

CONVERSION OF PLSS COORDINATES USING AUTOMATICALLY
GENERATED REGRESSION DOMAINS

A Thesis

by

MICHAEL ANDREW AUSTIN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Marian Eriksson
Committee Members,	Andrew Klein
	Sorin Popescu

Head of Department,	Kathleen Kavanagh
---------------------	-------------------

December 2016

Major Subject: Ecosystem Science and Management

Copyright 2016 Michael Austin

ABSTRACT

The Public Land Survey System (PLSS) serves as a means of legally identifying land in 30 out of 50 U.S. states. PLSS coordinates are frequently used to describe spatial data such as the locations of wildfires, but PLSS coordinates are largely incompatible with most software for spatial analysis. One means of translating PLSS coordinates into longitude and latitude had been a regression-based tool known as TRS2LL which was developed by Martin Wefald, but the model on which this tool is based only supported 17 out of 30 PLSS states and the geographic boundaries of the regions for which each regression was applied to (regression domains) had been delineated by hand on paper maps. We present a new model known as PLSS2LL based loosely on TRS2LL in which regression domain boundaries are procedurally generated via GIS and the coverage is extended to all 30 PLSS states. We observed an improvement in PLSS2LL's accuracy (mean error 133.62 m) in predicting longitude and latitude coordinates over its predecessor TRS2LL (mean error 220.58 m). While more accurate, the resulting domains are more fragmented with an average of about 280 domains for each of 30 states, 8,439 total as compared to Wefald's average of 105 for each of 17 states. Approximately 59.60% of PLSS2LL's predictions fell within 100 m, versus the 29.41% of TRS2LL's. The inverse conversions in which longitude and latitude are used to predict PLSS coordinates were also tested. For predicting PLSS coordinates, PLSS2LL and TRS2LL respectively yielded accuracies of 90.33% and 84.48%. Our model

effectively reduced the original data from the U.S. Bureau of Land management by factor of 143.75. We also explored the various sources of error and considered future improvements.

ACKNOWLEDGEMENTS

I would like to thank my committee chair Dr. Eriksson for my introduction to this research topic, and for her support and direction throughout the course of this research. I would also like to thank the other members of my committee Dr. Popescu and Dr. Klein for their interest in this work. And finally thanks to my father and mother for their supportiveness over the years. Were it not for them this research would not have been possible.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professor Dr. Marian Eriksson, Dr. Sorin Popescu of the Department of Ecosystem Sciences and Management and Dr. Andrew Klein of the Department of Geography. All work conducted for the thesis was completed by the student.

Funding Sources

There are no outside funding contributions to acknowledge related to the research nor the compilation of this document.

TABLE OF CONTENTS

	Page
1. INTRODUCTION.....	1
2. LITERATURE REVIEW	3
2.1 History of PLSS.....	3
2.2 What PLSS is.....	4
2.3 How Wefald's TRS2LL works.....	7
2.4 Examples of PLSS data in geospatial research.....	10
2.5 Examples of Wefald's TRS2LL in use.....	11
2.6 Examples of other implementations of PLSS converters.....	12
2.7 Tools for performing ellipsoidal trigonometry.....	13
2.8 Research objective.....	13
2.9 Contributions.....	14
3. METHODS.....	16
3.1 Data sources.....	16
3.2 Overview of analysis technique.....	20
3.3 Correcting, reformatting, and splitting datasets.....	22
3.4 Corner identification technique.....	25
3.5 Automation of regression domain generation.....	28
3.6 Exporting domains and corners as tabular data.....	36
3.7 Predicting latitude and longitude by regression.....	37
3.8 Sampling methodology for accuracy assessment.....	41
3.9 Predicting PLSS from longitude and latitude.....	42
4. RESULTS.....	44
4.1 Overview.....	45
4.2 Sampling for accuracy assessment.....	47
4.3 Magnitudes of errors (PLSS2LL versus TRS2LL).....	48
4.4 Directional bias of errors (PLSS2LL versus TRS2LL).....	49
4.5 Inspection of individual failures in PLSS2LL predictions.....	50
4.6 Latitude and longitude to PLSS accuracy assessment.....	59

5. SUMMARY	61
5.1 Concluding assessment.....	61
5.2 Future research.....	62
REFERENCES	63

LIST OF FIGURES

	Page
Figure 2.1: Map of lands surveyed under the Public Land Survey System (PLSS), colored by meridian.	5
Figure 2.2: Demonstration of the hierarchy of basic PLSS features, and how a Cartesian-like “plane” can be identified with the PLSS coordinates. This coordinate can be identified as “E1/2 SE1/4 S.1, T3N, R4E, Indian”.....	7
Figure 2.3: An example of how township misalignments can define the boundaries of regression domains. Township boundaries are in blue and domain boundaries are in red. The vertical domain boundary line is due to the left side being part of a different meridian than the right side. The horizontal domain boundary lines in the interior are due to latitudinal corrections.....	8
Figure 2.4: An example of PLSS irregularities in northeastern Oklahoma that may have been the result of surveying over a large body of water.....	9
Figure 2.5: An example of PLSS irregularities in California that may have been the result of surveying over mountainous terrain.....	10
Figure 3.1: Depicted above is a diagram of how the CadNSDI formats for township (PLSSID) and section (FRSTDIVID) ID strings correlate with the alternate source dataset’s format for townships (Indkey) and sections (mtrs). Note that the latter lacks space for storing duplicate township codes.....	17
Figure 3.2: Comparison of PLSS data from GeoCommunicator (orange) and National Atlas (blue) overlaid atop USGS topographic maps (PLSS features are red). National Atlas features occasionally exhibit large gaps in coverage (top-left) that GeoCommunicator correctly covers. Other times GeoCommunicator correctly omits parts of a township that USGS omits while National Atlas over-extends the township in an attempt to meet nominal PLSS criteria (top-right). In most cases GeoCommunicator has townships that align more closely to topographic maps than National Atlas. Occasionally GeoCommunicator’s townships align less closely to topographic maps due to obvious errors (bottom-right)	19

Figure 3.3: This diagram demonstrates a portion of the overall model, specifically the portion responsible for Oklahoma. There are 29 other states are unrepresented here. Note that the state of Oklahoma contains just two meridians: The Cimarron meridian (meridian code #11) and the Indian meridian (meridian code #17). For a state with more than two meridians, the first process produces more workflow paths. The dashed lines indicate optional inputs. In the final build of the model, these optional arguments were omitted so that only section level corner data was used in regressions, due to it being more granular and thus less sensitive to errors caused by truncated features. The states of Kansas, Nebraska, and Wisconsin utilized a slightly different model that omits any township data, and inputs section corner data into the domain identification.....	21
Figure 3.4: The above are examples of irregular, but not erroneous, section numberings exhibited by townships in Mississippi. While section numberings typically consist of values 1 to 36, we see values including 43, 52, and 35 here. Such section features can not be predicted by our model....	23
Figure 3.5: Example of the a coastal township who's corners are detected by the distance-to-bounding-box-corners method. Circular nodes represent identified corners, and dashed lines indicate the bounding box. Note that the southwest corner was not detected because no vertex of the township was within 200 meters.....	27
Figure 3.6: Above in the highlighted region is an example of irregular section geometry. Despite being having well aligned corners with their neighbors, these features can impair the quality of regressions if not encapsulated within a regression domain of their own.....	29
Figure 3.7: Depicted here are 9 steps in a domain expansion (A through I). Primary expansions are attempted in the order of East (blue), West (green), North (red), and then South (yellow). Black indicates the initial township from which expansion began. Note that between steps G and H a primary expansion northward was attempted, but halted by misalignments of neighbors along the north edge (the townships in white), so the next successful expansion is southward. The final graphic (J) demonstrates in detail which of its neighbor's corners came into play when testing adjacent alignments in the final secondary expansion.....	31

Figure 3.8: These are simplified examples of the geometry tested for by the various alignment tests within the domain expansion algorithm. The most common examples of a well-aligned pair of adjacent townships that permit domain expansion to pass (A) as well as poorly aligned pair of adjacent townships that would halt domain expansion (B) are demonstrated here. In these examples, all well defined township corners are pictured as circular nodes on the polygon. An example of the slice test (C) demonstrates how a misalignment can be determined from the outer corners of both townships when corners along the adjacent edge don't exist. Also present are two examples of the shear test (D), which detects internal misalignments within individual townships. And lastly, there is the case (E) of a pair of townships whose abrupt change in relative size would halt a domain expansion from one to the other.....	33
Figure 3.9: Depicted above is an example of a township at 3 North and 4 East, it's constituent sections, and a mapping of decimal range and decimal tiers to its features. The decimal range and decimal tier serves as useful intermediate coordinates that facilitate the prediction of latitude and longitude by transforming integer range, tier, and section values into a Cartesian coordinate.....	38
Figure 4.1: Regions in green indicate coverage for TRS2LL. The domain coverage of PLSS2LL encompasses regions in purple and green.....	44
Figure 4.2: Histograms of the distance from sections' centroids to the predictions of both PLSS2LL (A) and TRS2LL (B) reveal substantive differences the error distributions of each model. The error of one TRS2LL sample was in excess of 1,500 meters.....	49
Figure 4.3: Depicted above are the azimuths of geodesics calculated from the centroid of randomly sampled sections to the predictions of PLSS2LL (A) and TRS2LL (B).....	50
Figure 4.4: Blue lines indicate domain boundaries, the cyan selection in the middle indicates the randomly sampled section. And the highlighted region on the right indicates the column of townships with 5x6 section grids that skewed the predictive accuracy of PLSS2LL.....	51
Figure 4.5: Demonstrated here is a case of PLSS features with local horizontal shearing. This negatively impacts the accuracy of predictions under both PLSS2LL as well as TRS2LL. It is not uncommon within mountainous terrains that were difficult to survey.....	53

Figure 4.6: Depicted above is a map of all PLSS2LL domains, classified according to the value of either their ordinary least squares regression’s longitudinal R2 values or latitudinal R2 value (whichever is lower). Note that in Arkansas and Missouri several domains with slight rotations suffered a markedly reduced lowest R2 value relative to their neighbors. Also indicated here are domains with nonviable regressions due to too many missing corners, shown in red. These are typically due to townships or sections with geometry so thin that corner detection filtered out points in the domain due to their irregular dimensions. Future builds of PLSS2LL could use revised threshold values for handling features with irregular dimensions to be more forgiving so that enough corners are kept to maintain a viable regression in those areas..... 55

Figure 4.7: Alaskan domains generated by PLSS2LL. Note that Alaska’s domains are exceptionally simple in structure. This is likely due to the original data being free of surveying errors..... 57

Figure 4.8: Regression domains generated by PLSS2LL are indicated in blue, and TRS2LL domains are in red. Depicted here are California (A), North Dakota (B) and South Dakota (C). PLSS2LL domains tended to be more complex than those utilized by TRS2LL, but the former’s domains were automatically generated while the latter's were crafted manually. Mountainous regions tended to induce additional domain fragmentation in PLSS2LL, as seen in California. The regression domains between the two models were largely in agreement in North Dakota, but in South Dakota the triangular region in the North East required TRS2LL to implement many thin and long domains to accommodate it. PLSS2LL populated the triangular region with shorter and more compact domains along the triangle’s edges..... 58

LIST OF TABLES

	Page
Table 4.1: The table below presents the frequency of all fractional and duplicate townships per state. While domains composed of such townships posed a challenge for accurate predictions they remain uncommon features relative to the size of the dataset.....	46

1. INTRODUCTION

Following the American Revolution and the creation of the government, lands in territories seeking statehood needed to be surveyed in order to transfer ownership from the public domain into the private domain and, in those early years, to pay the debts incurred by the Revolutionary War (White, 1926, p. 11). In response to this need the United States Public Land Survey System was created, often abbreviated as USPLSS or as just PLSS and also commonly referred to as the Rectangular Survey System. Most of the United States west of the Mississippi River was surveyed by PLSS (Avery & Burkhart, 1994, p. 46). It has been and still continues to be the legal means for parcel identification in 30 out of 50 states (Committee on Integrated Land Data Mapping, 1982, p. 9). Lands not originally surveyed by PLSS were not resurveyed to accommodate it. Despite the availability and affordability of commercial GPS some modern data supplied by government agencies still relies on PLSS coordinates rather than latitude-longitude coordinates. Additionally, there is always going to be historical PLSS coordinate data that may need to be converted well into the future.

In order to satisfy the USDA National Forest Service's need for directing fire dispatch personnel based on PLSS locations, a technique for PLSS to latitude-longitude coordinate conversion was developed by Martin Wefald and it involves using the coordinates of the corners of PLSS features to construct regressions (Martin Wefald, personal communication, October 31, 2012). Because of errors in the original surveys, gross errors, errors arising from steep terrain, meandering around water bodies, etc., and

corrections made for the curvature of the Earth, the regressions were restricted to relatively small areas which Wefald referred to as “factors” and which will be referred to here as “regression domains”. Wefald fitted two regressions, one for latitude and one for longitude, for each of his 1,798 domains.

This project’s core goal was to develop an improved implementation of Wefald’s model for translating PLSS coordinates into latitude and longitude. Intended improvements included comprehensive coverage for all PLSS territories with the exception of certain Ohioan meridians (as they were the incipient land surveys and harbor irreconcilable peculiarities as compared to later standards), as well as complete automation of the generation of domain boundaries, removing the potential for human error and overgeneralization in the model. Automation also allows the model to be consistently rebuilt when BLM eventually publishes revised PLSS datasets in the future. We also investigated differences in accuracy between Wefald’s TRS2LL as well as their differences in domain geometry complexity.

2. LITERATURE REVIEW

2.1 History of PLSS

Prior to the development of the PLSS the United States utilized a simple method known as “metes and bounds” to define land ownership and property boundaries were often poorly delineated by temporary landmarks such as trees (Avery & Burkhart, 1994, p. 46). The geometry that these lands were divided into were often arbitrary and irregular. At the time, the new government was deeply in debt to France as well as many other creditors yet had no authority to levy federal taxes under the Articles of Confederation so the selling of new land was a much needed source of revenue (White, 1926, p. 11). In 1785, soon after the revolution Congress enacted the first law governing public land surveying in what would later become Ohio (Avery & Burkhart, 1994, p. 46). The first PLSS survey itself was executed under the leadership of the first appointed Geographer of the United States, Thomas Hutchins (Neunzert, 2011, p. 82). The only instruments at the disposal of the first surveyors were a sextant, hand compasses, and circumferentors (White, 1926, p. 18). The Land Ordinance of 1785 specified land would be divided up into *townships*, with each one exactly 6 miles by 6 miles, and further subdivided into 36 *sections*, where each section was 1 square mile. Each collection of townships is defined by a carefully surveyed starting point called the *initial point*. Later Congress modified the surveying rules to allow flexibility for minor deviations from the originally strict size specifications in order to accommodate issues caused by the Earth’s curvature (Avery & Burkhart, 1994, p. 46). Despite PLSS being

an improvement over metes and bounds, poor practices sometimes persisted, particularly in earlier surveys (White, 1926, p. 19). Any inaccuracies in surveying for PLSS are problematic since the surveys' original markers are legally binding regardless of whatever flaws they may possess. This is done in order to avoid litigation and resurveying (Avery & Burkhart, 1994, p. 46). Unfortunately, many of the original markers are lost or obliterated. An obliterated marker is a marker that has been physically destroyed. Despite obliteration, its location may be recoverable by the memory of those familiar with it or by acts of the landowners. If the location cannot be recovered by any means of external evidence then the obliterated marker is also considered to be a lost marker (White, 1926, p. 684). According to the Department of the Interior a failure to replace markers in a way that is unchanging from the original survey was a frequent problem and is an injustice to landowners (White, 1926, p. 684). The Department of the Interior received many letters over the years indicating a growing problem of lost or obliterated markers and in response the willful defacing of any original survey marker became a federal crime punishable by a fine of up to \$250 and up to 6 months of imprisonment (White, 1926, p. 754). Thus far there appears to be no evidence of a *comprehensive* effort to resurvey or reconstruct extinct markers.

