To: Appalachian Basin Geothermal Play Fairway Analysis Group

From: Jared Smith

Date: 15 October, 2015

Subject: Selection of Four Counties in Each State with “Best” Thermal Resources

Applicability: The methods described here were used to select the four “best” counties in each state according to the thermal resource. This analysis complements the Play Fairway maps that are based on the combination of the other three risk factors with the thermal resource, but this analysis is specific to thermal attributes.

Introduction

At the conclusion of Phase 1 of this Geothermal Play Fairway Analysis project, the four most favorable or attractive counties in each of New York, Pennsylvania, and West Virginia must be selected for further inspection of the commensurate favorability of geothermal development. Favorability is primarily determined by high thermal resource quality, specifically the depth to 80 °C as defined in the Statement of Project Objectives [SOPO]. Anticipating that the thermal resource quality will be a core factor in decisions regarding development of geothermal direct-use projects at specific locations, in this memo we investigate the values and uncertainty of the thermal resource in the counties that currently appear to be of best thermal quality and of high interest to potential users.

The following analysis could be applied to any county or site of interest to a potential user. To select counties for which to illustrate the insight that is gained from the analysis, additional factors we have taken into consideration are:

i) Whether or not reservoirs and population/utilization centers are available to use the resource in the county

ii) The number of wells within the county from which additional detailed lithology may be obtained to cross-examine the predicted thermal values, and from which to collect additional data in Phase 2

iii) The location of all selected counties within the region of study. Spatial variety was desired such that the selected counties did not all occupy the same hot spots.

High thermal resource quality was interpreted as a location having both a shallow predicted mean depth to 80 °C, and relatively high certainty in the predicted mean depth. No value was assigned to the certainty that was considered to be high; however in all but three counties the average uncertainty in the predicted mean throughout the county has a two-standard-error spread within ± 500 m.

The selected counties are presented in Figure 1. Each of the counties are represented below in a series of cutout maps of the predicted mean depth to 80 °C within the counties. The color scale on each map is different, tuned to the local temperature-depth relations in order to show variability within the counties of interest. A cross section through each county is provided, which depicts the uncertainty in the predicted mean as

$$\hat{\mu} \pm (2 * SE) \quad [1]$$

where $\hat{\mu}$ is the predicted mean and SE is the standard error of the predicted mean. These bounds can be thought of as 95% confidence bands about the predicted mean. A second cross section of the Thermal Play Fairway Metric (0-5 point scale, see ThermalResourceThresholds_final.doc memo for discussion) is also provided for each county.

