The Eastern Transportation Coalition
The ABCs of Conflation: TMC, LRS, OSM – What Happens When You Muck It Up – August 19, 2021
Question and Answer Summary

Connecting the Dots: Pennsylvania's Conflation Experience

Q: **Ioannis Tsapakis (Texas A&M Transportation Institute):** Have you documented this conflation process?

A: Steve Gault (Pennsylvania DOT): We had some consultants help during the original effort, and they prepared a document describing step-by-step what is done. We used that documentation to pull together our slides. The documentation isn't overly extensive.

Q: **Chris Snyder (ICF International):** What tools did you use for this conflation? Mostly ArcGIS? Or did you use some open-source tools as well?

A: Steve Gault (Pennsylvania DOT): It ended up being a python script, so not directly in ArcGIS.

Q: **Tom Edinger (DVRPC):** Have you conflated to the INRIX XD data besides just TMCs? Would that change the methodology?

A: Steve Gault (Pennsylvania DOT): We have. We started with TMCs and then we used the same process and replicated it for INRIX XD data.

Q: **Lauren Geng (Texas A&M Transportation Institute):** What is the accuracy rate in general with this conflation method?

A: Steve Gault (Pennsylvania DOT): We don't have a specific number, but we had some of our MPOs doing validations and they reviewed the shapefiles. Once we got to the point where we weren't seeing feedback on anomalies, we knew we were in the "good enough" range.

Q: **Stan Young (The Eastern Transportation Coalition):** Is there an objective threshold for 'good enough?'

A: Steve Gault (Pennsylvania DOT): It's a matter that's subjective, but the goal is something useable without getting to the point where it's too costly and cumbersome that we don't get useable data to derive reasonable decisions. We're not using this to engineer a project – just for screening and regional operations plans. The planning process is where we need to be focusing our efforts.

C: **Paul Bushore (MARC):** Probably a wider thought, but does the reporting process for average annual daily traffic (AADT) or speed limit need to be linear-based? Perhaps reporting the data by XY location can then be applied to the reporting agency's network as opposed to having segmentation be an issue. I know it comes with its issues and it still doesn't speak to the "one-segment-to-rule-them-all" - just a thought- we can talk about this later.

A: Steve Gault (Pennsylvania DOT): For federal performance reporting, the calculations are based on some definition of segments and the federal regulations do allow states to define their own segments. Most states just used TMC segments to avoid manipulating

data for all the other elements. We just wanted to get to the point where we had one speed limit and one volume for each TMC for federal reporting.

C: Stan Young (The Eastern Transportation Coalition): We come down to point-based attributes and segment-based attributes - we need to find a way to marry the two.

Probe Segments and Linear Referencing System (LRS): A Tale of Two Giants and How Ohio Made Them Friends

Q: **Stan Young (The Eastern Transportation Coalition):** Are there any metrics - voting logic, matching logic, or screening for questionable matches?

A: Bill Welch (Ohio DOT): None that are formalized, but we check in an ad hoc process. We have a general sense of where this method may not work as well. As long as we get most of our highways and major state routes that's our 'good enough' point.

Q: **Tom Edinger: (DVRPC):** How much QA/QC is still involved in cleanup work to update correct traffic volumes from one segment to another?

A: Bill Welch (Ohio DOT): Volumes are less important overall for us. We were making sure just the segments were matched up. The QA/QC required takes a few hours.

Q: **James Li (MWCOG):** I was wondering whether XD data or TMC data (or both) was conflated LRS in your presentation. The INRIX very short segment sample you showed seems like a TMC segment, was it?

A: Bill Welch (Ohio DOT): Only XD data – we stopped using TMC data once we got XD data because of the increased granularity.

Q: **Paul Bushore (MARC):** Did you find any segments on either side that completely fell out, what did you do with that? Does the Ohio LRS go down to local streets?

A: Bill Welch (Ohio DOT): We check somewhat closely the ones that don't match (almost always it's a segment that we don't have a match in our LRS). The LRS goes down to local streets but not all of them (mostly omitting ones in urban areas). We check a sample of them and then call it 'good enough' – we have some weird cases, like upper/lower road configurations in Cleveland that result in some issues and we just manually check those.

Q: **Stan Young (The Eastern Transportation Coalition):** Is this a once-a-year process?

A: Bill Welch (Ohio DOT): We do this twice a year because our timing is tied to INRIX map releases and we'll conflate whenever a new map is released.

Conflation Software and Its Methodology in Real World Applications

Q: **James Li (MWCOG):** What ground truth was used to reach the 88%+ accuracy? How was the QA/QC was processed?

A: Fred Hejazi (CitygateGIS): The Army assessed the accuracy. They had a manually conflated dataset and then compared it to our results. When we run our process, we generate a percent match – features that have a 1:1 match we know we're confident in and others we're less confident. If you want to do conflation from a process perspective, you get the best results when you look at the network, not just proximity. You identify the major roads and their major intersections and when you have that, you can start the matching process for other roads. Next, you can work out the secondary roads. That seems to be the way we've had the most success conflating large datasets.

Q: **Paul Bushore (MARC):** Has anyone tried using sources such as OpenStreetMap as a base for their LRS? It seems some data providers are using that to use to report some data. It comes with its issues and there isn't command and control over the segmentation so maybe it is a non-starter for people.

A: Fred Hejazi (CitygateGIS): One of the disadvantages we had was working in a classified environment was that we didn't see where the source material was coming from. It could have been OSM or any number of places. We would do simple processing. We haven't done anything with OSM on the civilian side.

A: Steve Gault (Pennsylvania DOT): No, not in Pennsylvania at this point.

A: Bill Welch (Ohio DOT): No, but we have thought about it. INRIX has released some open LR data that works with OSM, but it seems like we'd have to conflate back to our LRS anyway so we're not sure it would save us time.

Q: **Stan Young (The Eastern Transportation Coalition):** Are there generalized standards for mapping-based that are more portable?

A: Fred Hejazi (CitygateGIS): I haven't come across anything. Most of what we have seen is that in DOTs and local municipalities there are five different departments using five sets of data. It would be great if there was a standard.