2.2 What PLSS is

A territory surveyed by PLSS was first based on a carefully surveyed principle meridian line and baseline. The principle meridian is a line of longitude and the baseline is a line of latitude. Baselines were formerly known as geographer's lines (Neunzert, 2011, p. 82). Guide meridians are similar to principle meridians since they are a line of

longitude and are marked by monuments but guide meridians are not named and appear to be only used for reference by surveyors since there is little discussion of them. Closely related to the concept of a principle meridian is another term, simply just *meridian*, which instead of referring to a line refers to a collective set of all townships that are based on a common principle meridian. The intersection between the principle meridian and the baseline defines the initial point, which can be thought of as the origin of a meridian's coordinate system. Initial points are also sometimes referred to as principle points (Neunzert, 2011, p. 14). Many meridians do happen to cross state boundary lines (Figure 2.1).

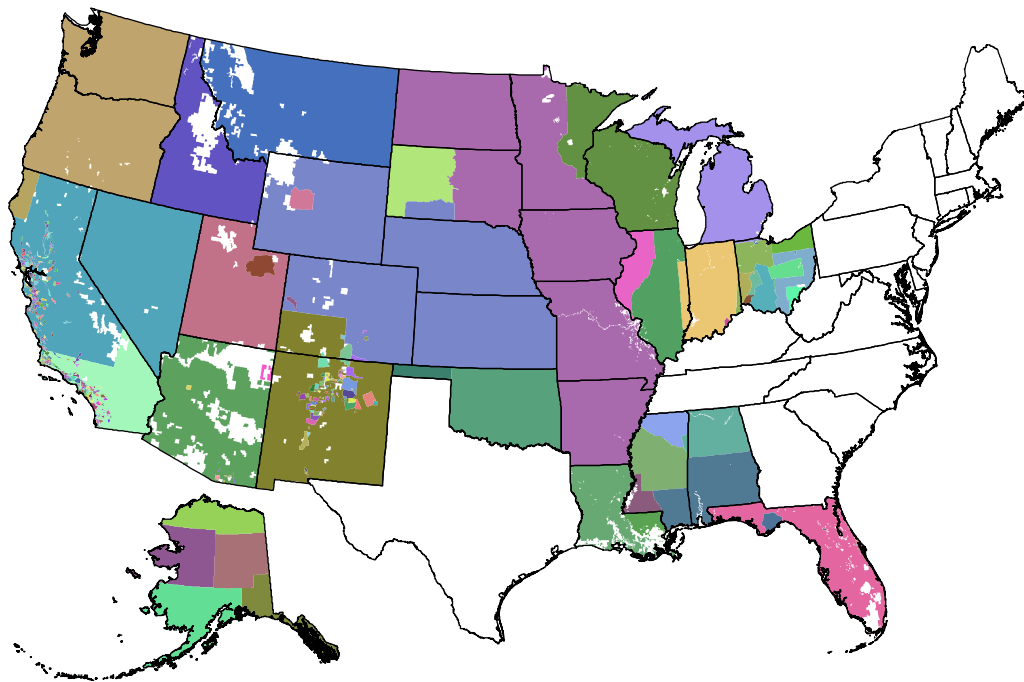


Figure 2.1: Map of lands surveyed under the Public Land Survey System (PLSS), colored by meridian.

A township is nominally a 6 mile square. In practice no township is exactly a square due to the curvature of the Earth, to limitations in survey accuracies, accommodations made to antecedent land surveys, as well as flagrant errors. A *quadrangle*, also known as a *24-mile tract*, is nominally a 24 mile square and defines the bounds of up to 16 townships (Neunzert, 2011, p. 18). Quadrangles were usually surveyed and marked with monuments before individual townships are surveyed and marked. Quadrangles at the same latitude are typically well aligned at their corners, but their corners will often not align with quadrangles at higher or lower latitudes due to adjustments that are necessary in order to account for the Earth's curvature. Townships were further subdivided into 36 sections (Figure 2.2), each approximately one mile square. Further subdivision is possible with binary fractions of sections (Figure 2.2). Sections are identified by a single number between 1 through 36 boustrophedonically, with the first section being at the northeast corner and values proceeding westward, then alternating eastward and westward with each lower row of sections (Figure 2.2). Townships are identified by two integers: Range number and a township number, but to avoid confusion with the previously discussed townships (6 mile squares) the township coordinates will be simply referred to here as tiers. A range is a measure of distance from the principle meridian and tiers are a measure of distance from the baseline. Range coordinates are identified by an east or west identifier and tier coordinates must be given with a South or North identifier.

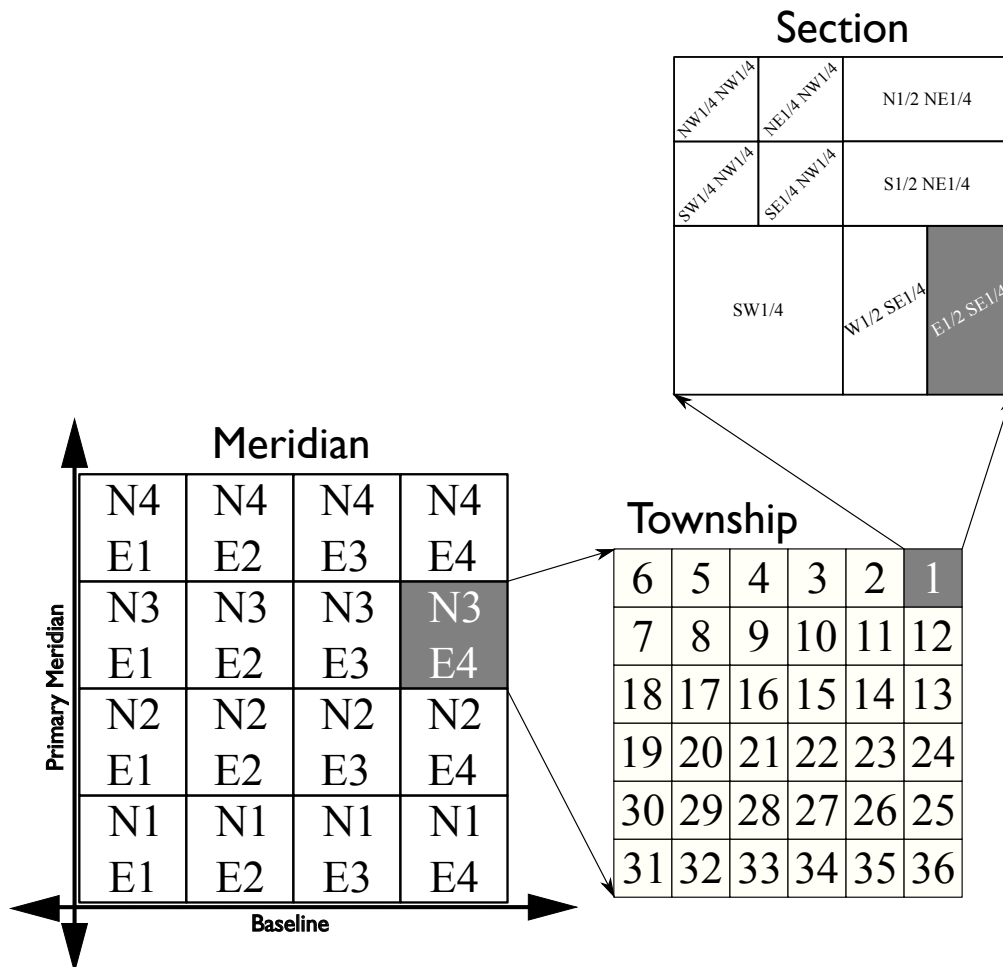


Figure 2.2: Demonstration of the hierarchy of basic PLSS features, and how a Cartesian-like “plane” can be identified with the PLSS coordinates. This coordinate can be identified as “E1/2 SE1/4 S.1, T3N, R4E, Indian”.

2.3 How Wefald’s TRS2LL works

Wefald’s method relies on the subdivision of meridians into rectangular aggregations of townships called regression domains. Wefald would then hand-select the corners from 3 or more sections and use their latitude and longitude as samples to form a regression to predict the values for all other townships within the domain (Wefald, 2009). The accuracy of the prediction is dependent on how regular the

dimensions of the townships are and how well aligned the edges are with cardinal directions. Wefald's implementation supported up to 17 states out of the possible 30 states that use PLSS and was based on maps using the NAD27 datum (Wefald, 2009).

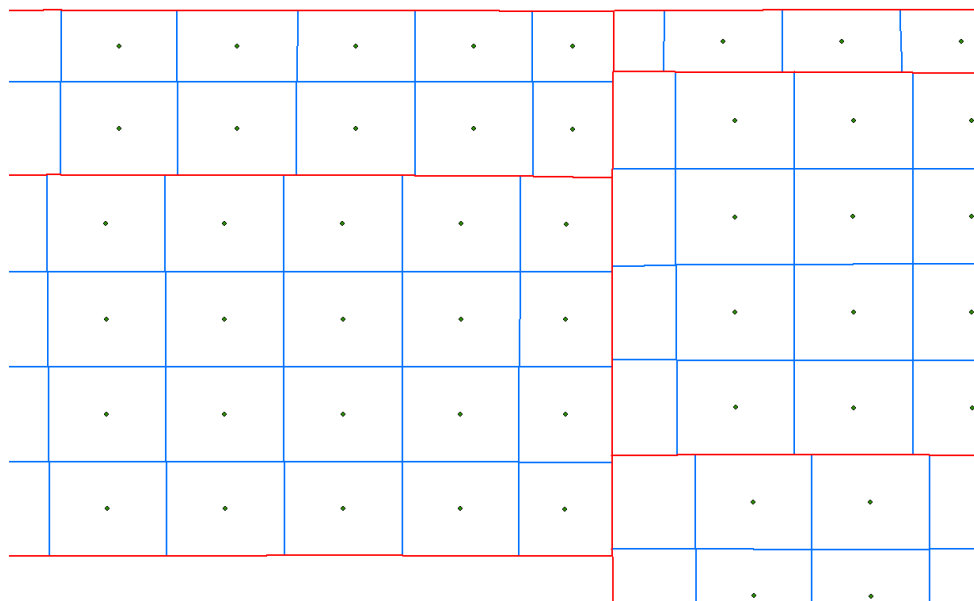


Figure 2.3: An example of how township misalignments can define the boundaries of regression domains. Township boundaries are in blue and domain boundaries are in red. The vertical domain boundary line is due to the left side being part of a different meridian than the right side. The horizontal domain boundary lines in the interior are due to latitudinal corrections.

Barriers in the regression domains can be due to where two different meridians contact with one another, deliberate choices made by surveyors to make a latitudinal correction for the curvature of the Earth (Figure 2.3), inaccuracies caused by large bodies of water (Figure 2.4), or may be simply the result of accidental errors in

surveying mountainous terrain (Figure 2.5). Regardless of the cause it leads to a barrier in how far regression domains can extent outward.

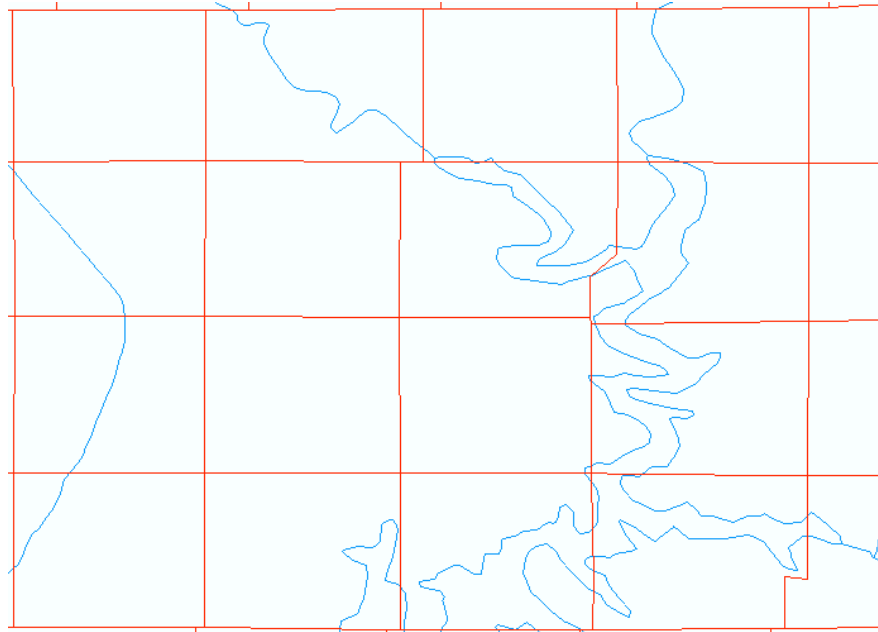


Figure 2.4: An example of PLSS irregularities in northeastern Oklahoma that may have been the result of surveying over a large body of water.

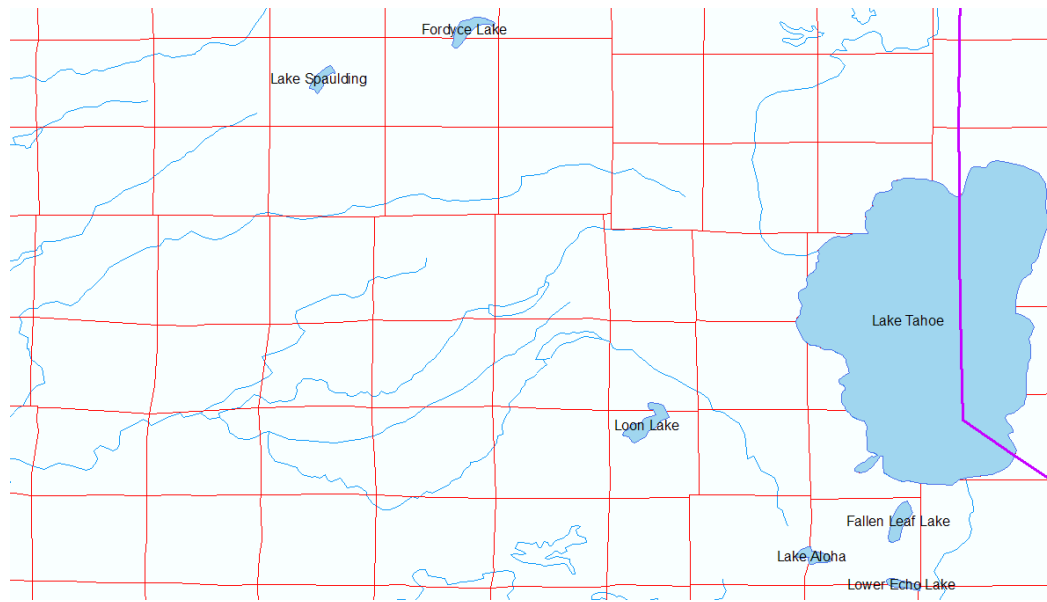


Figure 2.5: An example of PLSS irregularities in California that may have been the result of surveying over mountainous terrain.

2.4 Examples of PLSS data in geospatial research

While longitude, latitude, or UTM coordinates are preferable in geospatial research PLSS coordinates remain relevant as not all useful spatial data is necessarily recorded via UTM or longitude and latitude. Some examples of this include research concerning the prevalence of diseases in *Odocoileus virginianus* (white tailed deer). In one case, hunters were interviewed to ascertain locations of where deer had been fed since such supplemental feeding was a potential risk factor with respect to the prevalence of bovine tuberculosis (Miller et al., 2003). Many of the hunters they interviewed were unwilling to supply the locations of their hunting sights down to a section level, so they sufficed with just range and tier level data. In another case that examined chronic wasting disease in white tailed deer the geographic locations of deer

killed by vehicles along roadsides and submitted by hunters were collected in terms of range, tier, and section (Joly et al., 2003). And in order to better understand the risks *Canis lupus* posed to livestock in Wisconsin and Minnesota, range and tier locations of attacks on livestock have been collected and then used to model risk at a township level (Treves et al., 2004). In an instance of archeological research that examined the prehistoric hunting of bison the two most common formats for geographic locations of archeological sites were UTM and PLSS coordinates (Cooper, 2008). Wildfire databases also frequently rely on PLSS positions (Short, 2014), and it has been found that PLSS positions fire databases are potentially more accurately reported than longitude and latitude positions despite GPS being a more precise means of measurement.

2.5 Examples of Wefald's TRS2LL in use

Many states' records use PLSS coordinates to indicate locations outside of the context of legal land ownership and reaching into records of flora, fauna, and natural resources. The developers of HerbariaViz (Auer et al., 2011), a web interface for vegetation data, relied on Wefald's TRS2LL in handing data from the Consortium of California Herbaria (CCH). Much of the raw data from CCH is obtained from Californian counties that used PLSS and currently the CCH is still attempting to provide latitude-longitude coordinate data for as many specimens as possible (Moe, 2007). TRS2LL was also used in another instance to assist with the georeferencing of 702 specimens of grasses from five different herbaria, few of which had any georeferencing at all (Barkworth, Anderton, McGrew, & Giblin, 2006, p. 236).