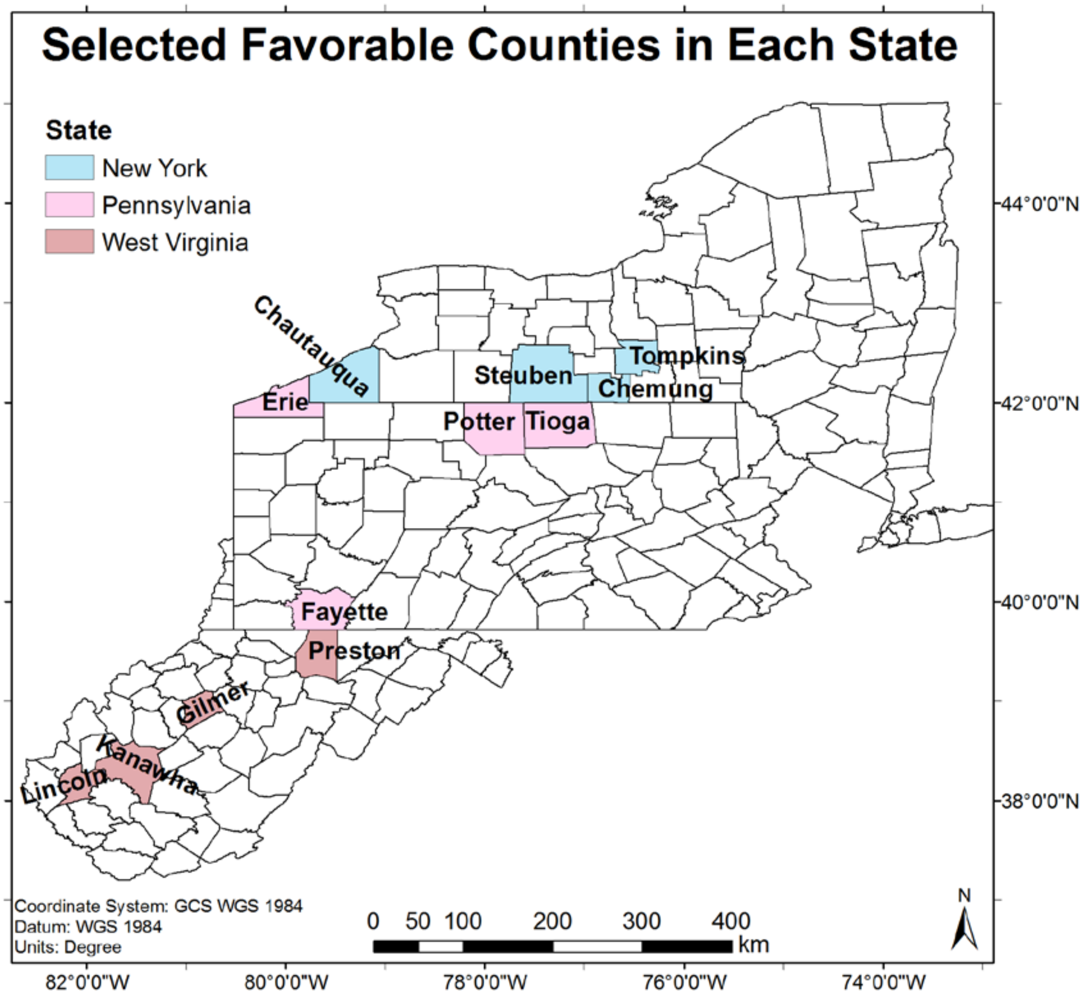


Figure 1. *Selected four best counties in New York, Pennsylvania, and West Virginia based on thermal resource, presence of reservoirs, population centers, and variety of location within the Appalachian Basin.*

Results and Discussion

The most promising counties have predicted mean depths to 80 °C that are shallower than 2500 m. These include Preston, Gilmer, Lincoln, and Kanawha counties in West Virginia; Chemung, Steuben, and Tompkins counties in New York, and Tioga and Potter counties in Pennsylvania.

Chautauqua, NY, is a great example of where there is high certainty in the prediction as a result of many wells (Figure 2). The highly clustered well data provides insight to the spatial variability of the thermal field on a small scale. For example, the variation in the predicted mean depth to 80 °C is in some locations on the order of a hundred meters on a horizontal scale of about 10 km in map distance (Figure 3). This provides insight to the spatial scale of thermal resource variability that may be expected everywhere in the basin, but is not captured on the maps as a result of fewer data available to support the predictions.

Many of the counties have prediction boundaries (e.g., interpolation zone boundaries) that cut through the county (prediction boundaries are not shown on the individual county maps, but are provided on the cross sections, and the regional thermal resource maps for the basin). As a result, predictions on one side of the boundary may be discontinuous compared to those on the other side of the boundary; however not all boundaries show a meaningful statistically significant difference (e.g. at the $\alpha \approx 5\%$ level) in the predicted mean on either side. The meaningful statistically significant differences are potentially indicative of real boundaries in the thermal field. An example of one meaningful statistically significant difference is in Gilmer County, WV (Figure 8, Figure 9). This boundary is thought to represent the Rome Trough – a feature of known structural importance in the basin. It is not clear if the concurrence of a statistically significant difference and the Rome Trough is a result of poor well sampling, or if this is a real boundary in the thermal field. One argument against poor well sampling is that a two-sample t test in the difference of the mean depth to 80 °C calculated for each well rejected the null hypothesis:

$$\mu_1 - \mu_2 = 0$$

with a p-value of 1.2×10^{-4} and 314 degrees of freedom. Assuming unequal variance, the p-value is 8.9×10^{-5} with 313.6 degrees of freedom. Even so, the wells included in this test are located predominantly in northern Gilmer County, so the test reflects the difference in two means across this northern boundary rather than the sharper difference in southwestern portion of the county. More sampling in southern Gilmer County may change the significance of this test.

At this time, it is also not clear if all statistically significant differences across interpolation boundaries coincide with features of similar importance to the basin, and further influence the thermal field. Even so, the gravity and magnetic potential field edges were used to define prediction boundaries in an attempt to capture differences in the subsurface that may correspond to variations in the thermal field (e.g. changes in the data generating process). Having one statistically significant difference appear along a feature of importance is encouraging support of this assumption.

Evaluation of the predicted mean depth to 80 °C was conducted for the region using a “leave one out” cross validation. For N points, this cross validation runs the kriging interpolation algorithm N times, with one point left out of the prediction in each of the N repetitions. The results of the cross validation are provided for each county below in Figure 2, Figure 4, Figure 6, Figure 8, and Figure 10 as ‘bubble plots’ that display the Z-Score: $Z = \frac{x - \mu}{SE}$, where x is the geothermal variable for the well point, μ is the predicted mean for the grid cell, and SE is the standard error of the predicted mean for the grid cell. Results indicate that on a regional scale about 98% of the data for the left-out points lie within 3 standard errors of the predicted mean.

