Integrating Least Squares Analysis with GIS for Cadastral Data Quality Enhancements

Tim HODSON and Christine LESLIE, United States

Key words: land records, boundary network, cadastral measurements, parcel fabric

New technology computing power, processing, services-oriented-architecture,

SUMMARY

The newest generation of the parcel fabric is built from advanced elements of today’s computing, mapping and information systems. When empowered by land information systems such as these, cadastral agencies can benefit from advanced spatial analysis and attain previously unreachable goals. This paper focuses on using least-squares analysis to improve the spatial accuracy of parcel boundaries in a GIS. The paper includes details on the integration of the DynAdjust least squares engine. Other topics include performance, scalability, data modeling, symbology, visualization, user experience, and harnessing the power of a service-oriented-architecture.

1. INTRODUCTION

Starting with ArcGIS Pro version 2.6, released in July 2020, the parcel fabric technology includes a least-squares adjustment engine called *DynAdjust*, from Geoscience Australia. This engine was used for the development of the GDA2020 Australian reference frame and is used by Australia's governmental geodesy departments to maintain their authoritative geodetic network. This least-squares engine became widely available as an open source product in August 2018. (Fraser et al, 2018)

Esri has integrated this technology for use with the parcel fabric and collaborates with the community of developers on the open source platform. The parcel fabric models property boundary networks; these have characteristics that are outside the norm of typical survey measurement networks.

A key goal of Esri’s implementation was to fully leverage the rich and varied abilities of feature layers and the power of map visualizations to represent the output statistics of the least-squares analysis. This is a break from the traditional adjustment engines’ numerical analysis, with the output of textual information.

The functionality is integrated into the Esri geoprocessing tools, and these capabilities are also available as REST API function calls. Thin clients can trigger analysis and adjustments to run on the back-end server. This provides the ability for survey field crews to get analysis results back directly in the field, instead of needing to return to the office to do this analysis.

An important outcome of this development work is that Esri is the world's first in providing a least-squares analysis that leverages a service-oriented architecture as a core part of its GIS platform.

2. PERFORMANCE AND SCALABILITY

2.1 Network Segmentation

One of the challenges of the least-squares adjustment is its limitations for working with very large networks. Networks that contain tens of thousands of measurements may take many hours, or even days to complete. This is due to the computationally intensive nature of calculating the inverse of the least-squares normal equation matrix (Ibid).
In such cases a phased least-squares adjustment can be used to partition a network into smaller blocks. The Dynadjust engine has been developed around algorithmic techniques to automate the segmentation of large networks, making it simple, efficient, robust, and configurable.

The segmentation can be configured and optimized for large or small CPU capacities. Large networks can be run on small memory computers with a basic CPU, albeit with a longer processing time. Supporting a wide spectrum of computer capability and performance allows a configurable scalability that ranges from low-performance desktop computers to national supercomputers.

2.2 Math Kernel Library (MKL)

In addition to the numerical techniques and algorithms used for these performance gains, the Dynadjust engine also uses Intel's Math Kernel Library (MKL). The linear algebra routines and functions of the MKL include matrix inversion processing that is considerably faster than the conventional inversion techniques. This combination gives performance that allows the use of least-squares analysis for more regular re-adjustment of networks, a notion that had previously been considered impractical for many organizations.

3. DATA MODEL

The ArcGIS Pro Parcel Fabric is a controller dataset that includes a set of simple feature classes, a geodatabase *Topology* and *Attribute Rules*.