The National Forest Genetic Electrophoresis Laboratory (NFGEL) which operates under the National Forest System records the locations of fauna samples in terms of tier, range, and section and Wefald's TRS2LL has been used to translate NFGEL coordinates to latitude-longitude in the study of the *Lewisia kelloggii* flower (Wilson, Hipkins, Rey-Vizgirdas, & Kaye, 2005, p. 346). Another example of TRS2LL being utilized is in the records of monarch butterflies from various sources such as museums and private collections. Sometimes these sources have recorded coordinates in a tier-range-section format rather than just strictly latitude-longitude, necessitating the use of Wefald's TRS2LL (Dingle, Zalucki, Rochester, & Armijo-Prewitt, 2005, p. 492). The Montana Department of Natural Resources and Conservation maintains locations of agricultural fields and the points of water diversion in terms of county, range, tier, section and quarter-sections and TRS2LL has been used to convert this to latitude-longitude (Aadland & Kolpin, 2004, p. 503).

2.6 Examples of other implementations of PLSS converters

Wefald's TRS2LL has not been the only available means of automated tier, range, and section conversion. One of the earliest converters was specific to just Kansas and converted range, tier, and section notations to Cartesian coordinates (Good, 1964). Good determined his mean horizontal and mean vertical accuracies to respectively be 719 meters and 270 meters.

More recent examples include MapTech and BioGeomancer (Auer et al., 2011, p. 97). MapTech is proprietary software suite (MyTopo, 2009). BioGeomancer is an open source project (BioGeomancer Working Group, 2007). Scantek Systems has also

provided an online proprietary converter that charges per-conversion (Scantek Systems Inc., 2011). With the exception of Good's, these software provide little documentation of methods or accuracy assessment.

2.7 Tools for performing ellipsoidal trigonometry

Charles F. F. Karney developed a programming library called GeographicLib which has been implemented across several languages enabling the efficient and accurate calculation of geometric features directly on an ellipsoid independent of the distortions of any projection. It was largely created out of his dissatisfaction with the simpler and more widely used algorithm by Vincenty, which lacked his desired accuracy and often failed to converge on solutions in the case of antipodal or nearly antipodal points (Karney, 2011, p. 1). Karney's library, GeographicLib, draws primarily on prior work from Bessel as well as improvements from several others.

At double precision (53 bits) Karney's algorithm produces direct and inverse geodesic solutions with maximum errors of only 15 nanometers (Karney, 2011, p. 14), giving it notable usefulness in any GIS application that demands exceptionally high accuracy. A version of his library is available in Python, allowing for straightforward interoperability with more common ArcGIS based Python scripting. By default it only supports the WGS84 reference ellipsoid, but it can be easily extended to support any other datum.

2.8 Research objective

This project's core goal was to develop an improved implementation of Wefald's model for translating PLSS coordinates into latitude and longitude. Intended

improvements included comprehensive coverage for all PLSS territories (with the exception of certain Ohioan meridians which did not conform to the standard PLSS format), as well as complete automation of the generation of domain boundaries, removing the potential for human error and overgeneralization in the model. Automation also allows the model to be consistently rebuilt when BLM eventually publishes revised PLSS datasets in the future. We also investigated differences in accuracy between Wefald's TRS2LL domain geometry and regression results and the new model.

2.9 Contributions

While modern technologies such as GPS and GIS typically rely on spherical or UTM coordinates, PLSS coordinates are still, and will remain, relevant and useful for an array of applications beyond the description of land titles for which they were originally intended.

Useful datasets, both private and public, have been compiled that frequently report positions in terms of PLSS. Both flora (Auer et al., 2011) and wildfire (Short, 2014) databases serve as examples. And despite the fact that GPS measurements are more precise, PLSS positions are sometimes more accurately reported in databases (Ball, 2014).

Additionally, a wide variety of research topics, including (but not limited to) archeology (Cooper, 2008) and wildlife ecology (Joly et al., 2003) employ PLSS data. The latter is especially true in cases where researchers are dependent upon interviewees

who may not ordinarily carry a GPS to precisely record their whereabouts, or may be unwilling to share exact positions (Miller et al., 2003).

Our research advances the handling of PLSS data in GIS by improving upon Wefald's model, both by expanding PLSS conversion coverage to 30 states and moving beyond Wefald's hand-delineated domain boundaries through automation. This automation facilitates both flexibility and repeatability, which in turn allows our model's predictions to improve as the dataset upon which it was built is refined. And finally, our model provides a fully documented methodology and accuracy assessment that previous models have often omitted.

3. METHODS

3.1 Data sources

Digitized PLSS data has been aggregated by two major publishers: The Bureau of Land Management (BLM) and the U.S. Geological Survey (USGS). Each source is associated with advantages and disadvantages that should be considered before relying on either one.

Data from the BLM is obtainable from their publication website *GeoCommunicator* (Bureau of Land Management, 2012). This is the repository for the Cadastral National Spatial Data Infrastructure (CadNSDI) products, for which the BLM is the steward. The data here was not given as a single nationwide geodatabase, but rather it was offered per state. At the start of this project only 10 out of 30 states had been made available in the CadNSDI format. For the remaining states, they only offered shapefiles from alternate-sources, most of which originated from the USDA Risk Management Agency (Bureau of Land Management, 2012). North Dakota was a special case where the state's data is split between both CadNSDI and alternate-source formats. After this project began, BLM had discontinued the availability of the alternate-source datasets and replaced them all with CadNSDI geodatabases, but we observed cases where some newer CadNSDI datasets actually had less coverage versus the older alternate-source datasets (for example, the newer Oklahoma CadNSDI product is missing all Cimarron meridian data). Given what was available at the time this project began, the amount of time required to sanitize datasets for our purposes, and the fact that alternate-source datasets helped fill in

gaps in the coverage of newer CadNSDI sources, this project incorporated a combination of data from both formats.

GeoCommunicator alternate source data was available down to a section level. CadNSDI products provide data down to a quarter-quarter section level. Special care had to be taken when attempting to use BLM’s CadNSDI and alternate-source datasets together because the attribute strings describing PLSS coordinates are formatted differently. Further complicating matters is the issue of how Washington and Oregon appeared to have consistently yet erroneously incorporated a feature of CadNSDI formatted strings into the alternate-source format; namely, they included duplicated township codes into “Indkey” and “mtrs” fields (Figure 3.1). The sanitation of these attributes was vital before using these datasets in any GIS scripting.

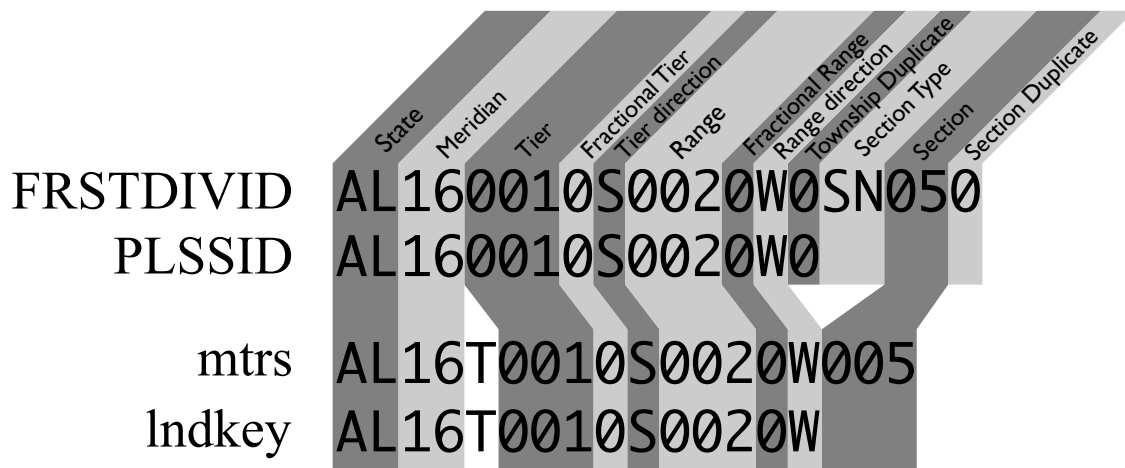


Figure 3.1: Depicted above is a diagram of how the CadNSDI formats for township (PLSSID) and section (FRSTDIVID) ID strings correlate with the alternate source dataset’s format for townships (Indkey) and sections (mtrs). Note that the latter lacks space for storing duplicate township codes.

Data from the USGS was obtained from their publication website *National Atlas* (U.S. Geological Survey, 2012). The *National Atlas* data was offered as a nationwide dataset, rather than being divided by state. Exactly from what original source data it was compiled is unclear but given that it often does not align with PLSS data obtained from the BLM it presumably was compiled from different original sources, or were digitized from the same source by a different technique. *National Atlas* data is provided only at a township level.

One thing that cannot be compared between the two sources of PLSS data is absolute accuracy since there has yet to be a comprehensive resurveying of markers with modern surveying techniques. Despite there not being a comprehensive resurveying of PLSS, it is not necessarily immutable as occasional surveys are performed and PLSS markers are sometimes lost or obliterated, and the nation has had disputes over defined state boundaries as recently as 1980 (Wilusz, 2002ab). The best that can be attempted in accuracy assessment is to assume that USGS topographic maps are the most authoritative source of PLSS data and then make comparisons of georectified USGS topographic maps to the *National Atlas* and *GeoCommunicator* data sets (Figure 3.2).

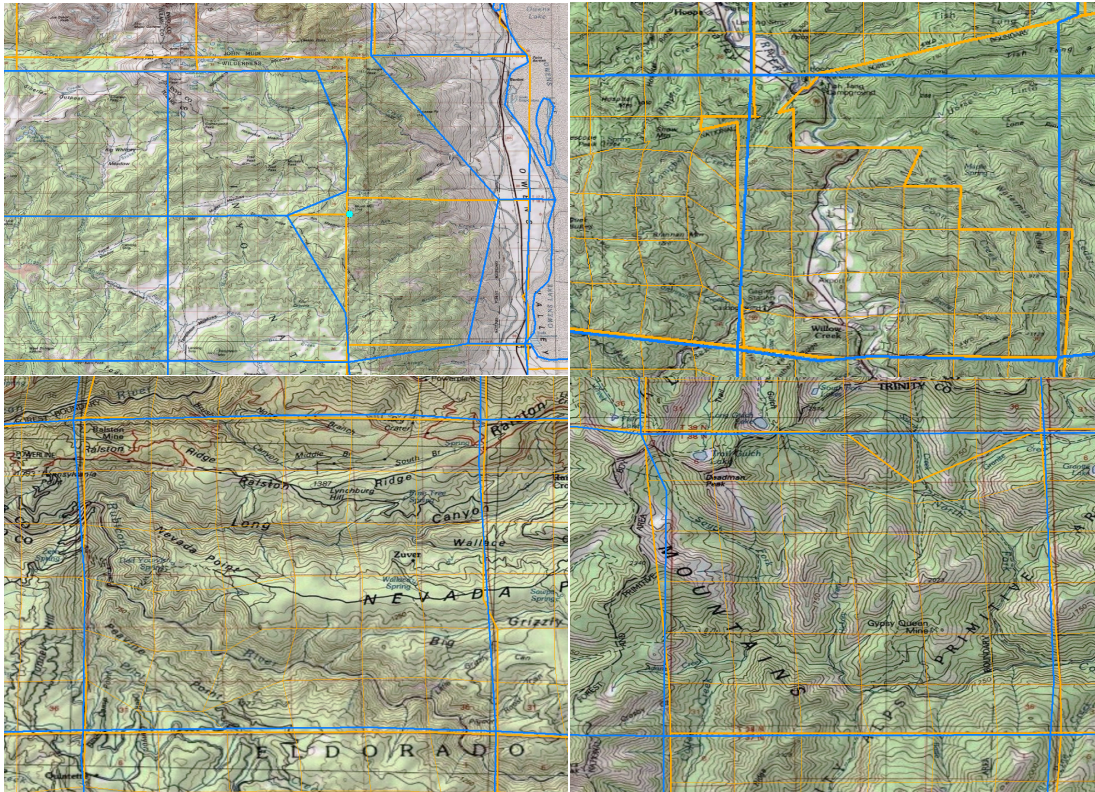


Figure 3.2: Comparison of PLSS data from GeoCommunicator (orange) and National Atlas (blue) overlaid atop USGS topographic maps (PLSS features are red). National Atlas features occasionally exhibit large gaps in coverage (top-left) that GeoCommunicator correctly covers. Other times GeoCommunicator correctly omits parts of a township that USGS omits while National Atlas over-extends the township in an attempt to meet nominal PLSS criteria (top-right). In most cases GeoCommunicator has townships that align more closely to topographic maps than National Atlas. Occasionally GeoCommunicator's townships align less closely to topographic maps due to obvious errors (bottom-right).

The *GeoCommunicator* PLSS dataset appeared to be in overall better alignment with USGS topographic maps than *National Atlas* maps. Again, it must be stressed that the PLSS definitions provided by USGS topographic maps are themselves a model and are not a perfectly accurate representation of PLSS. We only chose to evaluate the

accuracies of PLSS data with respect to them because they are presumed to be the best available PLSS definition.

3.2 Overview of analysis technique

The ultimate goal was the generation of at least two regression equations for each regression domain with a high degree of correlation, which implies that the features within each domain must possess consistent alignment and intervals of feature corners. Having sufficiently high correlations should suggest that the model provides accurate predictions of longitude, latitude, and PLSS coordinates, or at minimum, predictive of BLM's model of PLSS. Constructing this regression model required first reformatting all attribute data into CadNSDI format, then splitting a state's data by meridian. The next step was to accurately identify the corners of townships and sections from the vertices of their polygons. Next we generated regression domains based on the aforementioned township corners (or section corners in the case of Kansas, Nebraska, and Wisconsin, in order to encapsulate errors that would have been caused by their irregular section dimensions). Following that, we constructed a new continuous point-coordinate system analogous to range and tier with the ability to describe section-level corners (referred to here as decimal range and decimal tier), and performed linear regressions to relate longitude versus decimal range and latitude versus decimal tier. Then the coefficients of all the domain regression were merged into a single table that trivializes all PLSS to longitude-latitude conversions into simple first order polynomial equations (Figure 3.3).