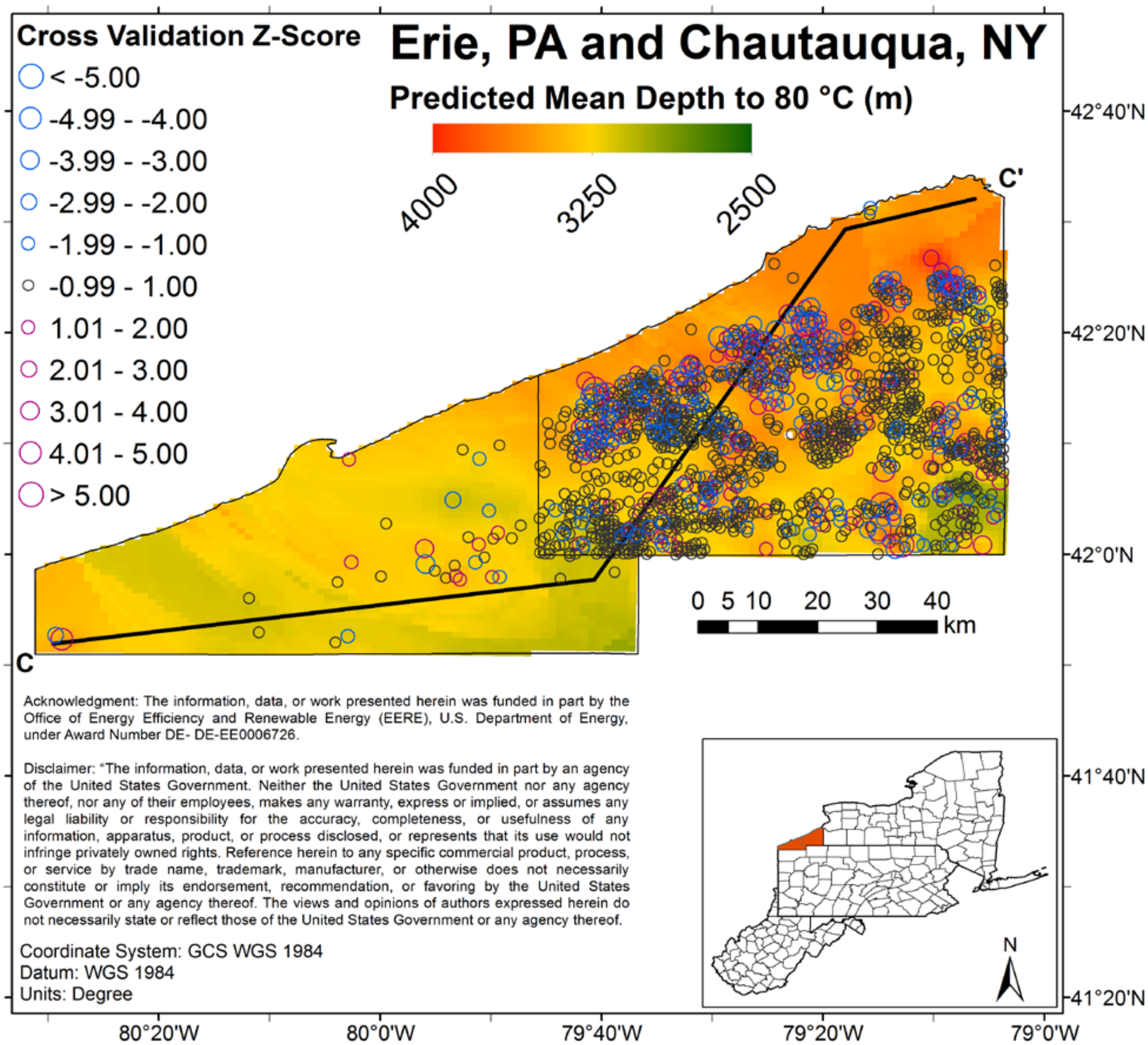
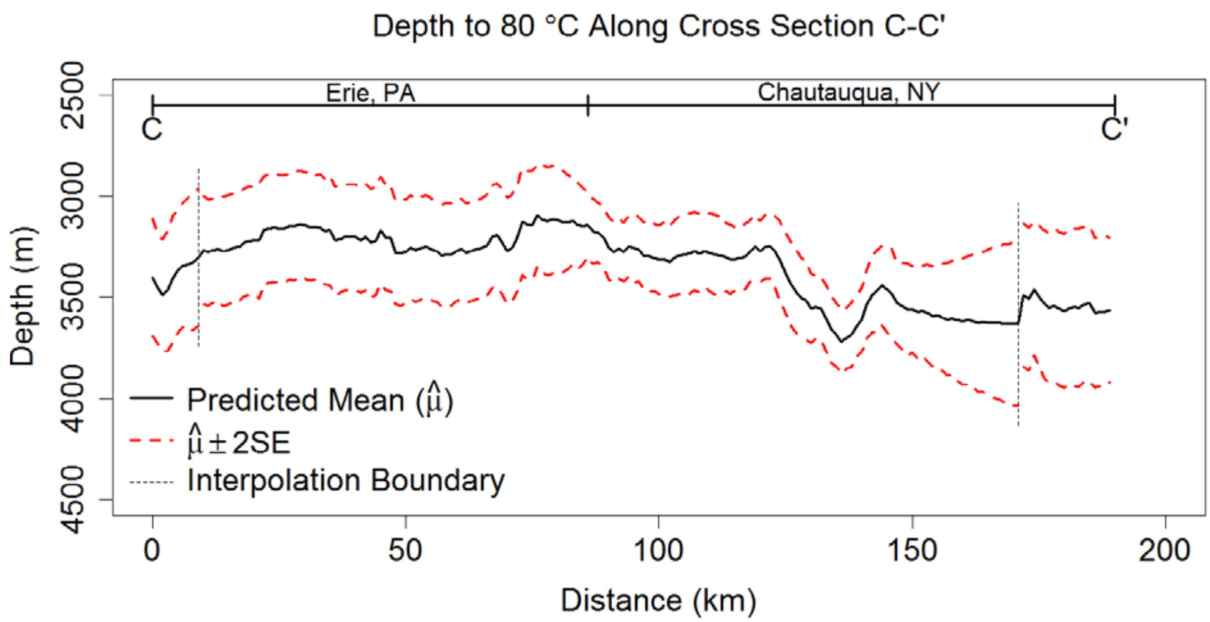


