



Using Conflation for Keeping Data Harmonized and Up-to-date

Dan Lee

Esri Inc., Redlands, California, USA

dlee@esri.com

Agenda

1. **Conflation for data harmonization**
2. **Conflation tools and workflows**
3. **Conflation scenarios**
4. **Work and research in progress**
5. **Conclusions**

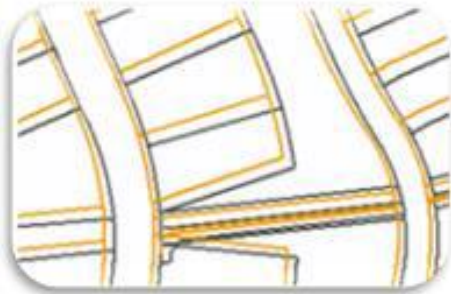
1. Conflation for data harmonization

Needs from NMAs and GIS organizations

- **Gathering needed information from various data sources**
- **Building and updating comprehensive, consistent, reliable databases for multi-purposes and collaborative work**

Challenges

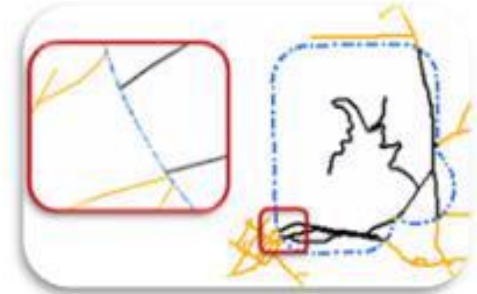
- **Spatial and attribute discrepancies in multi-source datasets**
- **Lack of efficient tools to harmonize overlapping and adjacent datasets**



(a) Parcels (LA Co. DPW)



(b) Roads (ICGC)



(c) Roads at border (RMS)

Conflation

- Identifies corresponding features
- Makes spatial adjustments and attribute transfer
- Ultimately reconciles overlapping and adjacent datasets for the best accuracy, completeness, and consistency
- Helps keep data harmonized and up-to-date
- Optimizes data quality and usability

Example efforts in automation

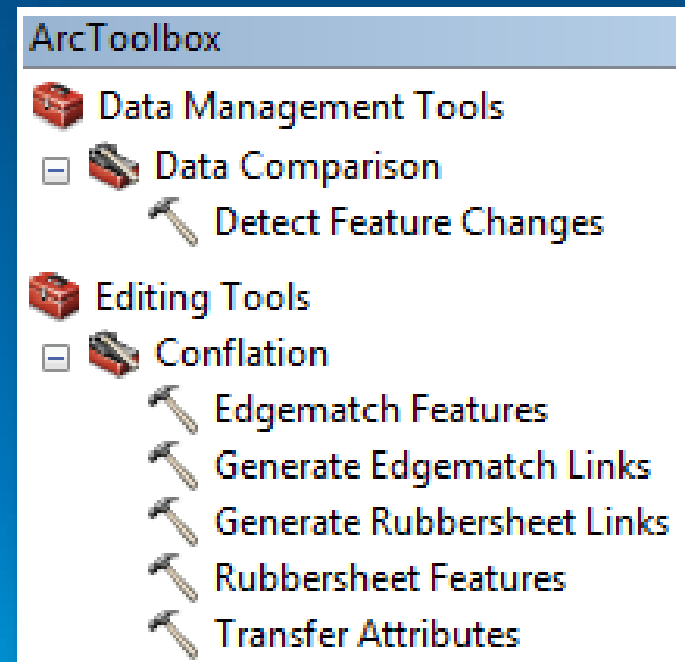
- For hydrographic data (Stanislawski et al, 2002)
- For road data (Li and Liu 2012; Abdolmajidi et al, 2014)

2. Conflation tools and workflows

Our initial focuses

- **Developing tools in Geoprocessing framework, aiming at high feature matching accuracy (not promising 100%)**
- **Starting with linear features (roads, parcel lines, etc.)**
- **Providing information to facilitate post-processing**
- **Building workflows**

ArcGIS 10.2.1



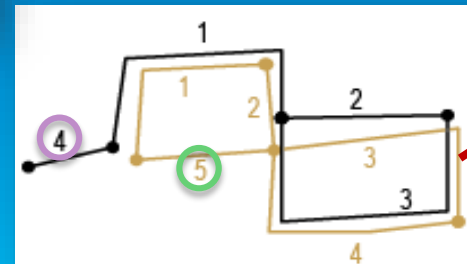
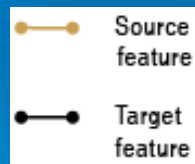
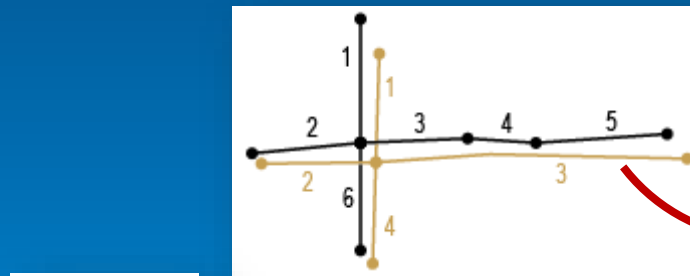
For overlapping datasets