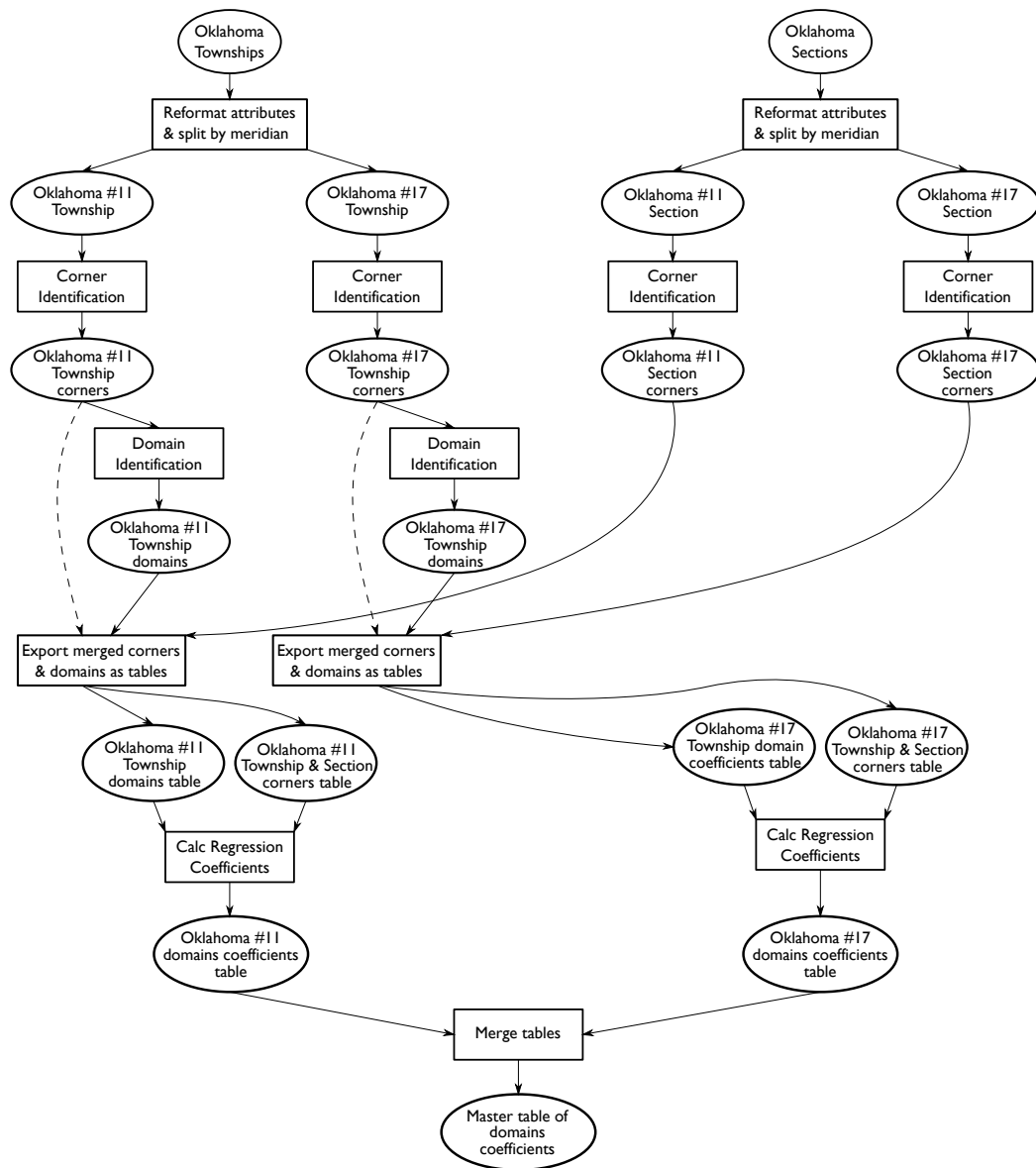


Figure 3.3: This diagram demonstrates a portion of the overall model, specifically the portion responsible for Oklahoma. There are 29 other states are unrepresented here. Note that the state of Oklahoma contains just two meridians: The Cimarron meridian (meridian code #11) and the Indian meridian (meridian code #17). For a state with more than two meridians, the first process produces more workflow paths. The dashed lines indicate optional inputs. In the final build of the model, these optional arguments were omitted so that only section level corner data was used in regressions, due to it being more granular and thus less sensitive to errors caused by truncated features. The states of Kansas, Nebraska, and Wisconsin utilized a slightly different model that omits any township data, and inputs section corner data into the domain identification.

All steps were assisted by a suite of scripts along with a new custom Python library (PLSS2LL). The library's core module for analysis of PLSS features is dependent on GeographicLib, mostly for the purpose of geodesic-distance calculations. The suite of scripts links against certain other modules in the PLSS2LL library that are dependent on the ESRI ArcPy© framework for convenience of reading and writing data.

3.3 Correcting, reformatting, and splitting datasets

Great effort was made to identify errors in the original data, and correct it whenever possible. But given that the dataset involves 6x6 mile squares and 1x1 mile squares with coverage of 30 US states, it's unreasonable to expect all errors were accounted for. The errors encountered include (but were not limited to) missing features, redundant overlapping features, erroneously digitized geometry, wrong PLSS attribute codes, and occasional violations of the PLSS ID formatting standards. Detection of these errors was initially performed manually, but as their frequency grew efforts were made to automate as much of the error detection as possible. Due to the varied and unpredictable sources of error, and since a particular type of error must be detected before code for that type can be created, the error checking remained partially manual.

In addition to erroneous data, some data was simply not appropriate to be included for the purpose of this project when it deviated from standard PLSS land description guidelines. For example, many of the section numbers along the Mississippi River in Mississippi's Washington meridian have irregular numberings that were unpredictable, sometimes having section numbers higher than 36 (Figure 3.4).

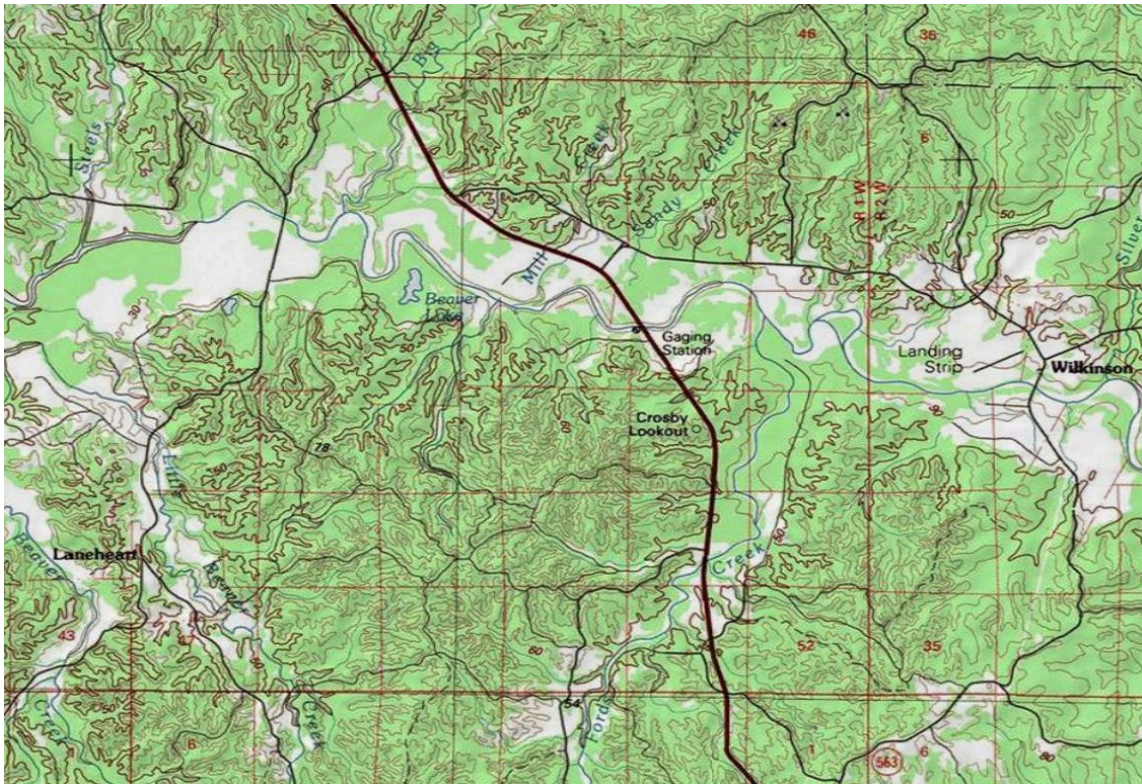


Figure 3.4: The above are examples of irregular, but not erroneous, section numberings exhibited by townships in Mississippi. While section numberings typically consist of values 1 to 36, we see values including 43, 52, and 35 here. Such section features can not be predicted by our model.

These were not technically errors because the section numberings were corroborated by USGS topographic maps and are thus likely their actual numbers which would make them legally binding descriptors, but they are unpredictable and would yield useless longitude and latitude regression later if retained in the dataset.

Within the CadNSDI datasets the PLSS identifiers include two characters reserved for the First Division Type Code. While the majority of these codes are “SN”, (which indicates “Section”) occasionally features have the same section number but a different

First Division Type Code such as “PB” (“Protracted Block”) or “UP” (“Unsurveyed Unprotracted”). When all the features of different First Division Type Codes but identical section numbers are examined as a whole typically resemble a complete 1x1 mile section feature. The distinctions between different First Division Type Codes are irrelevant to the purpose of this research as they don’t assist in predicting latitude and longitude, therefore all section level data derived from CadNSDI geodatabases had their First Division Type Codes coerced to be “SN” and were then dissolved. The intention of doing this was to make section features conform to normal 1x1 mile section geometry which in turn assisted the accuracy of the corner detection heuristics.

After the BLM provided data had been sufficiently sanitized, both in terms of feature attributes as well as feature geometry, each state’s dataset was split into separate feature classes by meridian code. This was to ensure none of the domains that are later generated span meridian boundaries.

The creation of a “new meridian” was also deemed necessary for a triangular region of land overlapping South Dakota and North Dakota. This land corresponds to the Lake Traverse Indian Reservation. The issue was a result of the fact that there were physically distinct townships with identical PLSS identifiers between the inside and outside of the triangular reservation. This violates the principle that a coordinate system should be able to guarantee uniqueness of a given coordinate, as well as the practical problem that it would later cause unexpected behavior for domain expansion algorithm if left unremediated. The features within the Lake Traverse Indian Reservation had their geometry clipped and attributes modified to reflect the creation of a new

meridian with a meridian code defined as 90 and a meridian name of “Sisseton”, named after the tribe that resides there (South Dakota Department of Tribal Government Relations, 2011). Special care must be taken when attempting to correlate the features in that region with official BLM PLSS formatted coordinates.

In addition to splitting the state data by meridian, all PLSS identifier attributes were then coerced into the format used by CadNSDI geodatabases. This format was preferred because it is more comprehensive in describing PLSS coordinates given that it didn’t require additional fields to store duplicate township codes. The entire PLSS coordinate can be represented down to a section level by a single string. Conversion of PLSS strings from the format used by the alternate-source data and into the CadNSDI geodatabases format was done by means of the PLSS2LL library.

3.4 Corner identification technique

Most PLSS townships or sections superficially resemble squares, but are polygons that frequently have many dozens of vertices. The vertices that best characterize each feature are the northeast, northwest, southeast, and southwest corners. These are also locations at which physical monuments were placed during the original and later surveys. Being able to distinguish these four corners was critical to generating the datasets that are later relied upon for domain generation. This was accomplished largely through a combination of 4 heuristics.

The internal angles of all township and section vertices were calculated. Any vertices with an internal angle of $90^\circ \pm 16^\circ$ were considered as potential candidates for

being corners, although this condition alone is not sufficient and serves only as a prerequisite for other corner detection tests.

The first and strongest test of a township or section vertex being a corner is whether it is touching 2 or more other corners from neighboring features. For each feature, the PLSS ID of its 8 neighbors is predicted, features with those PLSS IDs are queried, and their polygon vertices are searched for any matching points. While predicting for neighboring PLSS IDs to query was very generally reliable, it cannot always guarantee a perfect prediction of neighbors given that PLSS features with fractional range codes, fractional tier codes, and duplicate township codes are effectively unpredictable. In practice, these unpredictable features are rarely a problem given their rarity or the fact that many of the features with fraction ranges or tiers, or duplicate codes were the result of irregular surveying that was already going to result in poor geometry to begin with. A more significant shortcoming of testing for the coincidence of 2 neighbors' points was that along the meridian edges where neighbors cease to exist. Other corner detection tests helped serve as redundancies in such situations.

The second test is similar to the previous test, but is specialized for operating at the edge of a meridian by only requiring the presence of 1 touching vertex from neighboring features (rather than 2) and confirmation that the vertex is at the edge of the meridian by searching for vertices in a polyline dissolve of the township or section features which represents the hull of the meridian.

The final corner detection technique seeks to remedy the shortcomings of the first two. While the former tests rely heavily on the polygons possessing flush edges with

neighbors and complete rectangular shapes with minimal truncation, this final test does not, making it more suited for highly irregular geometry, such as the kind encountered along coastlines. All corners that have been identified by previous tests are first labeled according to their intercardinal orientation as being “northwest”, “northeast”, “southwest”, or “southeast”. If any of those 4 corner directions have not yet been matched with a vertex from the feature polygon, a search to find them begins. For those missing corners, it takes the bounding box corner for that intercardinal direction, and calculates the distance of every vertex in the feature polygon to that bounding box corner. The vertex with the shortest distance to the bounding box corner is identified as the corner for that intercardinal direction provided that it is within at least 200 meters (Figure 3.5).

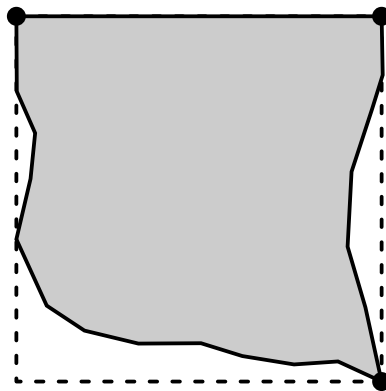


Figure 3.5: Example of the a coastal township who’s corners are detected by the distance-to-bounding-box-corners method. Circular nodes represent identified corners, and dashed lines indicate the bounding box. Note that the southwest corner was not detected because no vertex of the township was within 200 meters.

During corner identification efforts were also made to discern whether the 4 (or fewer) corners that are identified within a single township or section were suited for inclusion in later regressions by testing that the distances between adjacent corners were within a tolerance of $\frac{1}{12}$ that feature's normal side-length distance (± 0.500 miles for townships and ± 0.083 miles for sections). If a distance between adjacent corners exceeded this tolerance then one of the two corners was assumed to be a result of the feature being truncated by the meridian's edge, and was ignored in later regressions. The question of which corner was the corner afflicted by truncation was decided by offering priority to trusting corners closer to the meridian's initial point.

Finally, it should be noted that corner identification was the most time and RAM intensive step throughout the entirety of this research. Calculations for large meridians took many hours and RAM allocation can easily exceed the limits of what 32-bit versions of Python permit, particularly when handling larger meridians like those found in Arizona and New Mexico. This could not be easily worked around through parallelization of tiled datasets because information about a feature's neighboring features was required throughout corner identification. Therefore, 64-bit builds of Python were required.

3.5 Automation of regression domain generation

A feature class of township corners were the primary input for domain generation, except for Kansas, Nebraska and Wisconsin where section corners were used instead. This exception was a workaround to resolve an issue where certain columns of sections within townships were abnormally wide (Figure 3.6), which could reduce the correlation

of any longitude versus range or latitude versus tier on regressions later, particularly in the case of very large regression domains that would have spanned vast tracts of longitude.

26	25	30	29	28	27	26	25	30	29	28	27	26	25	30
T. 018 S R. 006 E				T. 018 S R. 007 E						T. 018 S R. 008 E				T. 018 S R. 009 E
35	36	31	32	33	34	35	36	31	32	33	34	35	36	31
02	01	06	05	04	03	02	01	06	05	04	03	02	01	06
11	12	07	08	09	10	11	12	07	08	09	10	11	12	07
14	13	18	17	16	15	14	13	18	17	16	15	14	13	18
T. 019 S R. 006 E				T. 019 S R. 007 E						T. 019 S R. 008 E				T. 019 S R. 009 E
23	24	19	20	21	22	23	24	19	20	21	22	23	24	19
26	25	30	29	28	27	26	25	30	29	28	27	26	25	30
35	36	31	32	33	34	35	36	31	32	33	34	35	36	31
02	01	06	05	04	03	02	01	06	05	04	03	02	01	06
11	12	07	08	09	10	11	12	07	08	09	10	11	12	07
14	13	18	17	16	15	14	13	18	17	16	15	14	13	18
T. 020 S R. 006 E				T. 020 S R. 007 E						T. 020 S R. 008 E				T. 020 S R. 009 E
23	24	19	20	21	22	23	24	19	20	21	22	23	24	19
26	25	30	29	28	27	26	25	30	29	28	27	26	25	30
35	36	31	32	33	34	35	36	31	32	33	34	35	36	31
02	01	06	05	04	03	02	01	06	05	04	03	02	01	06
T. 021 S R. 006 E				T. 021 S R. 007 E						T. 021 S R. 008 E				T. 021 S R. 009 E
11	12	07	08	09	10	11	12	07	08	09	10	11	12	07

Figure 3.6: Above in the highlighted region is an example of irregular section geometry. Despite being having well aligned corners with their neighbors, these features can impair the quality of regressions if not encapsulated within a regression domain of their own.

The domain generation algorithm is generalized enough such that it is agnostic as to whether the input data are section-level or township-level features. Consequently, the

operative logic between the two cases are identical. Although section-level domains were performed for the 3 aforementioned states, for the remainder of the discussion on domain generation I shall refer to domain generation in the context of townships.

The primary attributes of interest in domain generation were range and tier, but other minor components including fractional range, fractional tier, and duplicate township codes are also utilized in evaluating domain boundaries. For example, a township with a duplicate code of “A” would not be grouped into a domain with different duplicate code, otherwise the regressions later could be corrupted by the fact that two townships with identical tiers and ranges contained uncorrelated longitudes and latitudes. The domains involving a non-zero duplicate township codes naturally tend to be single-township domains since non-zero duplicate townships usually exist in isolation.