Figure 2. *Depth to 80 °C in Erie County Pennsylvania, and Chautauqua County New York. Leave one out cross validation results are shown as 'bubbles' that increase in size as the error, measured as a Z-Score, increases in magnitude. Red indicates that the value of the point was greater than the predicted mean at the location (grid cell) of the left-out point. Larger circles may indicate that the point had more influence in the spatial prediction.*



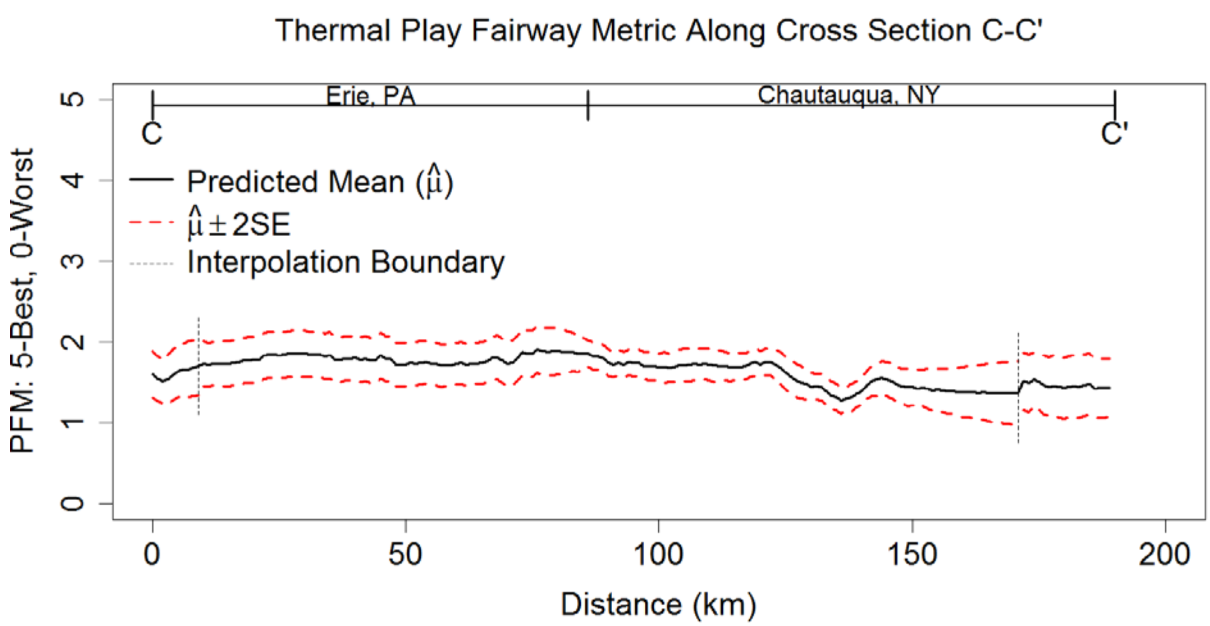


Figure 3. *Variation in depth to 80 °C and the thermal Play Fairway Metric along cross section C-C' in Figure 2. These two counties display moderate depths with relatively high certainty compared with other locations ($\hat{\mu} \pm 2SE$ spans approximately 500 m in most locations).*