Feature matching (FM)

- Based on proximity, topology, pattern and similarity analysis, and optional attributes

Match table

- Storing match information



1:1 and 1:m matches

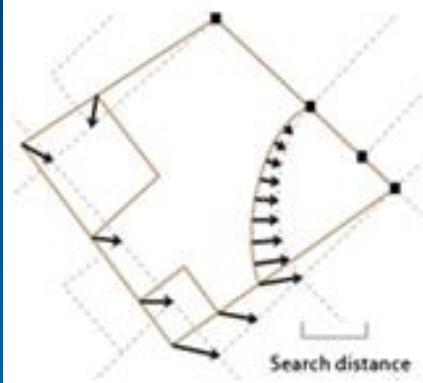
| OBJECTID * | SRC_FID | TGT_FID | FM_GRP | FM_MN | FM_CONF |
|------------|---------|---------|--------|-------|---------|
| 1 | 1 | 1 | 1 | 1:1 | 100 |
| 2 | 2 | 2 | 2 | 1:1 | 100 |
| 3 | 3 | 3 | 3 | 1:3 | 100 |
| 4 | 3 | 4 | 3 | 1:3 | 100 |
| 5 | 3 | 5 | 3 | 1:3 | 100 |
| 6 | 4 | 6 | 4 | 1:1 | 100 |

m:1 and m:n matches

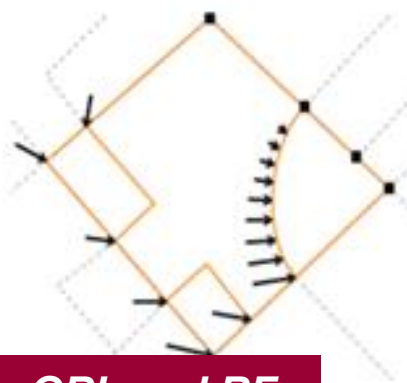
| OBJECTID * | SRC_FID | TGT_FID | FM_GRP | FM_MN | FM_CONF |
|------------|---------|---------|--------|-------|---------|
| 1 | 1 | 1 | 1 | 2:1 | 100 |
| 2 | 2 | 1 | 1 | 2:1 | 100 |
| 3 | 3 | 2 | 2 | 2:2 | 100 |
| 4 | 3 | 3 | 2 | 2:2 | 100 |
| 5 | 4 | 2 | 2 | 2:2 | 100 |
| 6 | 4 | 3 | 2 | 2:2 | 100 |
| 7 | 5 | -1 | -1 | N/A | 0 |
| 8 | -1 | 4 | -1 | N/A | 0 |