The process by which domains were generated starts with an arbitrary township. From this initial township begins an iterative process which cyclically attempted eastward, westward, northward, and then southward expansions (Figure 3.7). These are referred to as the primary expansions. Perpendicular to each primary expansion were up to 2 secondary expansions, which are both perpendicular to the primary expansion. These extend just far enough to complete a row or column of townships such that an overall rectangular shape is maintained. Each primary and secondary expansion is subjected to alignment tests. Should even a single misalignment be detected in a primary expansion or its associated secondary expansions, they are aborted, and a new primary expansion direction is attempted.

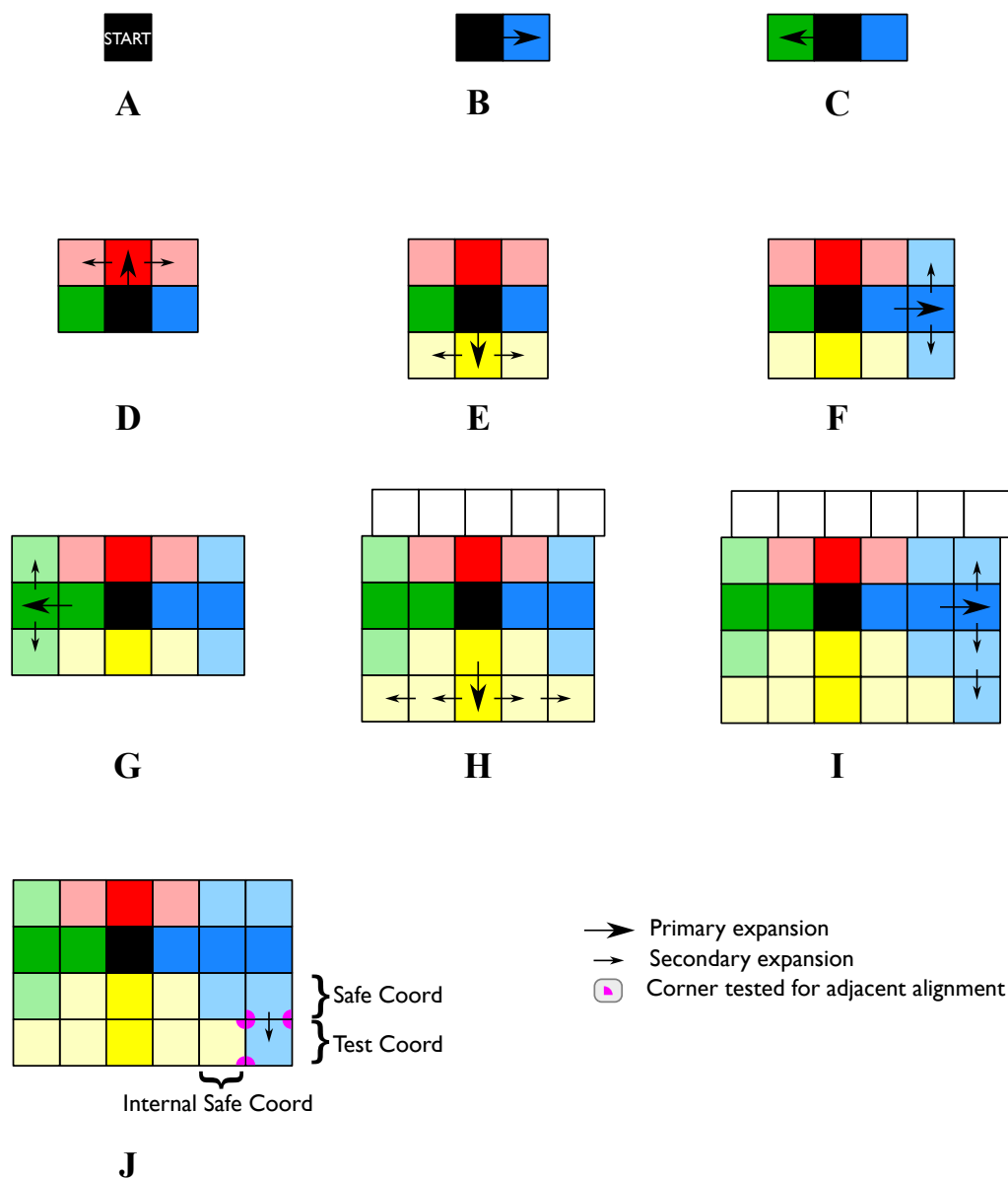


Figure 3.7: Depicted here are 9 steps in a domain expansion (A through I). Primary expansions are attempted in the order of east (blue), west (green), north (red), and then south (yellow). Black indicates the initial township from which expansion began. Note that between steps G and H a primary expansion northward was attempted, but halted by misalignments of neighbors along the north edge (the townships in white), so the next successful expansion is southward. The final graphic (J) demonstrates in detail which of its neighbor's corners came into play when testing adjacent alignments in the final secondary expansion.

Should an entire cycle of east, west, north, and south primary expansions fail in a row, the domain ceases to attempt any further expansion, and the domain's extents are recorded in terms of domain edge coordinates. The domain's edge coordinates are the PLSS ID strings of the east-most, west-most, north-most, and south-most townships that were also in-line with the initial township at which domain expansion had begun.

While the ostensible morphology of individual PLSS features is simplistic, the original surveyors periodically deviated from nominal 6x6 mile square townships, which has led to an assortment of ways for two townships to be misaligned with one another. To accommodate the variety of misalignments, a variety of heuristics were developed to detect them: A simple adjacent point-distance test, a shear test, a slice test, and a size difference test.

The adjacent point-distance test takes the corners of two neighboring townships, and calculates the distance between the pairs of adjacent corners (Figure 3.8). Should the distance between a pair of adjacent corners exceed a particular threshold, it fails the alignment test and halts the expansion in that direction. Secondary domain expansions also include a slight variation of the adjacent point-distance test that is applied along an additional perpendicular edge of a tested township that faces the domain's starting position (Figure 3.7).

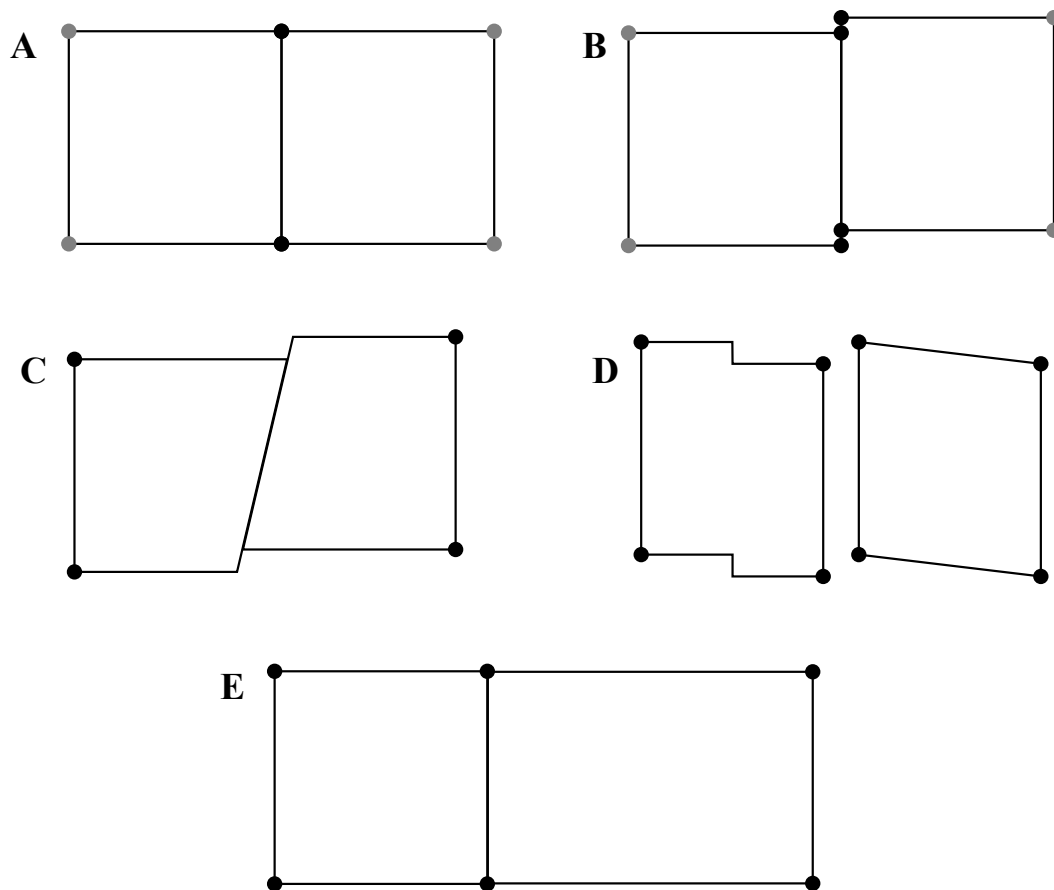


Figure 3.8: These are simplified examples of the geometry tested for by the various alignment tests within the domain expansion algorithm. The most common examples of a well-aligned pair of adjacent townships that permit domain expansion to pass (A) as well as poorly aligned pair of adjacent townships that would halt domain expansion (B) are demonstrated here. In these examples, all well defined township corners are pictured as circular nodes on the polygon. An example of the slice test (C) demonstrates how a misalignment can be determined from the outer corners of both townships when corners along the adjacent edge don't exist. Also present are two examples of the shear test (D), which detects internal misalignments within individual townships. And lastly, there is the case (E) of a pair of townships whose abrupt change in relative size would halt a domain expansion from one to the other.

In the case of two neighbors where one township is missing an adjacent corner, and the other township's adjacent corner does exist, alignment is presumed until proven to be misaligned. This is to accommodate the fact that occasionally poor geometry of features results in a failure to detect a corner, but is usually still well enough aligned. And usually when one pair of corners along a common edge is misaligned the other pair is as well.

Another alignment test is the slice test, which seeks to handle uncommon cases where two townships meet with diagonal edges that result in no pairs of corners along the common edge (Figure 3.8). In this situation rather than testing the distances between two adjacent corners it tests the axial-alignment between two far corners.

Another test that will halt domain expansion is the shear test (Figure 3.8). It's similar to the slice test but rather than being a test of alignment between neighbors the shear test is one of internal alignment within a single township. Two types of shearing appear in the data, one where opposite edges are misaligned connected via right angle turns in the joining edges, and another where they're connected via diagonal joining edges. In either case it is computed identically since the only data available to the domain expansion algorithm are the actual corners themselves (not all vertices in the polygons). Like the slice test, the shear test is one of axial-alignment.

The last major domain alignment test checks for abrupt differences in relative sizes between two neighboring townships (Figure 3.8). It was designed largely in response to inconsistently sized features in a minority of regions, which if incorporated into a larger

domain would have negatively impacted the correlation of the regressions across vast tracts of land.

All of the aforementioned tests of alignment were done by means of distance calculations using Karney's GeographicLib on the NAD83 ellipsoid. All tests of alignment also had adjustable thresholds. For the final build of the domain model, the thresholds were 200 meters for adjacent point alignments and 800 meters for axial alignments. Maximum size changes were thresholded with a maximum difference in edge size of 10% for township level domains and 25% for sections level domains.

One initial complication that the domain expansion algorithm encountered in early development was the fact that the edges of meridians often do not conform to rectangular shapes. Many states have diagonal boundaries or coastlines with highly irregular polygon geometry. So early implementations resulted in many small single-township domains along the meridian's edges. This issue was resolved by incorporating "virtual townships" which allows the domain expansion algorithm to behave as though townships exist in places where no townships actually exist at all by creating townships with no corner data, and in the absence of corners to compare for all the aforementioned tests it presumes alignment. The intended result is fewer and simpler domains along meridian edges, and better generalization of the model as a whole.

Once all domain expansions are complete and every township belongs to a domain, each domain's edge coordinates can be utilized to test whether any PLSS coordinate exists within the bounds of a particular domain. This functionality is later vital to performing a join of corner attributes to domain attributes, as well as being able to

efficiently look up what set of regression coefficients to use in a PLSS to longitude and latitude conversion. The edges being recorded in terms of PLSS has also the advantage of being less ambiguous than tests based on extents in terms of longitude and latitude since longitude and latitude based extents between 2 domains could meet with a perfectly flush edge, and a point feature along that edge could otherwise be interpreted to belong either domain. Preserving the extents in terms of PLSS coordinates prevents any ambiguity as to what point or polygon features ought to be associated with a domain, so long as those point or polygon features have PLSS attributes.

3.6 Exporting domains and corners as tabular data

Once all domain feature classes and corner feature classes had been generated they served as inputs for generating tables of domain data and tables of corner data which were later be utilized for calculating regression coefficients. The corner tables included fields for representing the positions of each section corner in terms of longitude, latitude, decimal range, decimal tier values, and the unique domain ID of the domain that the feature that the corner belonged to.

Ordinary range and tier values in the PLSS system are integers that refer to a specific township, but in order for us to be able to refer to features smaller than a township we had to incorporate non-integer range and tier values. They can be thought of in terms of either mixed fractions or decimal range and tier values.

Also, any records of corners in the tables that specify the same position in terms of decimal range and decimal tier and were within the same domain were merged into a single corner record where the longitude and latitude are an average of all the adjacent

corner records' longitudes and latitudes. The values for the field indicating the corner's domain affiliation is assigned by searching the list of domains for the one domain that the corner's PLSSID falls within the extents of. This field is necessary for the purpose of being able to efficiently subset the corner records while calculating the regressions.

Domain tables include the domain's unique identifier, as well as the domain's extents in terms of decimal range & decimal tier and also longitude & latitude. The latter is necessary for converting from longitude and latitude back to PLSS.

3.7 Predicting latitude and longitude by regression

Regression coefficients were calculated by looping through each domain record, and we then subsetted the corner records that corresponded to that domain. The subset of corners corresponding to that domain was then used as input for simple linear regressions, similar in principle to how Wefald used corners in his regressions (Wefald, 2009); one relating latitudes to the PLSS tier designations, the other relating longitudes to the range designations. Wefald identified the corresponding section corners by hand and for practical reasons this limited him to approximately 3 section corners for regression input per domain, but here the process was automated and all section corners in the domain were utilized.

Note that since sections are a numerically 1-dimensional boustrophedonic coordinates that represent a physically 2-dimensional grid, section numbers are not a practical input into the regressions without prior transformation into a Cartesian coordinate. Also, while standard range and tier values are non-zero integers the coordinate system we wish to predict (longitude and latitude) is continuous. Therefore it

was necessary to convert all range, tier, and section values into a continuous 2-dimensional Cartesian coordinate that was suited for regression. To accomplish this we employed intermediate coordinates referred to here as decimal range and decimal tier. Like latitude and longitude (but unlike PLSS coordinates) decimal range and decimal tier coordinates refer to exact points in the space of PLSS (Figure 3.9). The township's range and tier values correspond to the decimal range and decimal tier values of the township corner that is furthest away from the meridian's initial point.

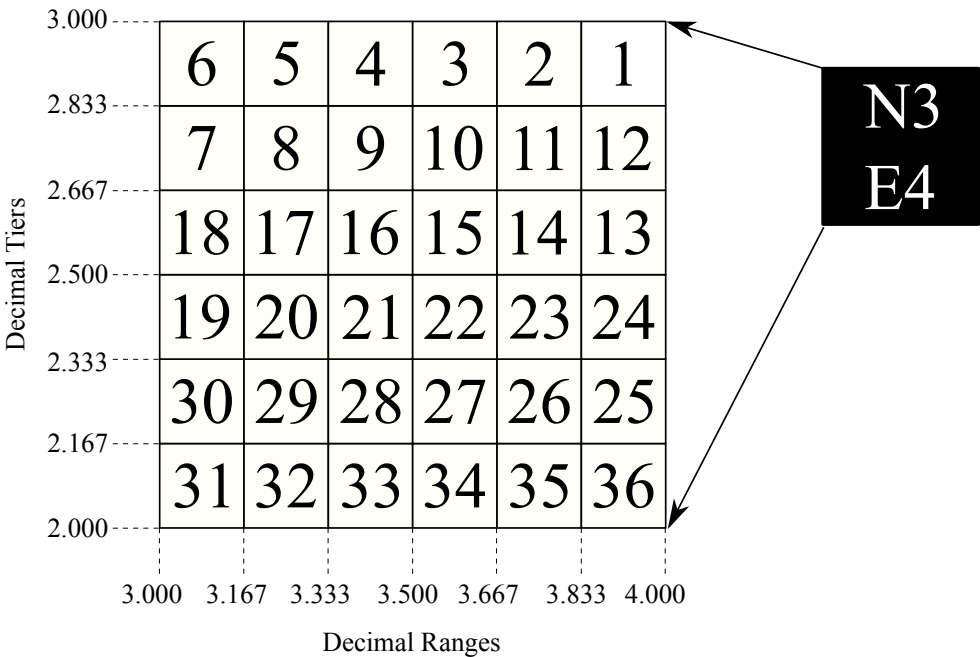


Figure 3.9: Depicted above is an example of a township at 3 North and 4 East, it's constituent sections, and a mapping of decimal range and decimal tiers to its features. The decimal range and decimal tier serves as useful intermediate coordinates that facilitate the prediction of latitude and longitude by transforming integer range, tier, and section values into a Cartesian coordinate.