The *Records* feature class stores the representation of each recording document that brings new parcels into the fabric. These documents

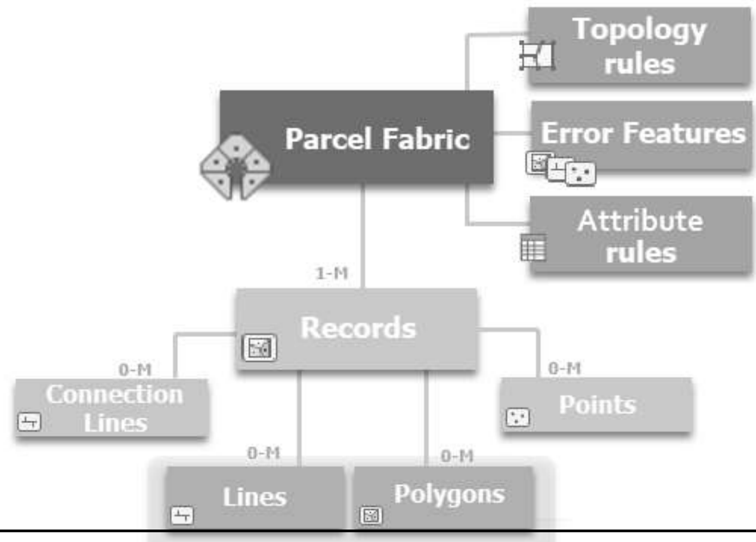

Figure 1. Parcel fabric model.

may represent one or multiple parcels, and the associated transactions also retire the parent parcels. These parent parcels are not removed and contribute to the parcel lineage. Each Record feature has a polygon footprint of the fabric features that it holds.

Multiple parcel types can be added to a fabric. Each parcel type is represented by a pair of feature classes, *Polygons* to store polygon features, and *Lines* to store the line features. The feature class pair is associated with the same parcel type. A parcel type is created with a given name such as 'Tax' or 'Lot', for example. These feature classes have a pre-defined set of attribute fields and they can be extended with additional fields, domains and subtypes.

Parcel boundaries are stored in the *Lines* feature class. The line feature classes in the fabric are automatically *COGO-enabled* for all parcel types. This means that they have additional attribute fields to store values from cadastral documents, such as the directions, distances for straight-line boundaries, or radius and arclengths for circular arc boundaries. It is also possible for these fields to be empty if one or all the values are not on the original document. Most COGO-lines are expected to be two-point, single-segment lines though there are exceptions such as natural boundaries like rivers and shorelines.

The *Connection Lines* feature class is used to store any type of line that is not used as part of a parcel boundary. They store recorded dimensions for lines that connect between parcels or that connect to control points. Some examples are street centerlines, and tie lines.

The *Points* feature class is used to define a network topology. Points have a mapped geometry location stored in the *Shape* field. The attribute fields include *Name*, *X*, *Y* and *Z* fields to optionally store coordinate attributes to complement the mapped geometry. Values in the *AccuracyXY* and *AccuracyZ* fields are used in the least-squares analysis. Points can be set as *fixed* and are used as constraints in a least-squares adjustment. Each parcel corner is represented by a single point feature in the point feature class and connects multiple measured lines in the network. These lines may come from different cadastral records with different recording dates.

4. PARCEL BOUNDARY NETWORKS

Applying a least squares adjustment on large parcel boundary networks with data of varying accuracy has been a technique employed by municipalities in Canada (Horwood et al, 2017) and Australia (Sandy et al, 2014), and by Bureau of Land Management in the USA. (BLM, 2001)
While the Dynadjust engine has support for twenty different measurement types, for the purposes of adjusting parcel boundaries Esri uses the following measurement types: *direction sets*, *distances*, and *geodetic latitudes and longitudes*. These are detailed in the sections that follow. The Dyndjust engine also supports the following types of point coordinate input:

- *Free*—These are regular parcel points. The point shape geometry is updated when the results of the least-squares adjustment are applied to the parcel fabric.

- *Constrained*—The spatial location of the point coordinate is fixed and does not move when the results of the least-squares analysis are applied.
- *Weighted*—The point is weighted by its associated accuracy in the XY Accuracy field. The higher the accuracy, the less movement allowed in the point shape geometry. This point type has a different treatment within the engine compared to the free and constrained points; this is described further in the section headed *Geodetic Latitude and Longitude.*

4.1 Distances

The distances recorded on plans and in deed descriptions in the USA are universally accepted to represent horizontal ground distances. (Brown, et al 1994.) While these values can be scaled to any projection surface, the values themselves are independent of the projection. (Figure 2)

The Dynadjust engine refers to this measurement type as the *direct* distance and is the best match for these parcel boundary values.