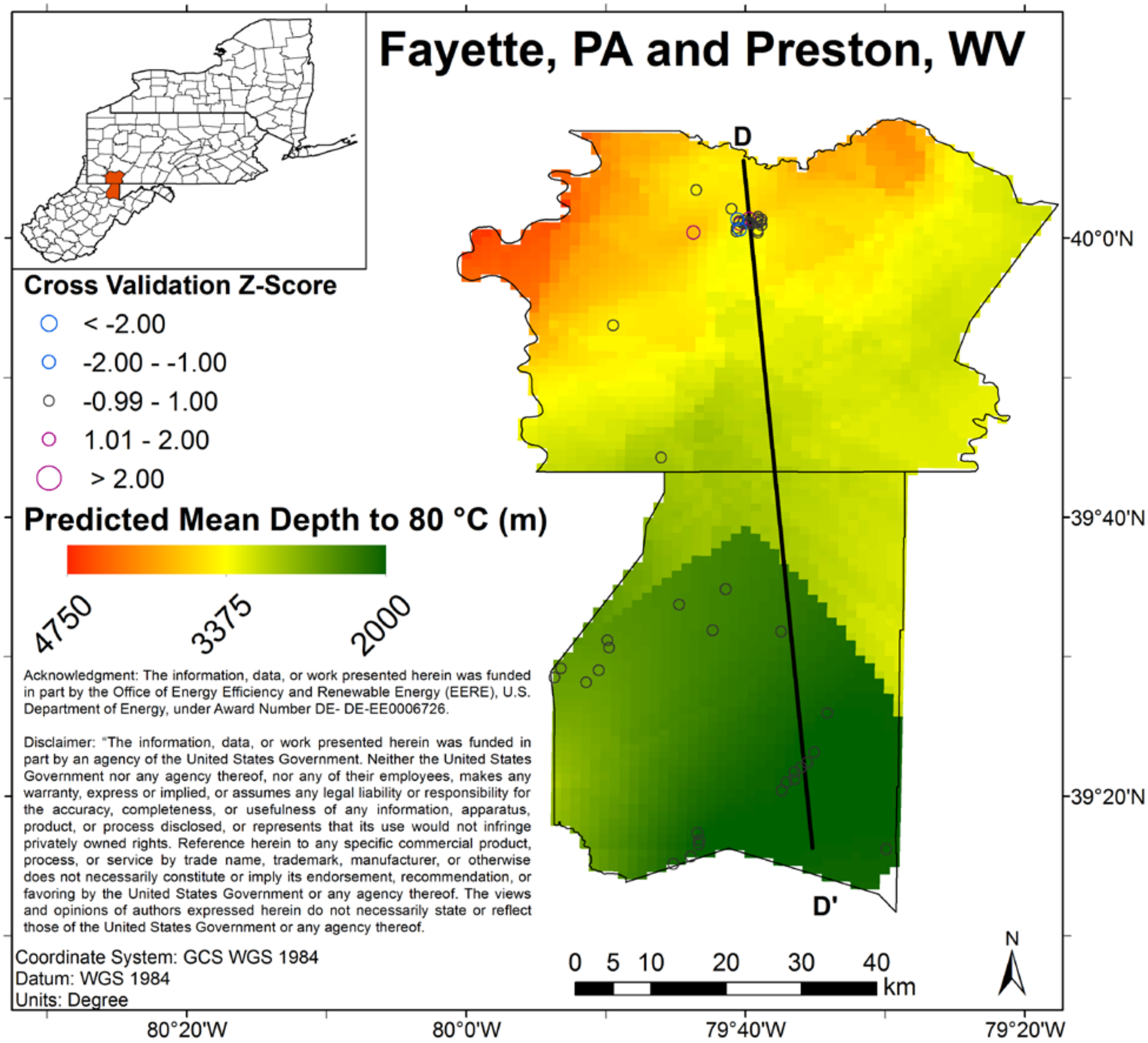
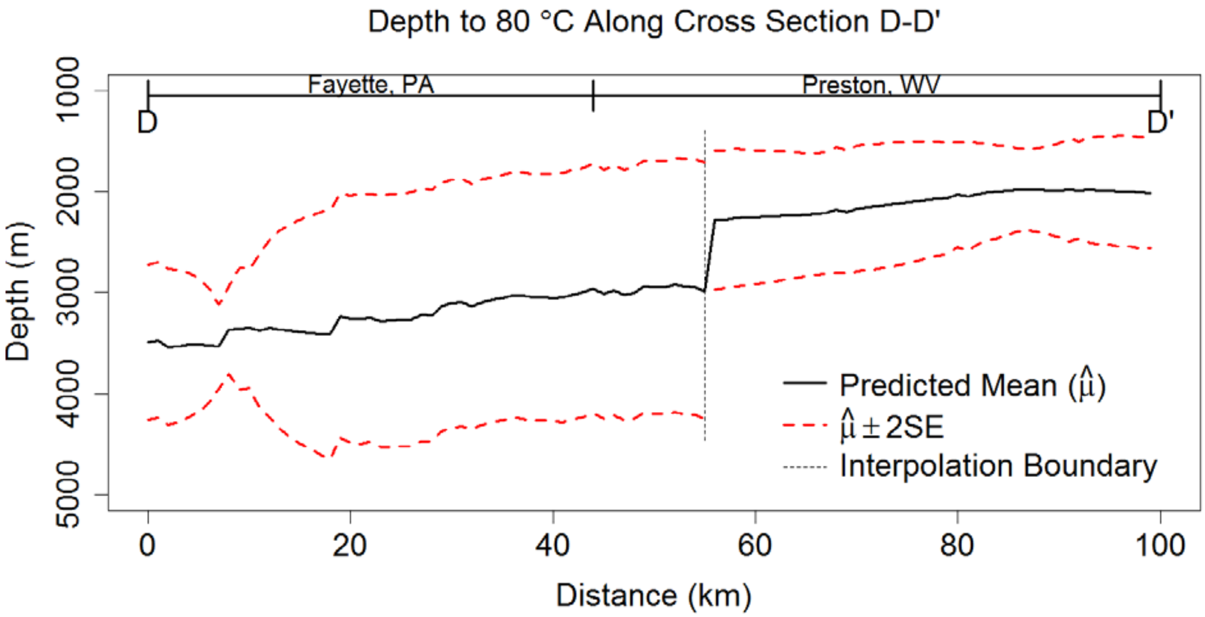


Figure 4. *Depth to 80 °C in Fayette County Pennsylvania, and Preston County West Virginia. Leave one out cross validation results are shown as 'bubbles' that increase in size as the error, measured as a Z-Score, increases in magnitude. Red indicates that the value of the point was greater than the predicted mean at the location (grid cell) of the left-out point. Larger circles may indicate that the point had more influence in the spatial prediction.*