The origin of the decimal range and decimal tier space corresponds exactly to the meridian's initial point.

It does merit mentioning that you can have unique PLSS coordinates that are physically distinct, yet become indistinguishable in decimal range and decimal tier space when encountering townships with non-zero duplicate townships codes, fractional tiers, and fractional ranges. However, since the domain-expansion algorithm is successful at preventing expansion through townships with different duplicate township codes, fractional tier codes, or fractional range codes, the features were not be conflated and impose undue variance on regressions later.

The transformation from range, tier, and section to decimal range and decimal tier can be modeled by 5 simple equations. Let n represent the section grid size (typically $n=6$), and S represent a PLSS section-number (typically 1, 2, ..., 36). Then $s = n^2 - S$ and

$$y_s = \lfloor s/n \rfloor \quad (1)$$

$$x_s = \lfloor s - (y_s n + (n - 1)[y_s + 1] \% 2) \rfloor, \quad (2)$$

respectively, represent the vertical and horizontal Cartesian coordinates, numbered 0 to 5, of the section's location within a township. The signed integers t and r are respectively the tier and range values. The values of t are signed such that north tiers are positive and south tiers are negative. The values of r are signed such that east and west ranges respectively correspond with positive and negative values. The transformation to

a continuous Cartesian coordinate is completed by calculating T and R which are respectively the decimal tier and decimal range, which are effectively equivalent to

$$T = t - \left(1 + \frac{\text{sign}(t)}{2}\right) + \frac{1+y_s}{n}, \quad (3)$$

$$R = r - \left(1 + \frac{\text{sign}(r)}{2}\right) + \frac{1+x_s}{n}. \quad (4)$$

The point (R, T) represents the outermost corner of a section that's furthest away from the meridian's initial point. Obtaining the other corners can be accomplished by adding offsets of $\pm \frac{1}{6}$ along the appropriate axis.

The decimal range and decimal tier values were then used as predictors to generate two regression equations: one to predict longitude (λ) from decimal range values and the other to predict latitude (ϕ) from decimal tier values for any position within the domain. Where T is decimal tier and R is decimal range, the equations are respectively

$$\phi = \beta_0 + T \beta_1 + \varepsilon_i \quad (5)$$

and

$$\lambda = \beta_0 + R \beta_1 + \eta_i \quad (6)$$

where in each equation β_0 and β_1 are coefficients to be estimated and ε is the error term.

These largely correspond to equations used by Wefald (Wefald, 2009) with the exception of 38 of his 1798 domains in which Wefald reduced the fitted β_1 values by a small amount whereas we did not. We suspect that he made this modification because he noticed trend in the residuals for these 39 of his domains. Residual plots for the approximately 8000 domains were examined for each fitted model, but no transformations were performed. Unlike Wefald, we chose to also employ the reverse

regression that predicts range from longitude and tier from latitude, as well as an alternative regression technique alongside the ordinary least squares linear regressions: orthogonal regression. Orthogonal regression is a type of error-in-variables model that doesn't presume all error exists on the predicted variable, but rather error exists on both the predictor and predicted variables (Fuller, 2006). Once all of the β_1 and β_0 coefficients and goodness of fit statistics are for each domain were calculated, they were then appended as fields to a table of domains for that meridian and state. Then the domains of all states and meridians were merged into a single table of all domain data. This table is then embedded into the PLSS2LL package such that it can be referenced for the purpose of conversions from PLSS coordinates to longitude and latitude. The converted PLSS coordinates can be either township or section level coordinates.

3.8 Sampling methodology for accuracy assessment

An ideal accuracy assessment would involve comparing groundtruthed locations of PLSS markers across the United States, determining their location with survey grade GPS measurements, and then comparing those locations against TRS2LL and PLSS2LL. Unfortunately groundtruthing PLSS markers across most of the nation is cost-prohibitive, so we instead used the centroids of BLM section level feature classes as a basis of comparison in lieu of surveyed markers. We randomly generated PLSS section coordinates within states and meridians that are within the coverage of TRS2LL. Then we generated longitudes and latitudes via TRS2LL and PLSS2LL for the randomly generated PLSS coordinates. Using GeographicLib we then calculated a geodesic

distance and azimuth from the centroid to the PLSS2LL prediction and the TRS2LL prediction.

3.9 Predicting PLSS from longitude and latitude

While the principle focus of this research was on producing longitudes and latitudes from PLSS coordinates since that is generally more useful for the purpose of GIS, the reverse conversion was also implemented. The orthogonal regression equations used for predicting longitude and latitude were algebraically inverted, which permitted us to predict decimal range and decimal tier. But for testing purposes we also implemented ordinary regressions for predicting longitude, latitude, and PLSS coordinates. (Fuller, 2006) We can then convert decimal range and decimal tier to a range, tier, and section by exhaustively testing what section's decimal range and decimal tier boundaries encapsulate the predicted point.

But to do the aforementioned prediction one must first know which domain's regression coefficients to call upon and the domain affiliation of a given longitude and latitude can be ambiguous in cases where the extents of multiple domains overlap the same point. The non-unique relationship between longitude and latitude and domains therefore implies multiple sets of regression coefficients could be used. The worst-case scenario would be a latitude and longitude that falls near the corner of a domain, overlapping as many as four domains. Another potential complication arises in which a latitude and longitude could fall upon a section or township boundary touching up to four PLSS features. Due to these inter-domain and intra-domain complications, multiple answers could exist for any given point and an attempt to reduce the solution to a single

answer will be a matter of interpretation. It is for this reason that PLSS2LL returns all possible answers and leaves the decision of which is best as an exercise for the end-user. This is in contrast to TRS2LL, which will only ever return a single answer.

4. RESULTS

4.1 Overview

The model for TRS2LL supported just 17 states, but PLSS2LL expanded coverage to all 30 PLSS states (Figure 4.1). TRS2LL's model included 1,798 regression domains. The PLSS2LL model initially generated 8,439 domains, but after filtering out domains that failed to calculate a sensible regression coefficients or had too few points to use in regression the model, the model was reduced to 7,388 regression domains. While 12.4% of the generated results being unusable may appear to be a problem, it is worth noting that these rejected domains were exceedingly small in terms of physical area (often only being one township in size or less), resulting in little practical impact to the model's coverage.

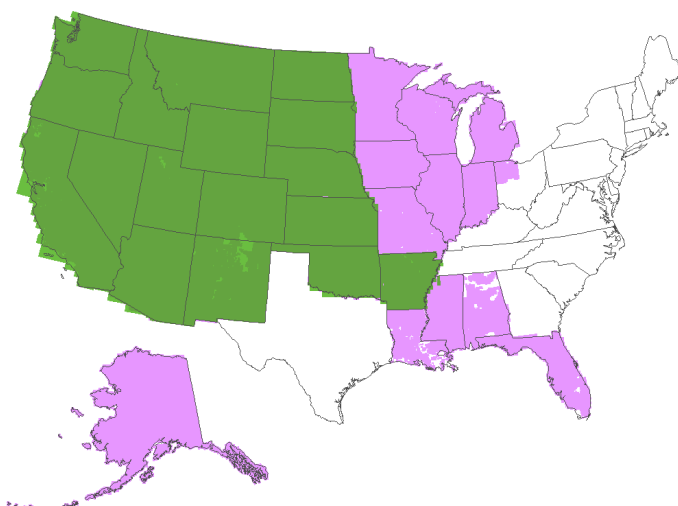


Figure 4.1: Regions in green indicate coverage for TRS2LL. The domain coverage of PLSS2LL encompasses regions in purple and green.

Within just the states that TRS2LL supported, PLSS2LL generated 5,975 domains, of which 5,751 were viable. This indicates an increase in the complexity of our model versus Wefald's by a factor of 3.2. This can likely be attributed to the fact that Wefald determined his domain boundaries by hand, whereas our model employed automation which was guided by user-defined parameters which had been set to relatively narrow tolerances.

There were also some differences caused by the fact that PLSS2LL does support handling of township duplicate codes while TRS2LL does not, but such differences account for a relatively small fraction of the land coverage. Duplicate townships only composed 0.05% of all townships (Table 4.1). PLSS2LL's regression domains cover 7,662,540 km², but domains that incorporate duplicate townships only cover 1,401 km².

Both models do support handling of fractional tier and fractional range codes. Although it should also be noted that PLSS2LL's corner detection thresholds were typically high enough to filter out corners of features that were of odd widths and heights. The consequence of this was that regression domains for features with non-zero fractional tier or fractional range codes were often missing too many corners to produce a viable regression. Despite the fact that this can cause gaps in the coverage of PLSS2LL, the rarity of PLSS features that actually use non-zero fractional tier and fractional range codes means that even when one such gap occurs, it's typically small in terms of geographic area, at least individually. Only 1.02% of all townships were fractional townships (Table 4.1).

Table 4.1: The table below presents the frequency of all fractional and duplicate townships per state. While domains composed of such townships posed a challenge for accurate predictions they remain uncommon features relative to the size of the dataset.

State	Fractional Tier Townships	Fractional Range Townships	Duplicate Townships	All Townships
Alabama	0	0	0	1304
Alaska	4	10	0	17808
Arizona	86	69	0	3409
Arkansas	1	0	0	1507
California	61	27	0	4572
Colorado	25	36	14	3059
Florida	0	0	0	1677
Idaho	0	4	0	2490
Illinois	0	0	0	1740
Indiana	0	0	0	1102
Iowa	0	0	0	1637
Kansas	0	0	0	2358
Louisiana	0	0	0	1190
Michigan	0	0	0	1836
Minnesota	0	0	0	2529
Mississippi	0	0	0	1450
Missouri	0	0	0	1971
Montana	3	7	0	4188
Nebraska	0	0	0	2235
Nevada	88	142	22	3354
New Mexico	6	10	0	3245
North Dakota	0	0	0	2054
Ohio	0	0	0	274
Oklahoma	0	0	0	2051
Oregon	28	104	4	2889
South Dakota	0	0	0	2377
Utah	42	58	0	2564
Washington	7	8	0	2042
Wisconsin	0	0	0	1633
Wyoming	9	12	0	2884
Total	360	487	40	83429
Total %	0.43%	0.58%	0.05%	

The combined area of all domains that incorporate fractional range or fraction tier townships is 22,742 km², and the combined area of all domains which have non-viable regressions is 17,651 km², much (but not all) of which was a result of fractional tier and fractional range townships with irregular dimensions. We expect that fine-tuning the thresholds used in our particular PLSS2LL model could remediate this issue in the future.

Ohio was only partially modeled by PLSS2LL given the irreconcilable irregularities regarding the way in which townships and sections were numbered in some meridians. The irregularities are likely due to the fact that these were some of the first meridians surveyed in the United States.

4.2 Sampling for accuracy assessment

We sampled 255 random section features. While PLSS2LL predictions included both ordinary least squares and orthogonal regressions, the typical variance between the two was found to be on the order of millimeters to centimeters. Given the negligible difference between PLSS2LL's ordinary and orthogonal regressions, for the purpose of comparisons between TRS2LL and PLSS2LL we shall focus on the orthogonal regressions in lieu of the ordinary least squares model. The accuracy assessment is limited to states and meridians within the coverage of TRS2LL's regression domains; vast tracts of additional states introduced by PLSS2LL are not compared. The sampling also makes no attempt to sample townships with non-zero fractional ranges, fractional tiers, or duplicate township codes.

4.3 Magnitudes of errors (PLSS2LL versus TRS2LL)

For each sampled section a geodesic distance was calculated from the section's centroid to the predictions of each model. The mean error of TRS2LL was found to be 221 meters, and the mean error of PLSS2LL was 134 meters. Median errors for TRS2LL and PLSS2LL were respectively 151 meters and 80 meters, and the standard deviations of TRS2LL and PLSS2LL were 284 meters and 151 meters, respectively.

A comparison of the distribution of errors shows that PLSS2LL had 59.61% of the samples occur within less than 100 meters of the centroid, whereas with TRS2LL only 29.41% were within 100 meters (Figure 4.2). Within the sample the errors produced by PLSS2LL was confined to a distance less than 1 kilometer. Despite all this, there was a minor peak in the frequency errors at 750 meters of PLSS2LL data. For some of these TRS2LL had smaller errors, but for the majority of them the corresponding TRS2LL errors were larger than the PLSS2LL errors.

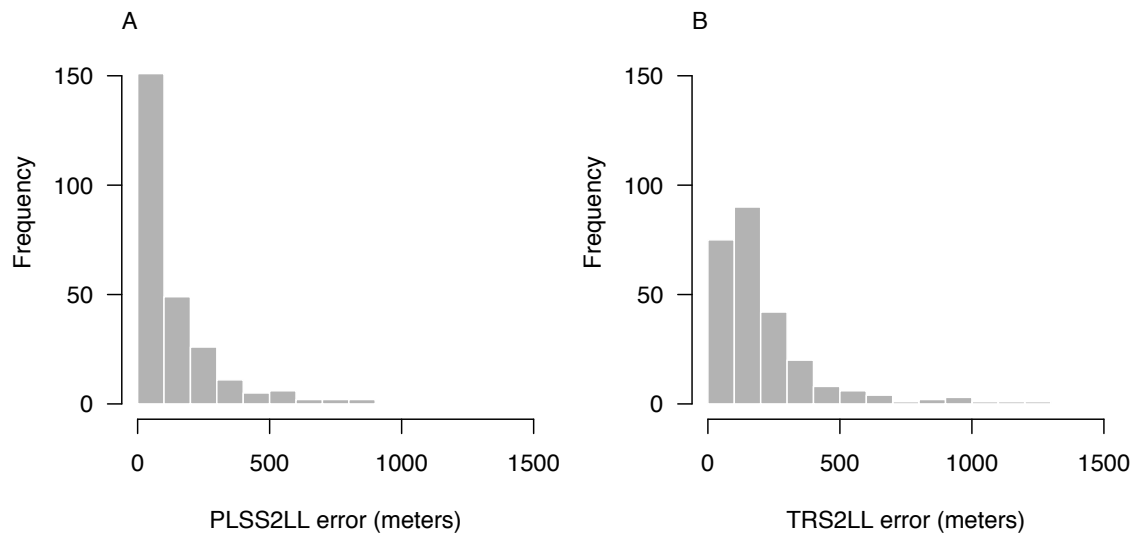


Figure 4.2: Histograms of the distance from sections’ centroids to the predictions of both PLSS2LL (A) and TRS2LL (B) reveal substantive differences in the error distributions of each model. The error of one TRS2LL sample was in excess of 1,500 meters.

4.4 Directional bias of errors (PLSS2LL versus TRS2LL)

In addition to geodesic distances, we also calculated the azimuths from each sampled sections’ centroid to the predictions of each model in an effort to detect any potential systemic biases in one model or the other. The distributions reveal what could be a directional bias in TRS2LL predictions toward the northwest and south (Figure 4.3).