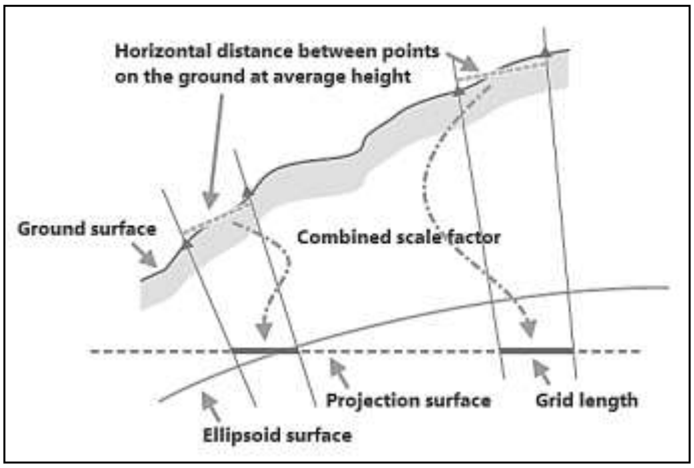

Figure 1. Combined scale factor

4.2 Direction Sets

A *direction set* is composed of an origin point (the from point), a backsight line (reference line), and multiple foresight lines. Direction sets are processed in the least-squares analysis as follows:

Figure 3. Direction set with one derived angle.

- The angles formed by the direction set are the measurements used in the least-squares engine. The angles are derived from the direction values of the backsight and foresight lines.
- In Figure 3, point 3762 is the origin point of the direction set. The backsight or reference direction is the line from point 3762 to point 3186. The foresight direction is the line from point 3762 to 3763.
- Any point in the parcel fabric that has multiple lines connecting to it may have multiple direction sets.
- When there are adjacent cadastral records, two direction sets are created for the same origin point. This is done to account for the possibility of different bases of bearings (rotations) being used for different cadastral records.

4.3 Geodetic Latitude and Longitude

The measurement types *geodetic latitude* and *geodetic longitude* in the Dynadjust engine are used in the parcel fabric model exclusively for specifying weights on parcel fabric points.

The coordinate values stored in the X, Y, and Z fields of weighted points are converted to *geodetic latitude* and *geodetic longitudes* by the Esri projection engine, and input to the DynAdjust least-squares analysis. Only the weighted points are converted to *geodetic latitudes* and *geodetic longitudes* for use as this measurement type. While the constrained and free point coordinates are used in the adjustment, they are not used as a measurement type.

These weighted point coordinates are considered “floating control” and can be assigned different accuracies based on the method of capture. (Fraser, et al, 2019) For example, a point may be captured by using a heads-up digitized position taken from a fence corner visible on orthophoto imagery. (Horwood et al, 2017) This point can be given a lower accuracy estimate and combined in the same adjustment process with the constrained and free points.

5. GEO-PROCESSING, USER EXPERIENCE & ANALYSIS LAYERS

Esri Parcel Fabric customers can use ArcGIS Pro geoprocessing and python scripting tools to schedule and run adjustments of very large parcel fabrics. Two geoprocessing tools are used for, (1) running a least squares analysis, and then (2) applying analysis results to update and improve coordinate positions of parcel fabric points.

Improvements to maintaining dimension attributes of property boundaries is becoming ever more important. Once a re-adjustable network of property lines has been built it serves as a data validation framework; as new property boundary data is added, or edits are made, the adjustment analysis can be re-run to confirm quality of the new data.

5.1 Improved User Experience

Rather than having the analysis and adjustment occur as part of a single tool, as has been the norm in prior GIS-based implementations, the workflow is designed around using an initial discrete step of generating the analysis results and saving these to a separate set of GIS feature classes, using the *Analyze Parcels By Least Squares Adjustment* geoprocessing tool.
This allows the information from the Dynadjust engine to be presented and visualized in separate analysis layers in a form that allows a much closer match with the output of the engine, while also giving Esri's parcel fabric data model a good amount of independence from that of the least-squares analysis results. The additional benefit to this approach is that the output from can be assessed without altering the original data.