Three FM-based tools

Generate Rubbersheet Links



Rubbersheet Features



GRL and RF

Transfer Attributes

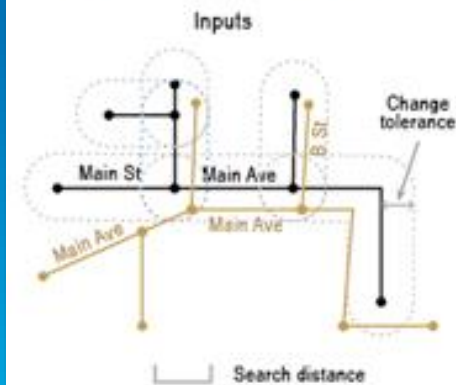


Transferred attributes

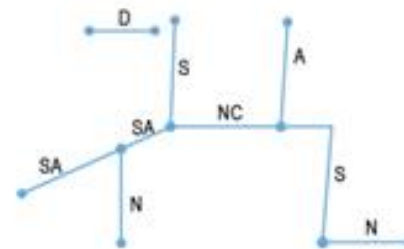


TA

Detect Feature Changes



Output with CHANGE_TYPE values

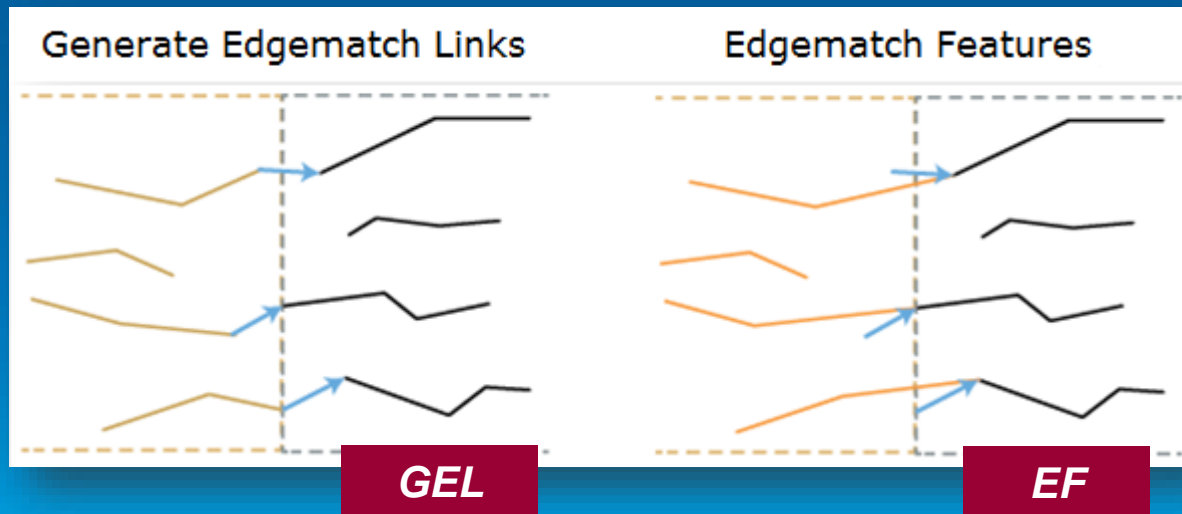


DFC

For adjacent datasets

Edgematching (EM)

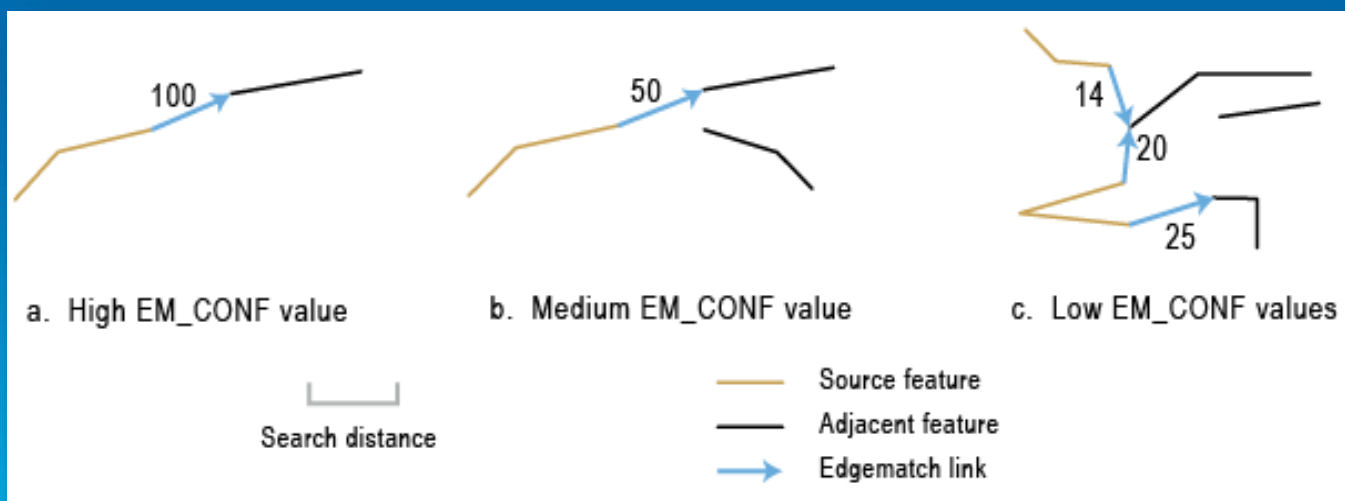
- Based on proximity, topology, continuity, and attributes (optional) analysis