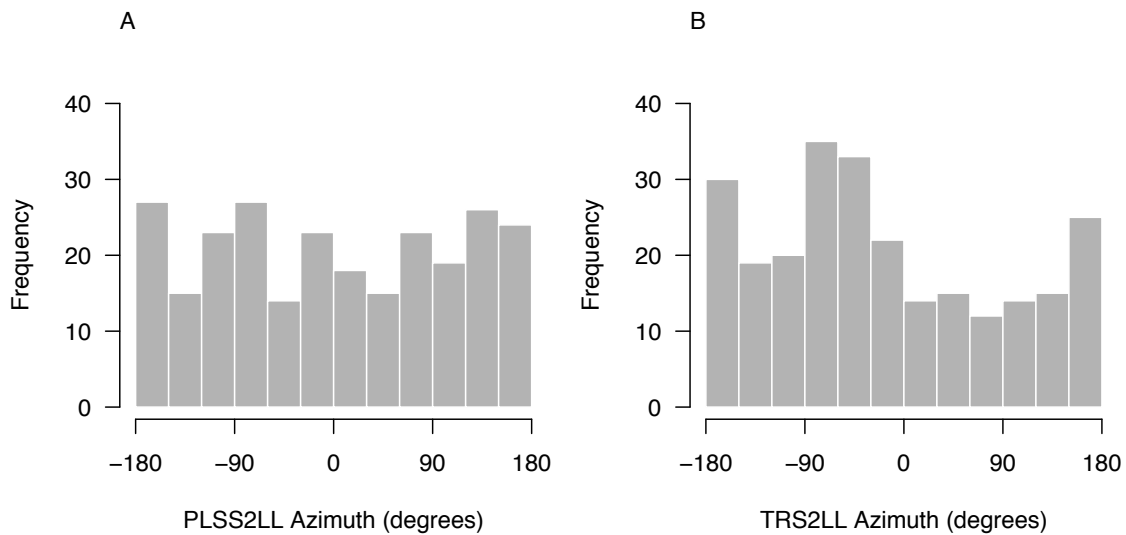


Figure 4.3: Depicted above are the azimuths of geodesics calculated from the centroid of randomly sampled sections to the predictions of PLSS2LL (A) and TRS2LL (B).

The source of the directional bias in TRS2LL was uncertain. PLSS2LL was generated from BLM data that was built against North American Datum 1983. Wefald relied on physically printed USGS maps with PLSS grids, which we know used North American Datum 1927 (Wefald, 2009). This bias occurs despite the fact that we corrected for the difference in datums by converting TRS2LL’s NAD27 coordinates into NAD83 coordinates before calculating the distance from each centroid to predictions by each model.

4.5 Inspection of individual failures in PLSS2LL predictions

Of the random samples that were found to have PLSS predictions that deviated from the centroid by more than 300 meters, a dozen were selected for hand inspection to determine the likely cause of their inaccuracy.

One case in Arizona involved a domain where a column of townships had 5x6 grids of sections instead the nominal 6x6 grid (Figure 4.4). Despite the fact that the decimal range and decimal tier equations we used do potentially support section grids other than 6x6s, we have yet to implement code that actually detects when such irregular townships exist so that the equations can have their arguments adjusted accordingly. Effectively, PLSS2LL applied 6x6 section grid math on a 5x6 grid, ruining the decimal range values for a column of sections. And since this column of townships wasn't at the outer edge of the domain, the error had enough leverage on the regression equation coefficients to impair longitude predictions throughout the domain. Despite all this, the ordinary least squares longitudinal R^2 prediction for this domain was 0.999457. This demonstrates that R^2 statistics alone are not sufficient for detecting poorly predictive domains.

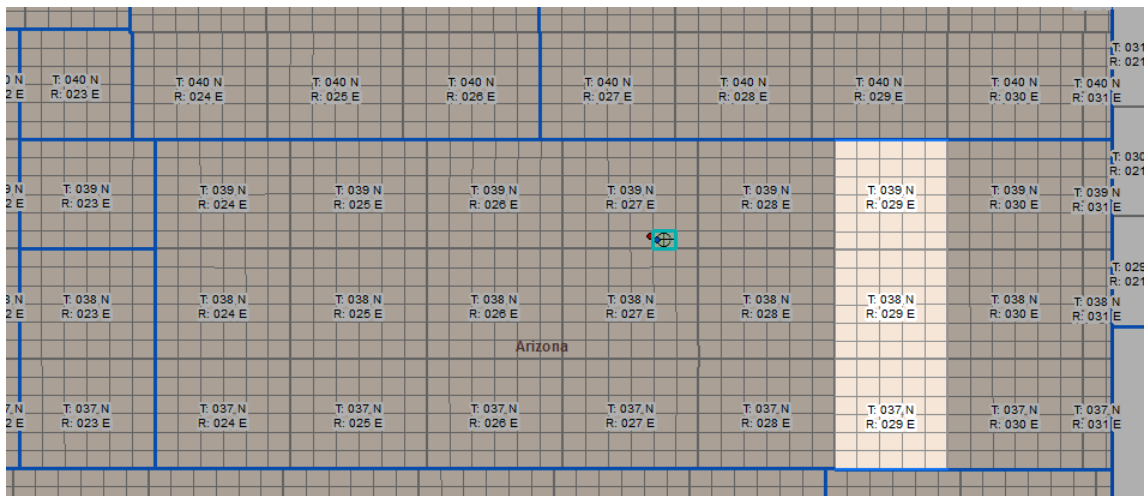


Figure 4.4: Blue lines indicate PLSS2LL domain boundaries, the cyan selection in the middle indicates the randomly sampled section. And the highlighted region on the right indicates the column of townships with 5x6 section grids that skewed the predictive accuracy of PLSS2LL. TRS2LL domains were too large to adequately represent here for comparison.

Another randomly sampled section in Montana demonstrated a similar failure in longitudinal accuracy, but it was caused by a domain that had expanded across a body of water. This was possible because the body overlapped where two townships would have had their edges meet, but those townships' geometry was clipped to the body of water. The result was along the edge where there were no detectable corners, and thus the domain expansion gave the benefit of a doubt to the missing corners' alignments.

The rest of the dozen audited samples were roughly evenly split between two more common sources of failure: Localized shear of townships and sections which represent only a subset of the domain's area or very gradual overall shear of an entire domain. The occurrence of localized shears within a domain is trivial to recognize upon inspection (Figure 4.5). Such cases are not uncommon in regions with mountainous terrain, particularly in California. While PLSS2LL's domain expansion algorithm does in fact check for local shearing from one individual township to another, if the magnitude of the shearing was just below the detection threshold or if the corners used for shear detection were not detected in the first place it can sometimes be joined with similar neighbors into a relatively small domain.

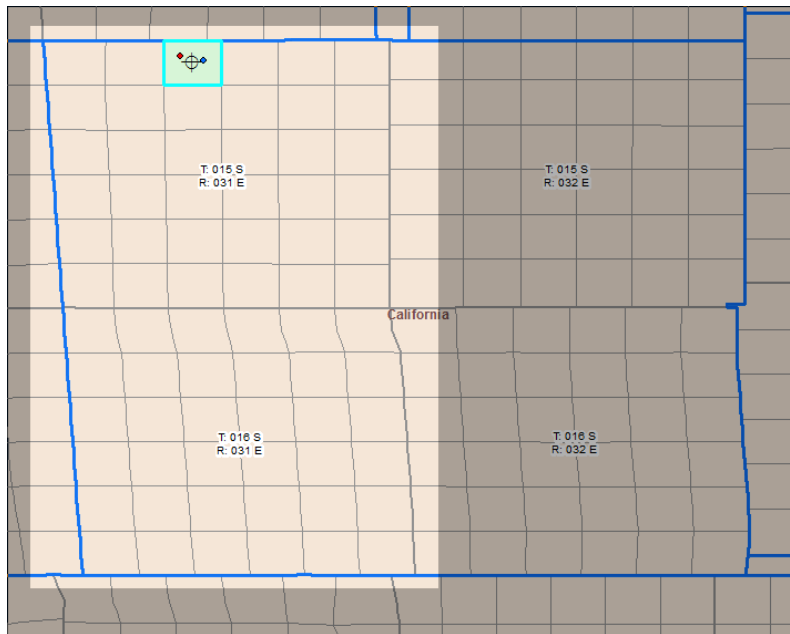


Figure 4.5: Demonstrated here is a case of PLSS features with local horizontal shearing. Blue lines depict PLSS2LL domains. TRS2LL domains are not depicted due to their greater size. This negatively impacts the accuracy of predictions under both PLSS2LL as well as TRS2LL. It is not uncommon within mountainous terrains that were difficult to survey.

The more problematic case we encountered was gradual shear over exceptionally wide domains. These were not readily identifiable by eye, and typically required comparing longitudes and latitudes of the domain's corners and calculating a geodesic from one domain corner and projecting toward the adjacent corner and comparing distances between the actual and projected corners. We presume they are a result of the very gradual accumulation of errors (typically latitudinal errors). Once a surveyor made such a mistake, they may have then perpetuated it northward and southward in an effort to keep townships' edges flush with one another. One such case occurred in Wyoming with a large domain that was 578 kilometers wide. The domain's southwest and

southeast corners differed in latitudinal alignment by 1.78 kilometers, which is more than the length of a section. The impact of this large yet subtle shearing was that PLSS2LL's error of the section's center deviated by 606 meters. One potential remedy for this issue in the future may be to put caps on the maximum width of domains during expansion.

While examining the regressions' R^2 value in states outside of the PLSS2LL-versus-TRS2LL comparison we also found another source of error that arose from slight rotations in at least 9 domains within Missouri and Arkansas (Figure 4.6). Wefald also noticed this inclination of Arkansas, and noted it to be $>1^\circ$ degree and that his solution for mitigating error "requires the areas to be broken down into smaller areas, and even then the accuracy is impaired" (Wefald, 2009).

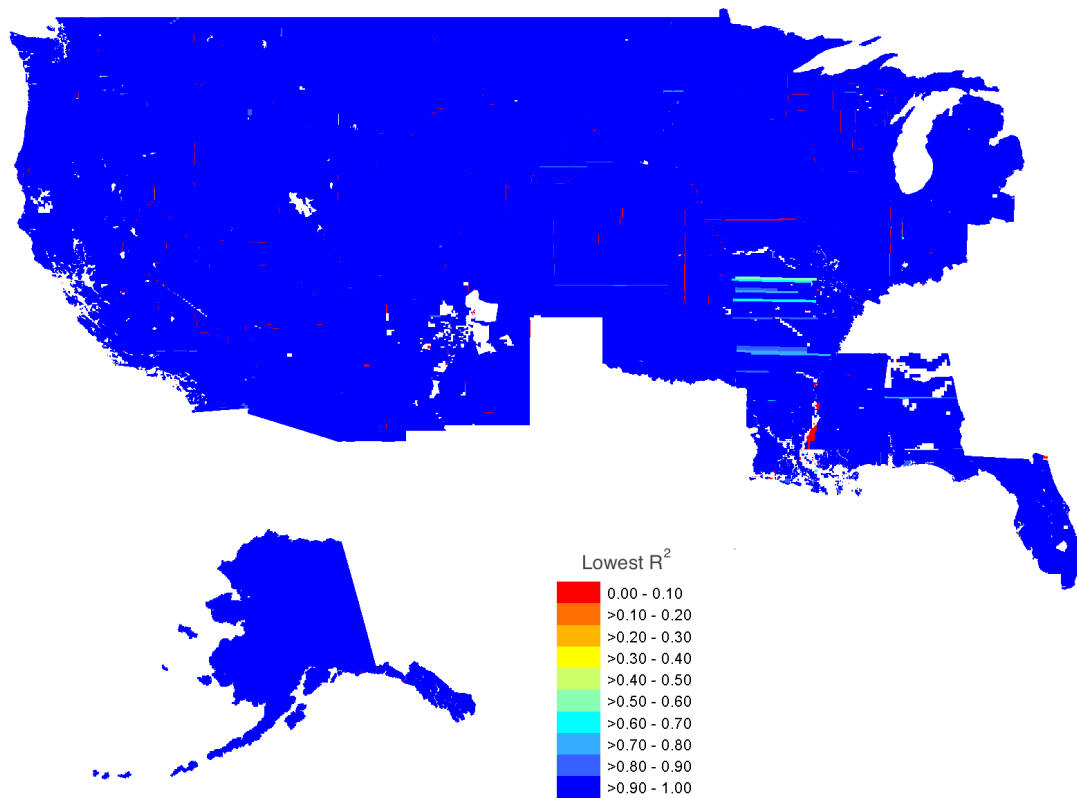


Figure 4.6: Depicted above is a map of all PLSS2LL domains, classified according to the value of either their ordinary least squares regression's longitudinal R^2 values or latitudinal R^2 value (whichever is lower). Note that in Arkansas and Missouri several domains with slight rotations suffered a markedly reduced lowest R^2 value relative to their neighbors. Also indicated here are domains with nonviable regressions due to too many missing corners, shown in red. These are typically due to townships or sections with geometry so thin that corner detection filtered out points in the domain due to their irregular dimensions. Future builds of PLSS2LL could use revised threshold values for handling features with irregular dimensions to be more forgiving so that enough corners are kept to maintain a viable regression in those areas.

When a domain's orientation rotated away from cardinality it introduced large amounts of variance to the latitudinal regressions. The afflicted domains tended to be particularly wide relative to their height (much like the domains most affected by gradual shearing). The most severe case of rotation-induced variance precipitated a

latitudinal R^2 of 0.58, despite having a longitudinal R^2 of 0.999996. Any domains with a longitudinal or latitudinal R^2 of less than 0.90 were removed from the final model.

Also, removed from the model were any domains that failed to make viable regressions both horizontally and vertically. These domains were frequently so thin that corner detection filtered out points in the domain due to their irregular dimensions, leaving too few corners for generating viable regressions. Future versions of our model might be able to generate regressions in the future by relaxing the thresholds for handling features with irregular dimensions so fewer corners are filtered out.

While a fair amount of discussion here has been devoted to the occurrence of errors, there is one case that merits attention for its quality. The regression domains in all of Alaska's meridians were observed to possess exceptionally simple and well-formed geometry (Figure 4.7). This stood in stark contrast to other states where domains were markedly more irregular and fragmented (Figure 4.8). We attribute Alaska's simple domain structure to the fact that dataset for Alaskan dataset was unique in that it was not the result of physical surveying. The original dataset for Alaska could be thought of as being equivalent to a state that was perfectly surveyed and is thus a kind of control. We take the uniquely well-formed domains of Alaska as confirmation that the more fragmented domain structures observed in other states are an accurate reflection of errors made by the original land surveyors.

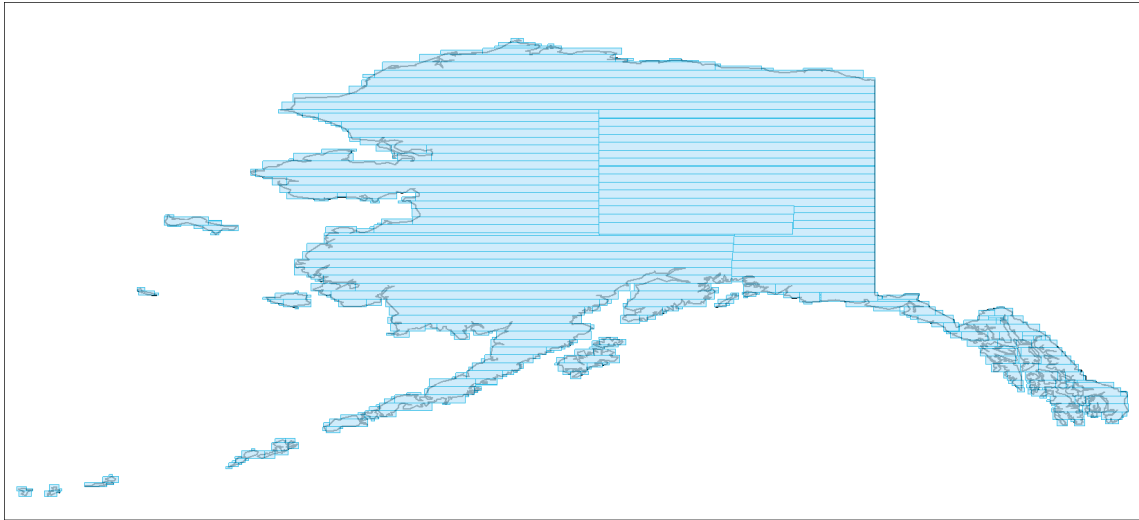


Figure 4.7: Alaskan domains generated by PLSS2LL. Note that Alaska’s domains are exceptionally simple in structure. This is likely due to the data not being a product of surveys, and thus is free of surveying errors. If Alaska is ever fully surveyed, it’s likely that the domain structure would still be well formed due to advances in surveying techniques and technology.