Once satisfied with the results a second geoprocessing tool, *Apply Parcel Least Squares Adjustment*, can be used to make the changes to the original parcels.

5.2 Visualizing Analysis Results

The analyze process generates the following feature classes and these are added to the map when the analysis results are created:

- AdjustmentLines—Stores and displays adjusted and statistical data for parcel lines' *direction sets*, *distances*, *geodetic latitudes*, and *geodetic longitudes.*
- AdjustmentPoints—Stores and displays adjusted and statistical data for parcel fabric points and coordinates.
- AdjustmentVectors—Stores and displays the vectors from parcel fabric points to their adjusted coordinate locations.

These layers can leverage the full panoply of ArcGIS Pro software to symbolize, label, and the information across multiple scales. The power of GIS layers rests in their ability to visually present complex attribute and spatial information in a meaningful way that is quickly and easily understood. The adjustment analysis produces a lot of valuable information that can be presented on a parcel map with different symbols and labels depending on the scale of the map. These scale-dependent views update dynamically as the user zooms in and out to varying scales in different map extents.

For example, a visual heat-map of accuracy for a large area of a city offers a powerful tool to help understand where additional measurement data or control points are needed.

The example in Figure 4 was created using the *AdjustmentPoints* layer and symbolizing on the XY Uncertainty field by graduated colors and size using Natural breaks method with six classes.