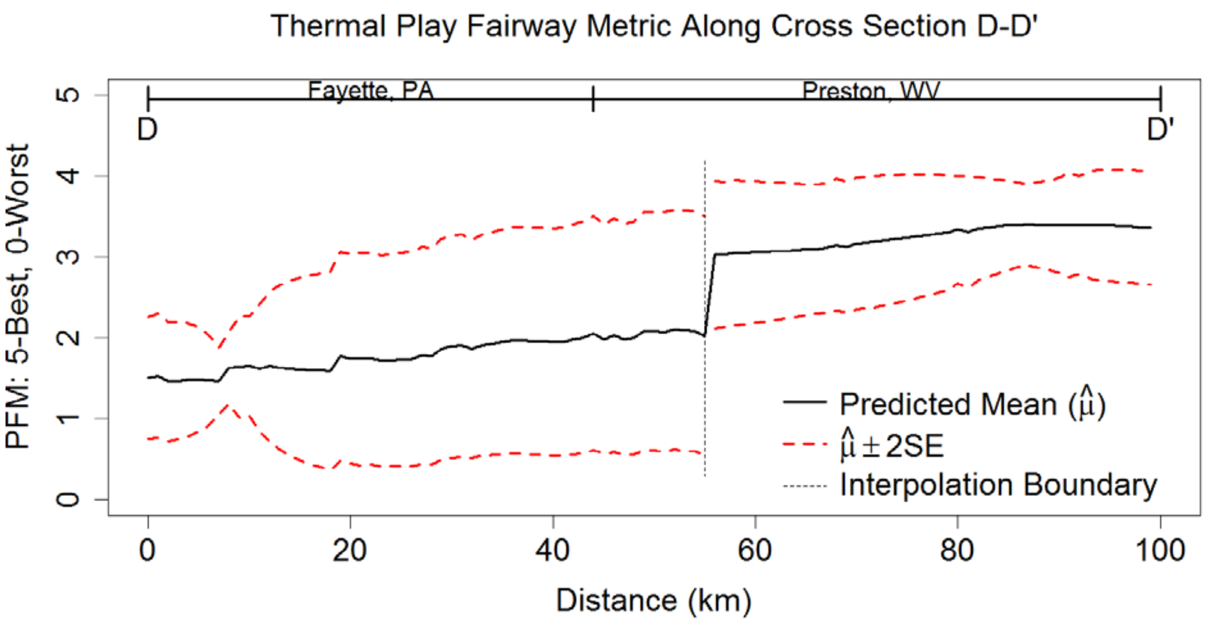


Figure 5. *Variation in depth to 80 °C and the thermal Play Fairway Metric along cross section D-D' in Figure 4. While Preston County has a higher predicted mean, it is not statistically different than the mean in Fayette County.*