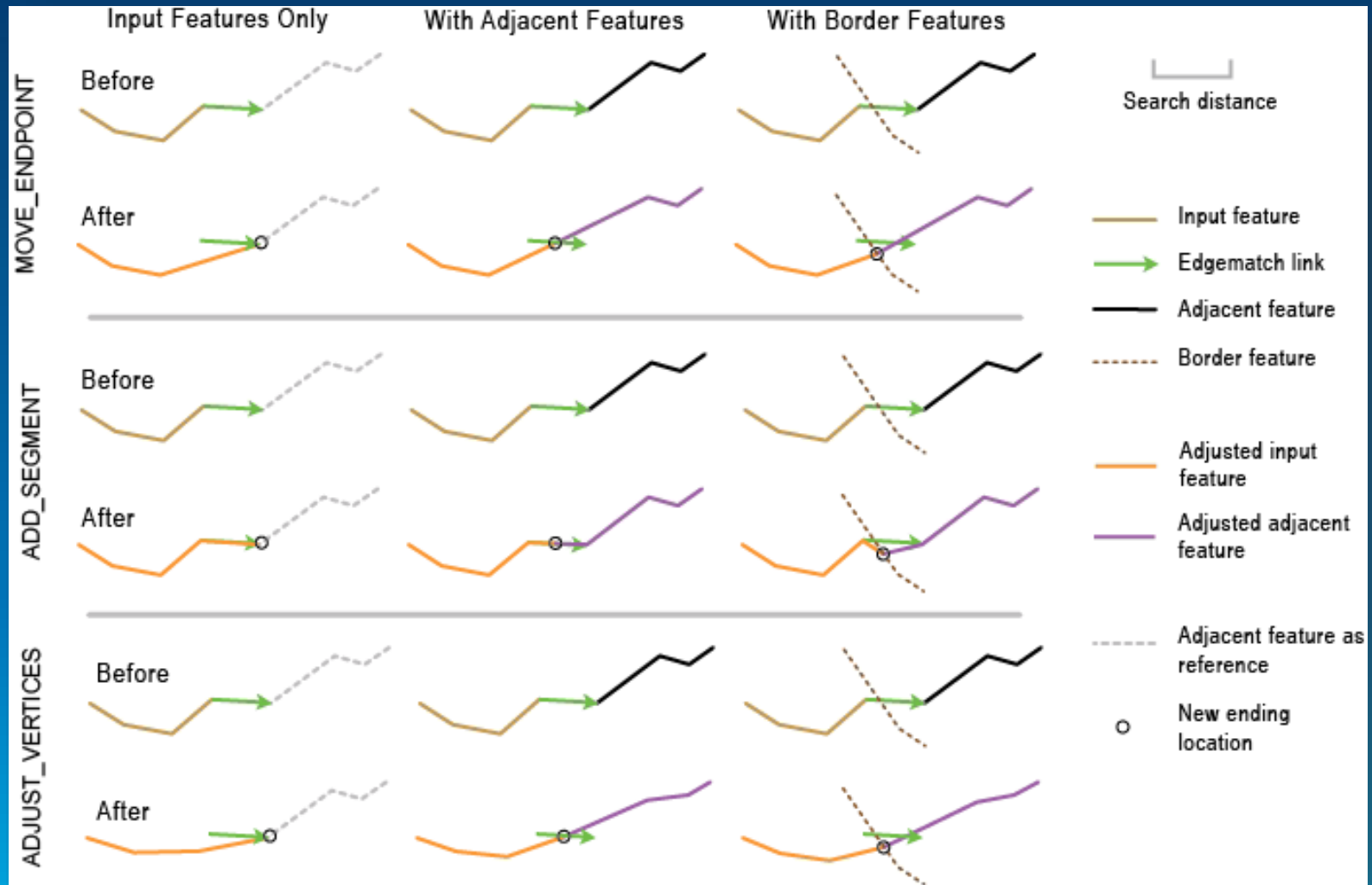
Edgematch links

- **Linking source --> target**
- **EM_CONF: the higher value the stronger match (less ambiguity)**

| emLinks | | | | | |
|---------|------------|----------|---------|---------|---------|
| | OBJECTID * | SHAPE * | SRC_FID | TGT_FID | EM_CONF |
| | 33 | Polyline | 923 | 3577 | 100 |
| | 34 | Polyline | 936 | 3323 | 100 |
| | 35 | Polyline | 963 | 2667 | 50 |
| | 36 | Polyline | 986 | 3071 | 100 |
| | 37 | Polyline | 987 | 3117 | 11 |
| | 38 | Polyline | 992 | 690 | 100 |
| | 39 | Polyline | 1002 | 590 | 33 |



Options for connecting features



General conflation workflows

Three components

Pre-processing

- In same projection
- Valid geometry
- Clean topology
- Consistent attributes
- Relevant features

Conflation and evaluation

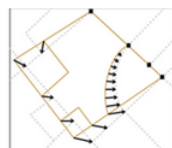
- Conflation workflow tools

Interactive review - editing

- Queued review
- Interactive tools

**Supplemental tools: –
download from arcgis.com
– search for conflation**

GP Conflation_Workflow_Guidelines



This item contains conflation workflow guidelines and supplemental t

Geoprocessing Sample by 1812

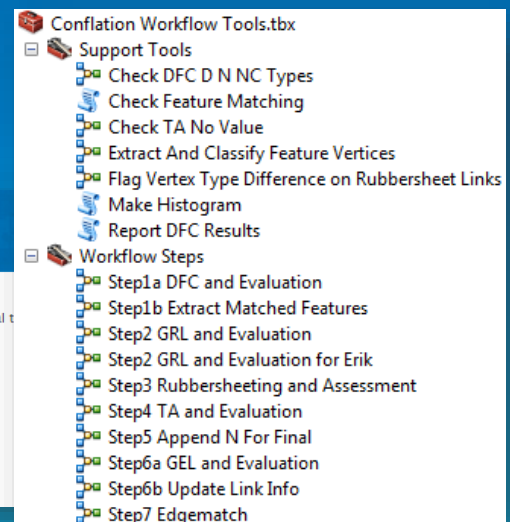
Last Modified: July 14, 2015

★★★★★ (0 ratings, 8 downloads)

Sign in to rate this item.

Facebook Twitter

OPEN



3. Conflation scenarios

Conflation plays essential roles in data harmonization

- **For overlapping datasets**
 - **Single task scenario**
 - **Comprehensive task scenario**
- **For adjacent datasets**
- **For multi-scale datasets**
 - **Establishing links and transferring attributes**
- **For data updating**
 - **Detecting feature changes**

Overlapping datasets spatial adjustment

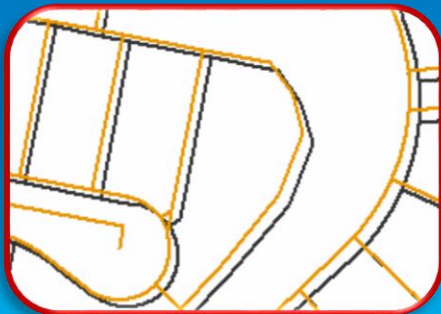
Single task on parcel lines (DPW, LA Co.)