Visual comparisons of overall domain structures between PLSS2LL and TRS2LL confirm that the former is more fragmented than the latter. Areas where this is particularly noticeable are mountainous regions such as those in California (Figure 4.8). We also observed that North Dakota was a case where domain boundaries between TRS2LL and PLSS2LL domains were in an unusually high degree of agreement, to the point that the boundaries were almost identical (Figure 4.8). But the two models offered extremely different domain boundaries in South Dakota, particularly around the Lake Traverse Indian Reservation (Figure 4.8). Part of the cause may be the fact that Wefald exploited a strategy where he would deliberately layer domains atop one another while we did not.

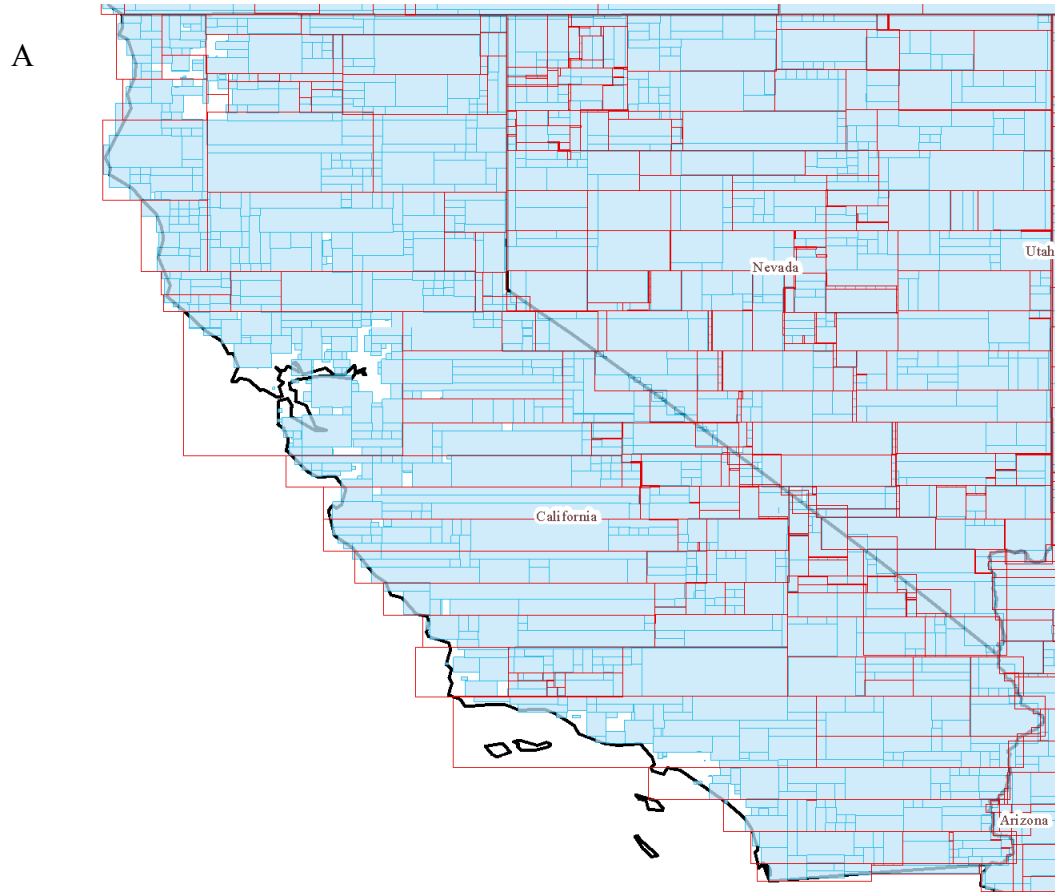


Figure 4.8: Regression domains generated by PLSS2LL are indicated in blue, and TRS2LL domains are in red. Depicted here are California (A), North Dakota (B) and South Dakota (C). PLSS2LL domains tended to be more complex than those utilized by TRS2LL, but the former's domains were automatically generated while the latter's were crafted manually. Mountainous regions tended to induce additional domain fragmentation in PLSS2LL, as seen in California. The regression domains between the two models were largely in agreement in North Dakota, but in South Dakota the triangular region in the North East required TRS2LL to implement many thin and long domains to accommodate it. PLSS2LL populated the triangular region with shorter and more compact domains along the triangle's edges.

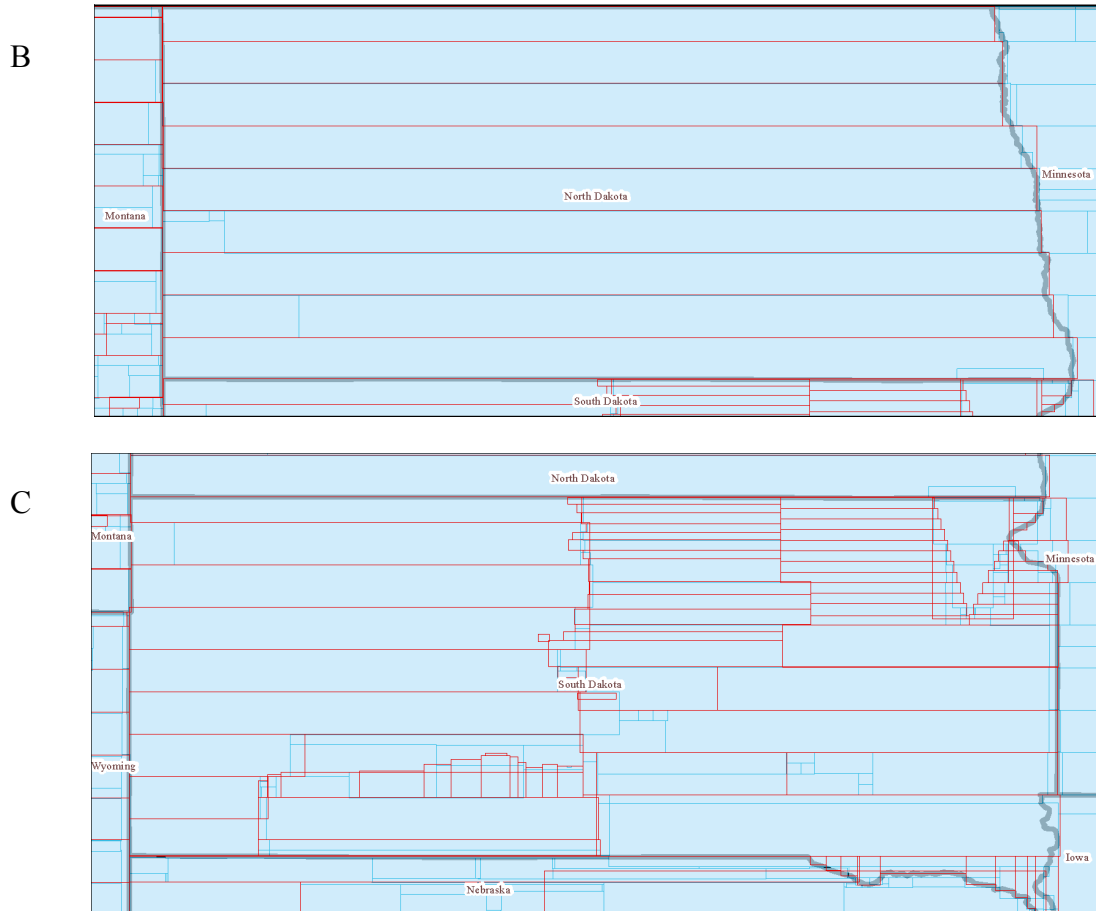


Figure 4.8: Continued.

4.6 Latitude and longitude to PLSS accuracy assessment

In order to assess the accuracy of the reverse conversions in which PLSS coordinates are predicted from longitude and latitude, we randomly generated 1,000 points within the states that that TRS2LL modeled, but due to an apparent bug in how TRS2LL returned Montanan meridian codes we were forced to omit any predictions affiliated with Montana leaving us with a sample of 890 points. We used a spatial join in ArcGIS to assess the true PLSS coordinate each random longitude and latitude was affiliated

with. We compared the known PLSS coordinates against the prediction of TRS2LL and the multiple predictions of PLSS2LL, respectively yielding accuracy rates of 84.48% and 90.33%. It is uncertain as to whether the superior accuracy of PLSS2LL is due to better domain structure and more predictive regression coefficients, or if it is merely a consequence of the fact that TRS2LL has restricted itself to only returning a single prediction in ambiguous cases where domains overlap while PLSS2LL returned all possible answers.

5. SUMMARY

5.1 Concluding assessment

Relative to its predecessor, our model has produced vast gains in land coverage, finally encompassing all PLSS states. We also observed gains in accuracy. The fact that TRS2LL was based on NAD27 also highlighted the need for an improved model such as ours which put the onus of datum conversion/consistency on the shoulders of the program and not the end-user.

Wefald's original work produced thousands of domain boundaries that were all delineated by hand with physical topographic maps. Doing such work manually is tedious and makes long-term maintenance difficult, particularly when attempting to expand the coverage to all territories. For these reasons we elected the route of automation for the construction of our model. Consequently, it can be rapidly rebuilt to reflect future updates made to the Cadastral NSDI and alternate source datasets by the Bureau of Land Management. This automation also affords enhanced flexibility as it can be rebuilt with custom optimizations for either accuracy or simplification depending on the values assigned to the domain expansion thresholds. This potential for customization is useful given that improvements to accuracy or simplification tend to be inversely related, and different applications may need to prioritize the former or the latter. Adjustments to the models that make gains in accuracy would likely produce more numerous and smaller domains, whereas a more simplified model would call for fewer and larger domains that tend to be less predictive. Additionally, TRS2LL only

produced conversions for the centers of section features, whereas our model optionally offers the centers or corners of townships or sections.

5.2 Future research

We anticipate there to be merit in pursuing additional research on this topic, first and foremost with respect to finding additional means of mitigating error. The aforementioned problem of gradual domain shearing may be solvable by restricting the maximum longitudinal length of domain expansion. This may also mitigate some of the variance introduced into latitudinal regressions in rotated domains, although that may be better remediated by employing more complex predictions than simple linear regressions. Given that the corner identification process accounts for most of the time required to build the model (taking approximately a week), speed optimizations to the corner identification tool could assist with rapid redeployment and testing of any new variations of the model.

The input dataset was 575 MB after stripping away the attributes that were unnecessary for the construction of our model. The final PLSS2LL model that we used reduced this to approximately 4 MB of text. This indicates a reduction in size by a factor of 143.75, making it feasible for deploying PLSS conversion applications on mobile devices that can't rely on large amounts of memory nor processing power.

REFERENCES

- Aadland, D., & Kolpin, V. (2004). Environmental determinants of cost sharing. *Journal of Economic Behavior & Organization* , 53, 495-511.
- Auer, T., MacEachern, A. M., McCabe, C., Pezanowski, S., & Stryker, M. (2011). HerbariaViz: A web-based client–server interface for mapping and exploring flora observation data. *Ecological Informatics* , 93-100.
- Avery, T. E., & Burkhardt, H. E. (1994). *Forest Measurements* (4th Edition ed.). Boston, Massachusetts: McGraw-Hill.
- Ball, K. C. (2014). *Wildland Fire Data: Issues and Proposed Solutions*. Dissertation, Texas A&M University, Ecosystem Science and Management, College Station.
- Barkworth, M. E., Anderton, L. K., McGrew, J., & Giblin, D. E. (2006). Geography and Morphology of the *Bromus Carinatus* (Poaceae: Bromeae) Complex. *Madroño* , 53 (3), 233-243.
- BioGeomancer Working Group. (2007). Retrieved November 2, 2012, from BioGeomancer: <https://sites.google.com/site/biogeomancerworkbench/>
- Bureau of Land Management. (2012, 11 15). *GeoCommunicator Home*. Retrieved 11 15, 2012, from GeoCommunicator: <http://www.geocommunicator.gov>
- Bureau of Land Management. (2012, 11 15). *Public Land Survey System Download* . Retrieved 11 15, 2012, from GeoCommunicator: http://www.geocommunicator.gov/GeoComm/lsis_home/home/index.htm#plss
- Committee on Integrated Land Data Mapping. (1982). *Modernization of the Public Land Survey System*. Washington D.C.: National Academy Press.

- Cooper, J. R. (2008). *Bison Hunting and Late Prehistoric Human Subsistence Economies in the Great Plains*. Southern Methodist University. Ann Arbor: ProQuest.
- Dingle, H., Zalucki, M. P., Rochester, W. A., & Armijo-Prewitt, T. (2005). Distribution of the monarch butterfly, *Danaus plexippus* (L.) (Lepidoptera: Nymphalidae), in western North America. *Biological Journal of the Linnean Society* , 85, 491-500.
- Fuller, W. (2006). *Measurement Error Models*. New York: John Wiley & Sons.
- Good, D. I. (1964, June). Mathematical Conversion of Section, Township, and Range Notation to Cartesian Coordinates. *Kansas Geological Survey* , 170 (3).
- Joly, D. O., Ribic, C. A., Langenberg, J. A., Beheler, K., Batha, C. A., Dhuey, B. J., Rolley, R. E., Bartelt, G., Van Deelen, T. R., Samuel, M. D. (2003). Chronic Wasting Disease in Free-Ranging Wisconsin White-Tailed Deer. *Emerging Infectious Diseases* , 9 (5), 599-601.
- Karney, C. F. (2011, February 7). Geodesics on an ellipsoid of revolution. *ArXiv* , 1-28.
- Miller, R., Kaneene , J. B., Fitzgerald , S. D., & Schmitt, S. M. (2003). Evaluation of the Influence of Supplemental Feeding of White-Tailed Deer (*Odocoileus Virginianus*) on the Prevalence Of Bovine Tuberculosis in the Michigan Wild Deer Population. *Journal of Wildlife Diseases* , 39 (1), 84-95.
- Moe, R. (2007, September 18). *Consortium of California Herbaria: About*. Retrieved October 24, 2012, from The University and Jepson Herbaria: <http://ucjeps.berkeley.edu/consortium/about.html>

- MyTopo. (2009). Retrieved November 2, 2012, from Terrain Navigator Pro:
<http://maptech.mytopo.com/land/terrainnavigatorpro/index.cfm>
- Neunzert, G. M. (2011). *Subdividing the Land*. Boca Raton, Florida: Taylor & Francis Group.
- Pattison, W. D. (1957). *Beginnings of the American Rectangular Land Survey*. Chicago, Illinois: University of Chicago Press.
- Scantek Systems Inc. (2011). *US Public Land Survey System*. Retrieved October 22, 2016, from Convert Legal Land Description to Latitude and Longitude:
<http://legallandconverter.com/p43.html>
- Short, K. C. (2014). A spatial database of wildfires in the United States, 1992–2011. *Earth System Science Data* , 6 (1), 1-27.
- South Dakota Department of Tribal Government Relations. (2011). *Sisseton Wahpeton Oyate*. Retrieved October 1, 2016, from South Dakota Department of Tribal Relations: <http://www.sdtribalrelations.com/sissetonwahpeton.aspx>
- Treves, A., Naughton-Treves, L., Harper, E. K., Mladenoff, D. J., Rose, R. A., Sickley, T. A., Wydeven, A. P. (2004). Predicting Human-Carnivore Conflict: a Spatial Model Derived from 25 Years of Data on Wolf Predation on Livestock. *Conservation Biology* , 18 (1), 114–125.
- U.S. Geological Survey. (2012, 11 15). *Map Layers*. Retrieved 11 15, 2012, from National Atlas: <http://www.nationalatlas.gov/maplayers.html>

- Wefald, M. (2009, November 22). *TRS2LL*. Retrieved August 8, 2015, from Marty Wefald's Home Page:
<https://web.archive.org/web/20150808002342/http://wefald.com/trs2ll.html>
- White, C. A. (1926). *A History of the Rectangular Survey System*. Washington D.C.: U.S. Dept. of the Interior, Bureau of Land Management.
- Wilson, B. L., Hipkins, V. D., Rey-Vizgirdas, E., & Kaye, T. N. (2005). Variation in *Lewisia Kelloggii* (Portulacaceae) with Description of a New Species Endemic to Idaho. *Western North American Naturalist* , 65 (3), 345-358.
- Wilusz, J. P. (2002, February). Part 1 The Colorful History of the California/Nevada State Boundary. *Professional Surveyor* , 22 (1), pp. 6-12.
- Wilusz, J. P. (2002, February). Part 2 The California/Nevada State Boundary. *Professional Surveyor* , 22 (2), pp. 22-28.