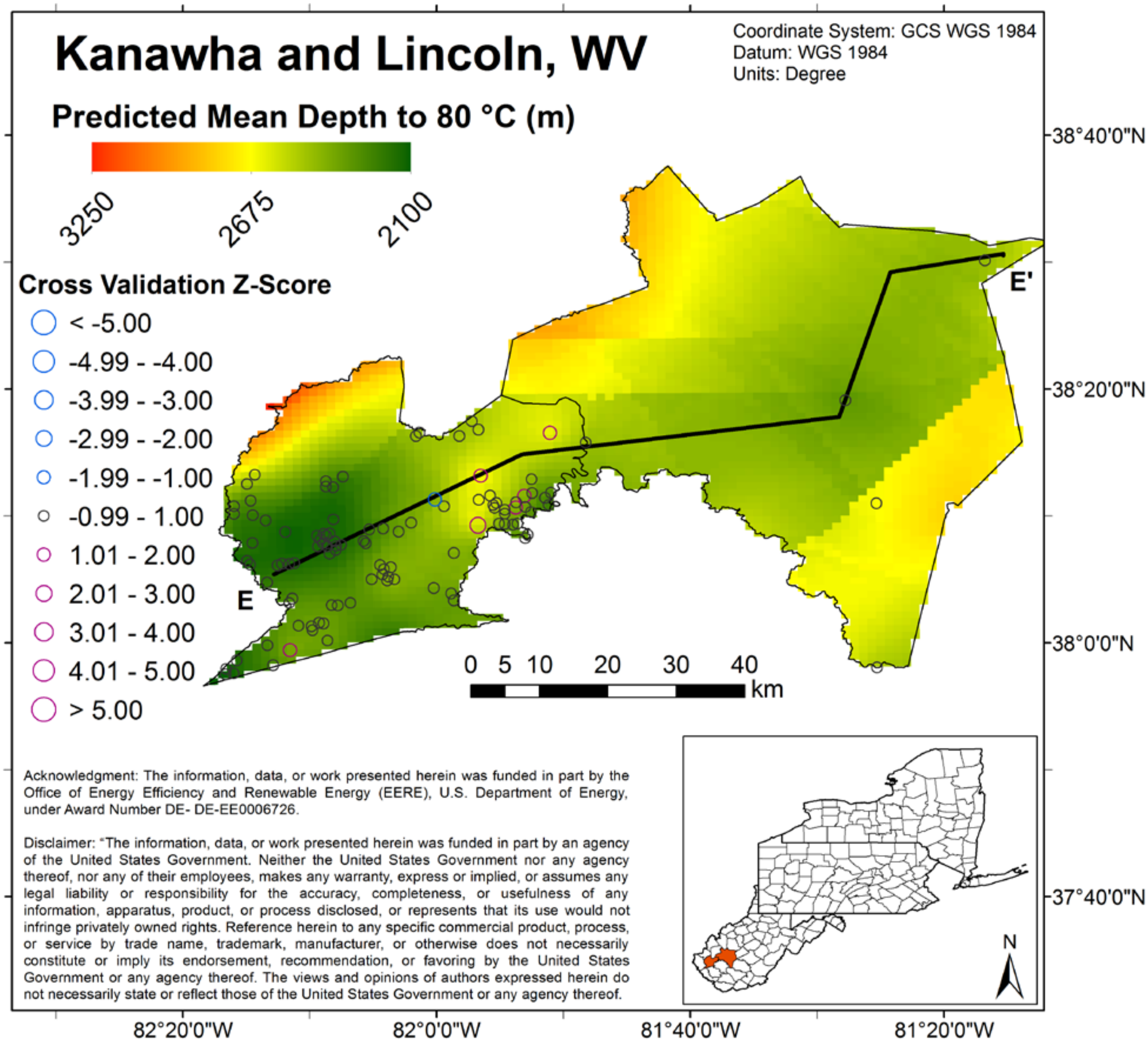


Figure 6. *Depth to 80 °C in Lincoln and Kanawha counties in West Virginia. Leave one out cross validation results are shown as 'bubbles' that increase in size as the error, measured as a Z-Score, increases in magnitude. Red indicates that the value of the point was greater than the predicted mean at the location (grid cell) of the left-out point. Larger circles may indicate that the point had more influence in the spatial prediction.*