- **Generate rubbersheet links, QA, and adjustment**
- **Estimated matching accuracy 93.84%**

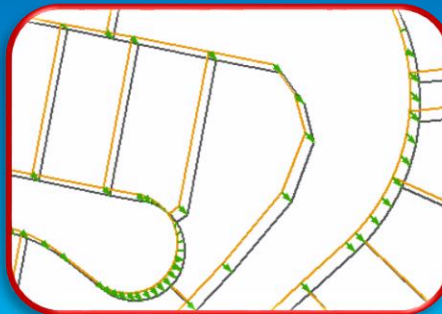
(a) Input parcels



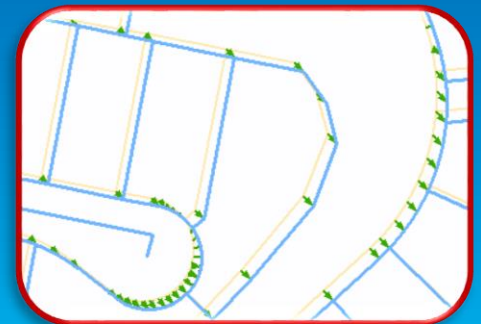
(b) Results of automated process with links and potential issues flagged



(c) Inputs enlarged



(d) Rubbersheet links



(e) Adjusted results

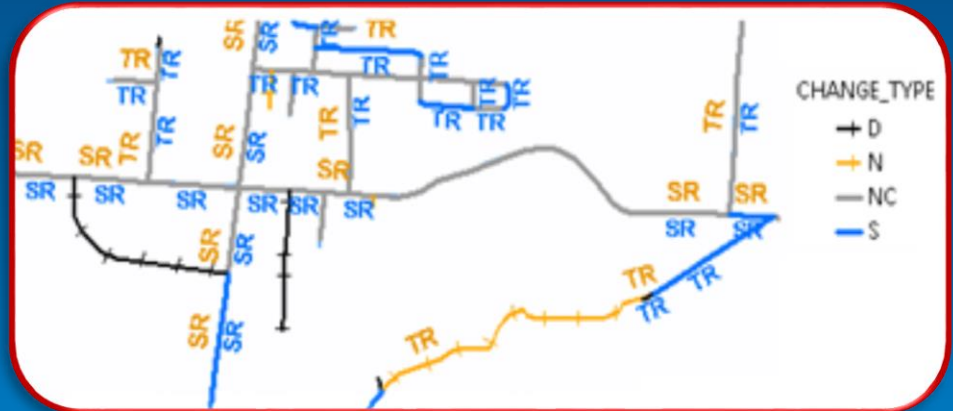
Overlapping datasets unification

Comprehensive task on roads (ODOT)

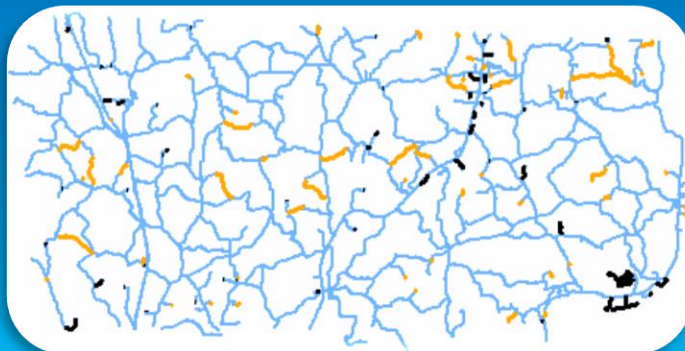
- **Five sub-workflows (automated and interactive QA)**
- **Estimated matching accuracy 98.74%**