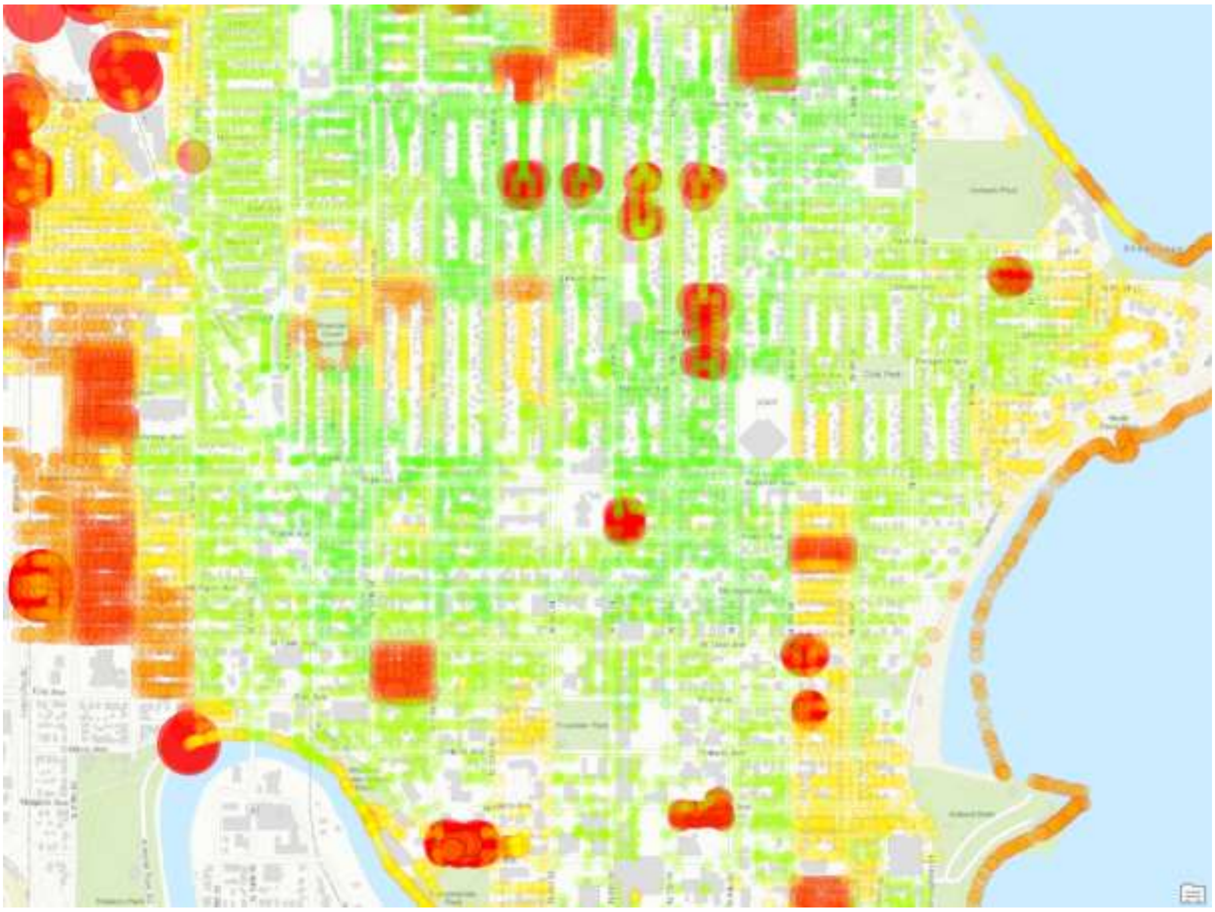

Figure 4. Point symbols for XY Uncertainty.

The example in Figure 5 shows the default symbology of the *AdjustmentLines* layer. In this case the outliers can be easily picked out on the map from the bold red and orange colors, and the relative thickness of the lines.

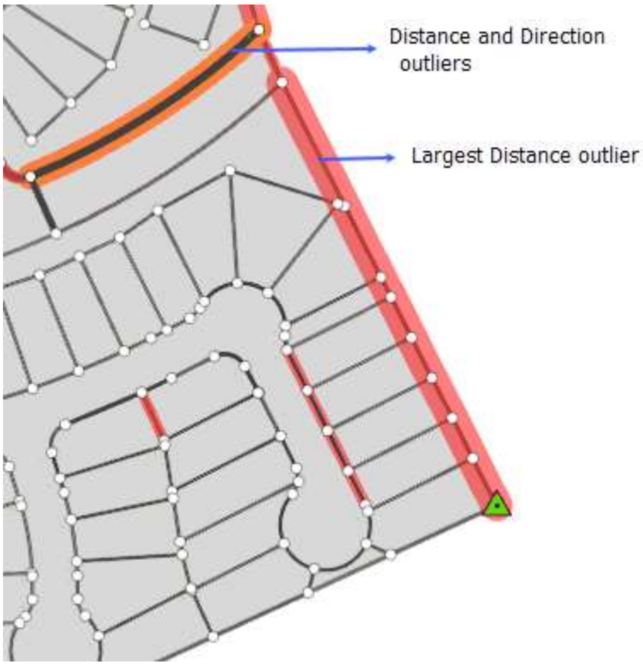

Figure 5. Outliers using the Student's t Statistic value.

The thickness of the line symbols for these outliers are based on the information in the analysis field *StudentStatistic*, showing the largest distance outlier as the thickest red line. Lines with larger Standardized Student's t Statistic attributes are deviating more than expected from the best-fit solution computed by the least-squares adjustment. (Esri, Documentation Reference 2)

Figure 6 shows the map at a large scale when zoomed in to a point's location. This analysis view presents information from the *AdjustmentVectors* layer, showing a single vector line, labeled with the vector's length, and also the *AdjustmentPoints* layer, showing the error ellipse created from the values in the *ErrorEllipseSemiMajor* and *ErrorEllipseSemiMinor* fields.

