conflation Proposal Review Tool

The screenshot displays the Conflation Proposal Review Tool interface. The main window shows an aerial map with various colored overlays representing different data layers. A left-hand pane lists the layers, including 'PARTNER_FL18A2000_F000887_04', 'PROPOSAL_FL18B2000', and 'EDGE_FL18_BUDG_FL18A2000'. The 'PROPOSAL_FL18B2000' layer is expanded, showing options like 'Add - No Flag Set', 'Add - Accept', 'Add - Reject/Interactive', 'Reshape - No Flag Set', 'Reshape - Accept', 'Reshape - Reject/Interactive', and 'Invalid ARB or Type'. The 'EDGE_FL18_BUDG_FL18A2000' layer is also expanded, showing options like 'MTDB - Roads', 'MTDB - Hydro', 'MTDB - Railroad', 'MTDB - Non-Visible Boundaries', 'MTDB - Other Linear Features', and 'MTDB - Other Edges'. The 'PARTNER_FL18A2000_F000887_04' layer is selected, showing options like 'rv_boundary', 'extension', 'smallFace', 'Floater', 'Cluster', and 'Current'. The 'Current' layer is expanded, showing options like 'Red - Band_1', 'Green - Band_2', and 'Blue - Band_3'. The right-hand pane is the 'Conflation Proposal Review Tool' control panel. It has tabs for 'Data Load', 'Review', 'Flag TOOR', 'Error', and 'Review'. The 'Review' tab is active. The 'Review Type' is set to 'Cluster'. There is a checkbox for 'Include resolved proposals in this record 3 of 141'. Navigation buttons include 'Previous', 'Next >', and 'Jump to:'. There are buttons for 'Select Cluster', 'Remove Selection', 'ACCEPT', and 'REJECT'. The 'ACCEPT' section has buttons for 'As Is', 'New Construction', and 'Major Modification'. The 'REJECT' section has buttons for 'Out of Slope', 'Paper Street', 'Inaccurate', and 'Matching'. There is a 'SEND TO INTERACTIVE' section with buttons for 'OK' and 'New Construction'. At the bottom, there is a checkbox for 'Capture metadata from imagery' and buttons for 'Set to ADD', 'Set to REMOVE', and 'Set to DELETE'.

Integration of data – Conflation Proposal Review Tool

The image displays three overlapping screenshots of the Conflation Proposal Review Tool interface, illustrating its various functional areas.

Left Window (Review Options): Shows the 'Review' tab with a 'Review Type' dropdown set to 'Floater'. It includes a checkbox for 'Include resolved proposals in tour record' (1 of 7), navigation buttons for 'Previous' and 'Next', and a 'Jump to' field. Below are buttons for 'Select Cluster/Failure' and 'Remove Selection'. The 'ACCEPT' section contains 'As Is', 'New Construction', and 'Major Modification'. The 'REJECT' section contains 'Out of Scope', 'Paper Street', 'Inaccurate', and 'Matching'. A 'SEND TO INTERACTIVE' section includes 'Other' and 'New Construction'. At the bottom, there is a checked checkbox for 'Capture metadata from imagery' and buttons for 'Set to ADD', 'Set to REALIGN', and 'Set to DELETE'.

Middle Window (Actions): Shows the 'Review' tab with a vertical list of actions: 'Reshape Edge', 'Delete Edge', 'Extend Edge', 'Trim/Shorten Edge', 'Holy Cow, What a Mess!', 'Boundary Problem', 'Some Other Problem', and 'Remove Flag'.

Right Window (Parameters): Shows the 'Parameters' tab with several numerical input fields and checkboxes. The parameters are: 'Snapping tolerance (meters)' (5), 'MSP tolerance (meters)' (1), 'Coincidence tolerance (meters)' (5), 'Small area threshold (square meters)' (200), 'Floater should snap tolerance (meters)*' (5), 'Maximum overshoot distance (meters)' (10), 'MSP Report Tolerance (meters) (for UDB report)*' (5), and 'Zoom factor (for feature tour)*' (1.5). Below these are text input fields for 'Imagery web service location' and 'TMeta web service location', and a 'Validate Parameters' button. Footnotes at the bottom indicate that '*' denotes a Finalize parameter and that some parameters take effect in the Review tab.

Geographic Partnership Program

Integration of technology architecture

§ COTS software integration

§ ArcGIS

- § Partner data is spatially enabled using ArcSDE

- § ArcMap is used to view and analyze data

§ FME (Feature Manipulation Engine)

- § Extraction of reference layers, ETL (to load partner data into an SDE geodatabase)

§ 1Spatial Radius Studio

- § Used for automation of conflation processes

Geographic Partnership Program

Integration of technology architecture – cont'd

§ COTS software integration

§ ADE (Acquis Data Editor)

- § Built on MapViewer

- § Utilized by GATRES (Geographic Acquis-Based Topological Real-time Editing System) for interactive editing

§ Redwood Cronacle

- § Used for process control, scheduling, and monitoring

Thank you!

§ Questions?

§ For more information:

§ Jay E. Spurlin

§ jay.e.spurlin@census.gov

§ U.S. Census Bureau

§ <http://www.census.gov>

§ GSS-I:

§ <http://www.census.gov/geo/www/gss/>

§ TIGERWeb:

§ http://tigerweb.geo.census.gov/tigerwebmain/tigerweb_main.html