(a) Input road datasets



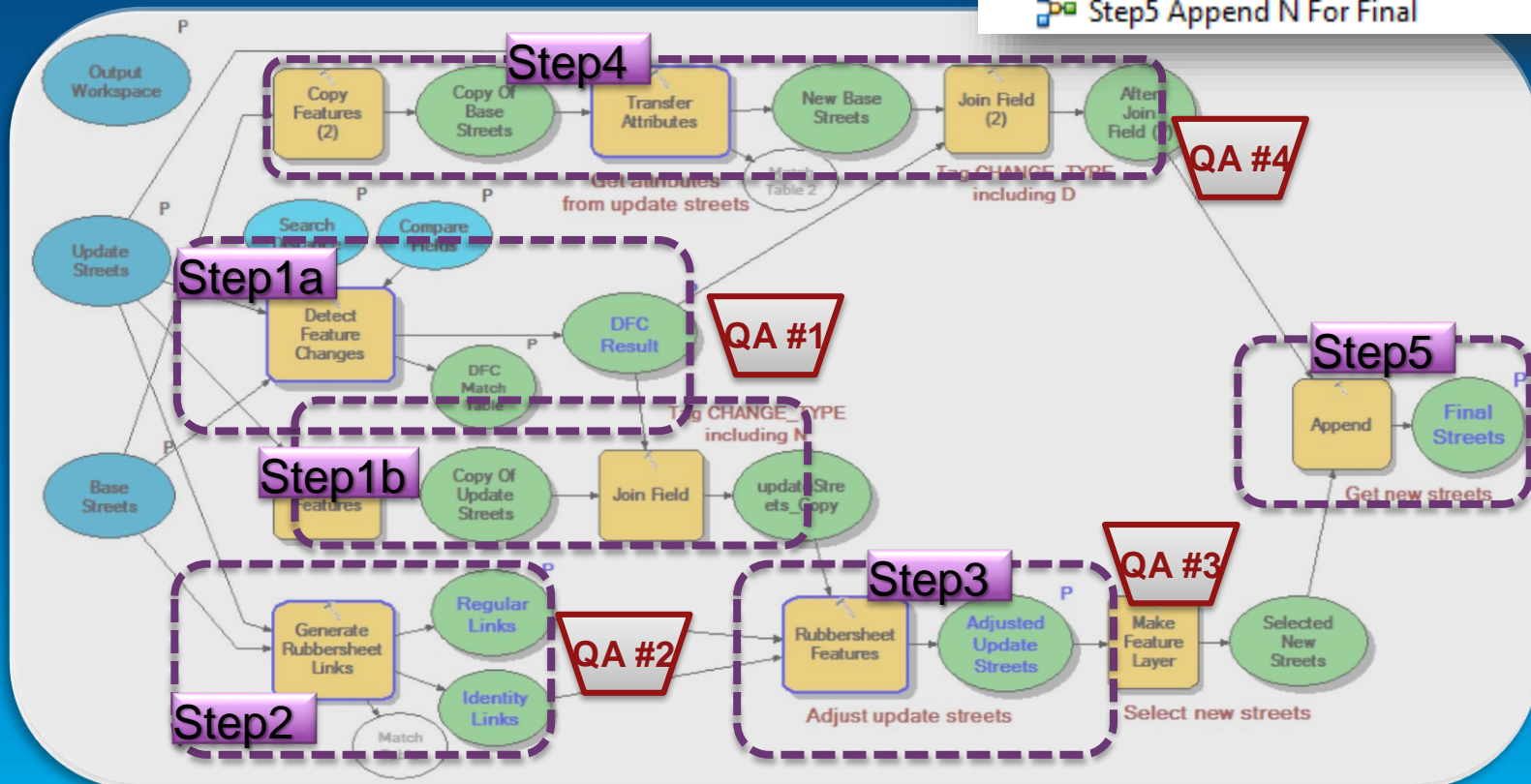
(b) Feature differences identified by DFC tool and transferred attributes



(c) Unified results

Workflow Breakdown

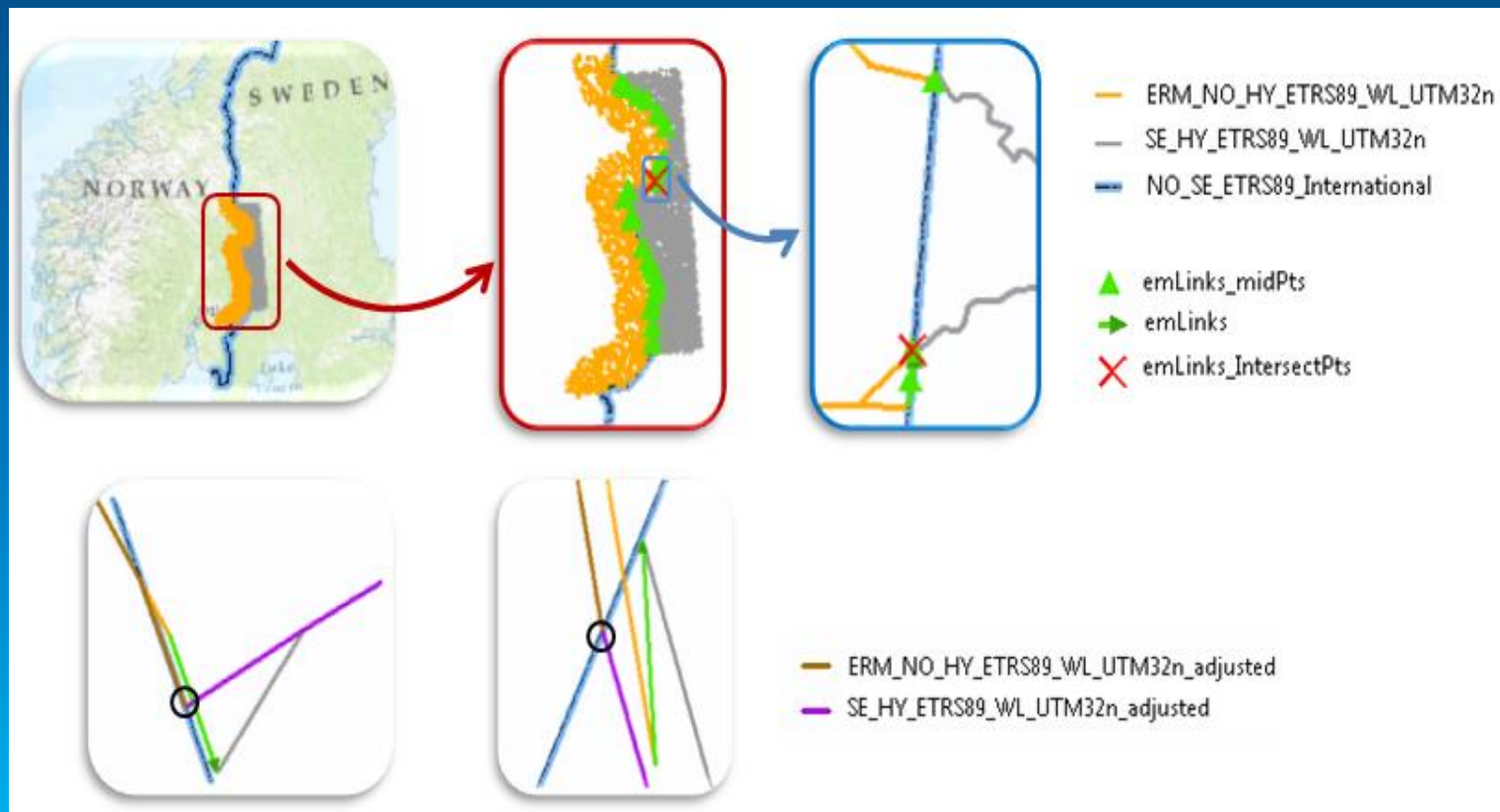
- Workflow Steps**
- Step1a DFC and Evaluation
 - Step1b Extract Matched Features
 - Step2 GRL and Evaluation
 - Step3 Rubbersheeting and Assessment
 - Step4 TA and Evaluation
 - Step5 Append N For Final