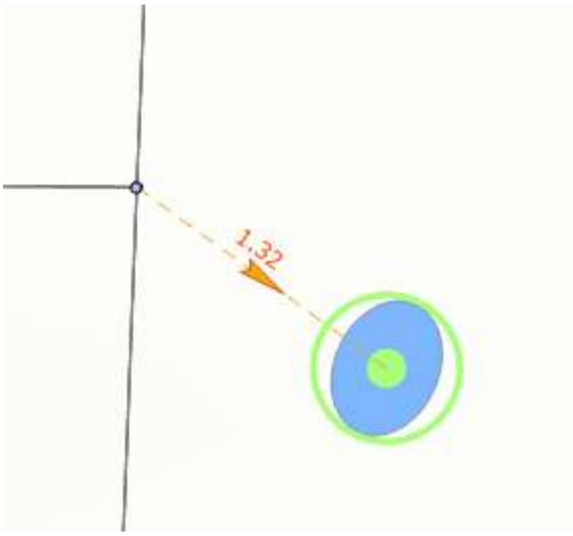

Figure 6. Error ellipse and vector with labeled length..

These images of analysis layer results touch on a few examples of contextual, scale-relevant visualization of information about accuracy at a range of map extents from the entire dataset, to subdivisions, to individual parcels, to much larger scales showing individual parcel corner point features. ()

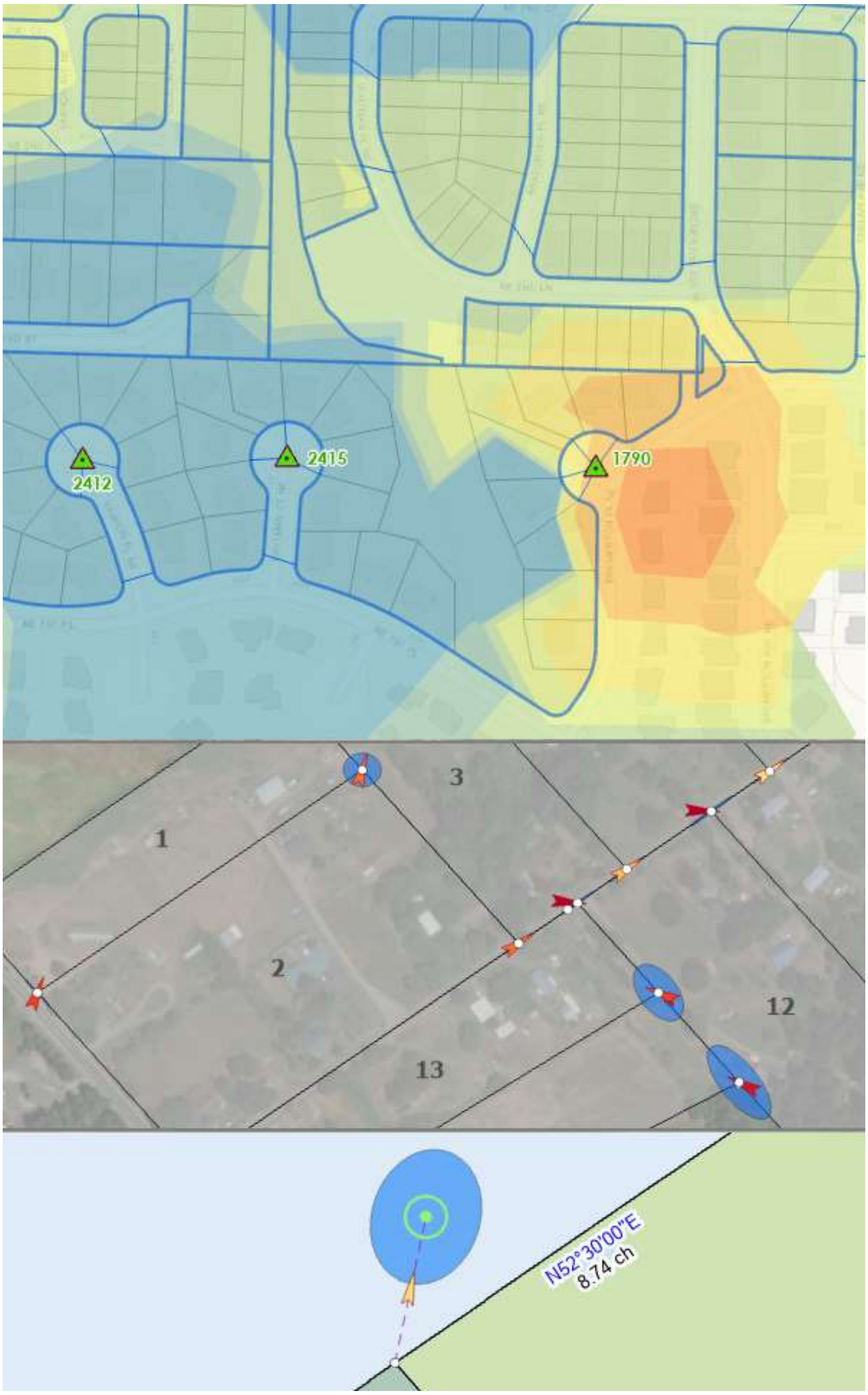

Figure 7. Visualizing least squares analysis results across different map scales.

6. SERVICE-ORIENTED ARCHITECTURE

ArcGIS Pro is designed to easily and seamlessly work in a multi-user environment that allows users to focus on the features and interact with the feature layers in the map, without needing to know how and where the data is hosted. The land records professionals who manage the data through ArcGIS Pro sign in to a portal connection, and then connect to a parcel fabric feature service that is managed in a service oriented architecture. (SOA) The user creates a *version*, performs their edits, and then reconciles and posts their changes to make the updates through these services.

Although ArcGIS Pro is the primary client of these services, they are also avilable to thin clients allowing immediate access to the system through mobile devices in the field. New measurements and other relevant cadastral data can be captured by field crews and assessed in real time, saving repeated return trips to the site. Where the infratructure does not support a real-time environment, or business requirements preclude it, the data can be captured in the field and synchronized later. (Bar-Maor, 2021 , Brinkman, 2021)

The *branch versioning* model used in this feature service editing model allows the dynamic changes to the fabric coordinate positions to be recovered from any previous *moment* in time; this means that the previous and current positions of fabric points can be used to maintain spatial relationships with features in other non-parcel layers, such as utilities that are located at known offsets from boundary lines. (Ibid)

REFERENCES

Bar-Maor, A., 2021, *Parcel Fabric 2.0 - Fit for (Multi) Purpose*, FIG e-Working Week 2021.
Brinkman, T. C., 2021, *Cadastral system modernization: the technology and business requirements that drive the next wave of disruption*, FIG e-Working Week 2021.
Elfick, M.H., 2003, *Building & Managing a Survey-Accurate Cadastre*, Survey Summit, Twenty Third Annual Esri User Conference, San Diego.
Esri, Documentation Reference 1:
https://pro.arcgis.com/en/pro-app/latest/help/data/parcel-editing/whatisparcelfabric.htm
Esri, Documentation Reference 2:
https://pro.arcgis.com/en/pro-app/latest/help/data/parcel-editing/analyzeadjustmentresults.htm
Fraser R., Leahy F., Collier P., 2018, *DynAdjust Users Guide*, ICSM open source: github.com/icsm-au/DynAdjust (Geoscience Australia, www.ga.gov.au)
Horwood, D.M., 2012, *Mapping and Coordinates, Ontario Professional Surveyor*, Winter 2012
Horwood, D.M., Chan, A., 2017, *Estimating Accuracy of a City Wide Parcel Fabric*, Thirty Seventh Annual ESRI User Conference, San Diego. (Session Id: 1866_315)

CONTACTS

Tim Hodson
Environmental Systems Research Institute
Address 111 Market Street NW, Suite 250
City Olympia
COUNTRY USA
Tel. +1 360 754 4727
Email: thodson@esri.com
Web site: www.esri.com

Christine Leslie
Environmental Systems Research Institute
Address 380 New York Street
City Redlands
COUNTRY USA
Tel. +1 909 793 2853
Email: cleslie@esri.com
Web site: www.esri.com