Adjacent datasets edgematching

Cross-border (Norway-Sweden) hydro datasets (ELF)

- **Five sub-workflows (automated and interactive QA)**
- **Estimated matching accuracy 95%**

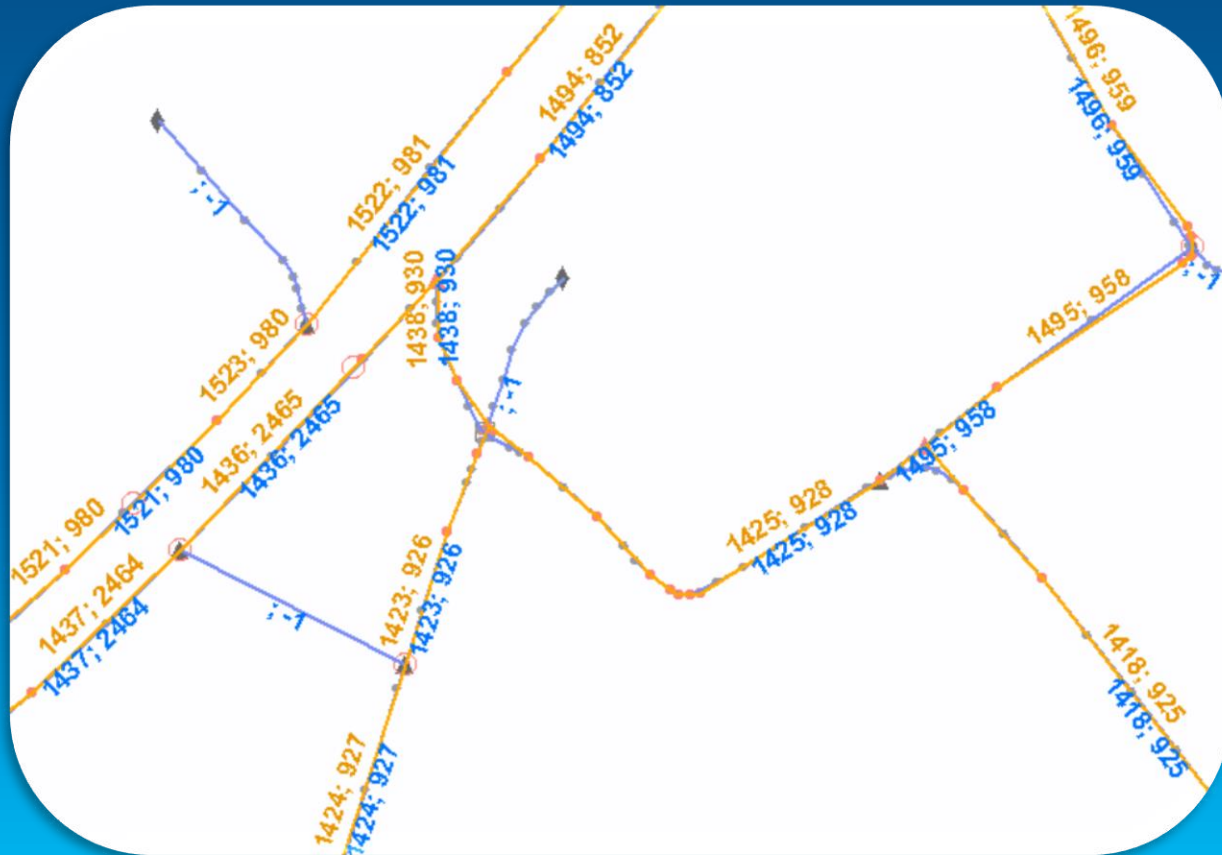


Edgematching workflow



BT-5M and BT-25 roads (ICGC)

- Match group IDs as links; attributes transferred 25k -> 5k
- Estimated matching accuracy 95%



Change detection and updating

Update and base streets (in-house test data)

- Unify spatial and attribute changes using workflows
- Propagate changes across scales through appropriate actions, e.g. generalization.

Update features with new streets and attributes

| UpdateStreets | | | | |
|---------------|----------|----------|----------|--------------|
| OBJECTID * | Shape | UniqueID | Priority | Shape_Length |
| 1 | Polyline | <Null> | | 88.544127 |
| 2 | Polyline | 2 | L | 298.351168 |
| 3 | Polyline | 3 | L | 166.198499 |
| 4 | Polyline | 4 | L | 83.759794 |
| 5 | Polyline | 5 | L | 109.060742 |
| 6 | Polyline | 6 | L | 378.689485 |
| 7 | Polyline | 7 | L | 44.406315 |
| 8 | Polyline | 8 | M | 178.995089 |
| 9 | Polyline | 9 | M | 82.505747 |
| 10 | Polyline | 10 | M | 326.699698 |
| 11 | Polyline | 11 | <Null> | 113.142624 |
| 12 | Polyline | 12 | <Null> | 59.711738 |
| 13 | Polyline | 13 | <Null> | 90.582727 |
| 14 | Polyline | 14 | <Null> | 18.331248 |

Base features

| BaseStreets | | | | |
|-------------|----------|------------|----------|--------------|
| OBJECTID * | Shape * | NAME | Priority | Shape_Length |
| 130 | Polyline | W FERN | M | 82.64188 |
| 149 | Polyline | W FERN | M | 316.419224 |
| 156 | Polyline | W FERN | M | 87.413082 |
| 106 | Polyline | SHERWOOD | L | 138.32864 |
| 129 | Polyline | SHERWOOD | L | 154.974493 |
| 73 | Polyline | SAN RAFAEL | L | 94.960556 |
| 82 | Polyline | SAN RAFAEL | L | 86.952636 |
| 87 | Polyline | SAN PABLO | L | 259.025112 |
| 102 | Polyline | SAN PABLO | L | 162.609722 |
| 114 | Polyline | NOTTINGHAM | L | 138.470439 |
| 140 | Polyline | NOTTINGHAM | L | 217.51654 |
| 66 | Polyline | MILLS | L | 90.060797 |
| 69 | Polyline | MILLS | L | 91.095418 |



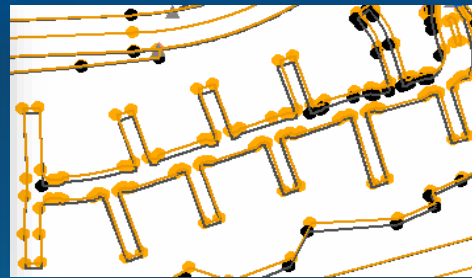
Detected changes



4. Work and research in progress

Feature matching enhancements

- **Pattern recognition and similarity analysis**



Q: How can data modeling better support feature matching?

Optimization of workflows

- **Improvements on evaluation**
- **Integration of QA processes**

Handling multi-theme data

- **Cross-border features; features in context (spatially related features)**

Q: How to describe and store interrelationships among them?

Research on potential hybrid approaches

- **GPS, imagery, and lidar data – accurate sources of ground truth**
- **Identification and extraction of target locations and features for conflation**
- **Change detection and updating**

5. Conclusions

Thanks to:

- Department of Public Works (DPW), Los Angeles County, USA.
- Institut Cartogràfic i Geològic de Catalunya (ICGC), Barcelona, Spain.
- Ohio State Department of Transportation, USA.
- National Institute for Water and Atmospheric Research (NIWA) and Land Information New Zealand (LINZ) - Crown Copyright Reserved.
- Resource Management Service, LLC, Birmingham, AL, USA.
- All others who supported us along the way.

5. Conclusions – cont.

It is important to keep data consistent and up-to-date

- **Ensuring reliable analysis and high quality mapping**
- **Supporting multi-purpose applications and collaborations**

Conflation tools and workflows help achieve the goals

- **Highly automated processes (accuracy ~ 85-95%)**
- **Manageable interactive inspections and editing**
- **Resulted in unified and streamlined quality data**

We will continue working closely with NMAs and others

- **Incorporating your feedback**
- **Improving conflation solutions to meet the challenges**



Any comments or questions?

Thanks