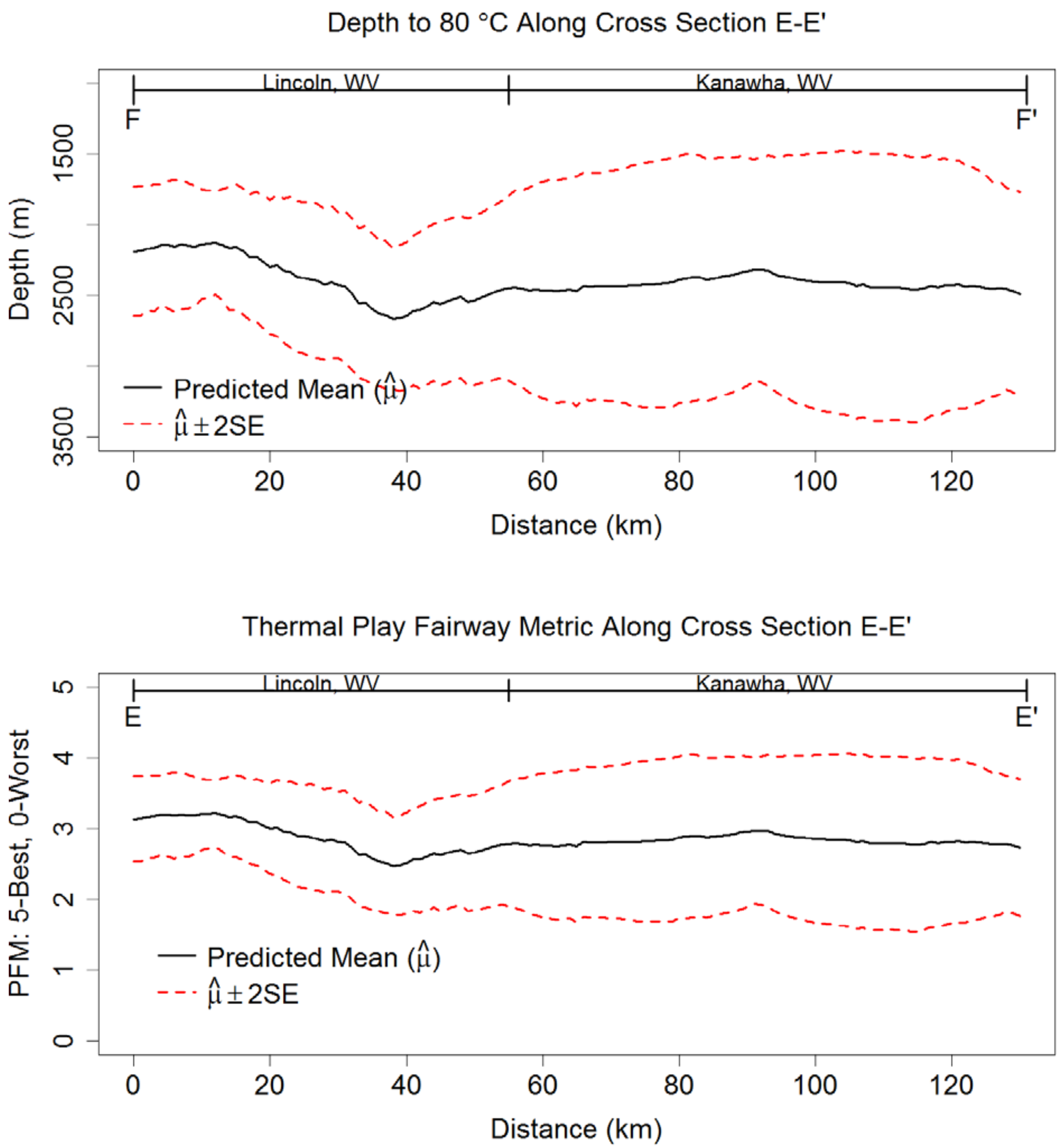


Figure 7. *Variation in depth to 80 °C and the thermal Play Fairway Metric along cross section E-E' in Figure 6. The uncertainty in the predicted mean increases along the cross section as a result of decreasing well density.*

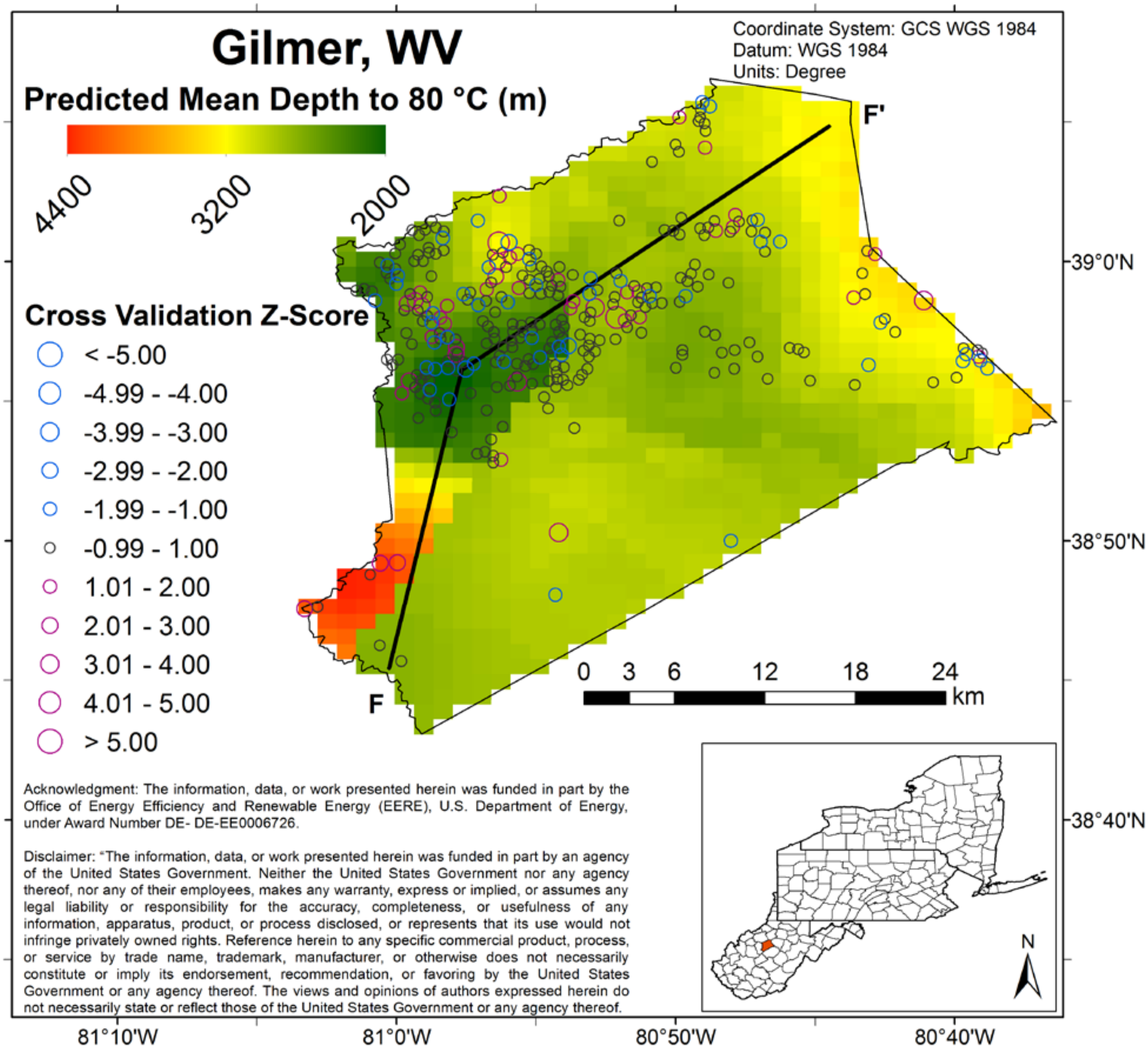


Figure 8. *Depth to 80 °C in Gilmer County West Virginia. Leave one out cross validation results are shown as 'bubbles' that increase in size as the error, measured as a Z-Score, increases in magnitude. Red indicates that the value of the point was greater than the predicted mean at the location (grid cell) of the left-out point. Larger circles may indicate that the point had more influence in the spatial prediction.*

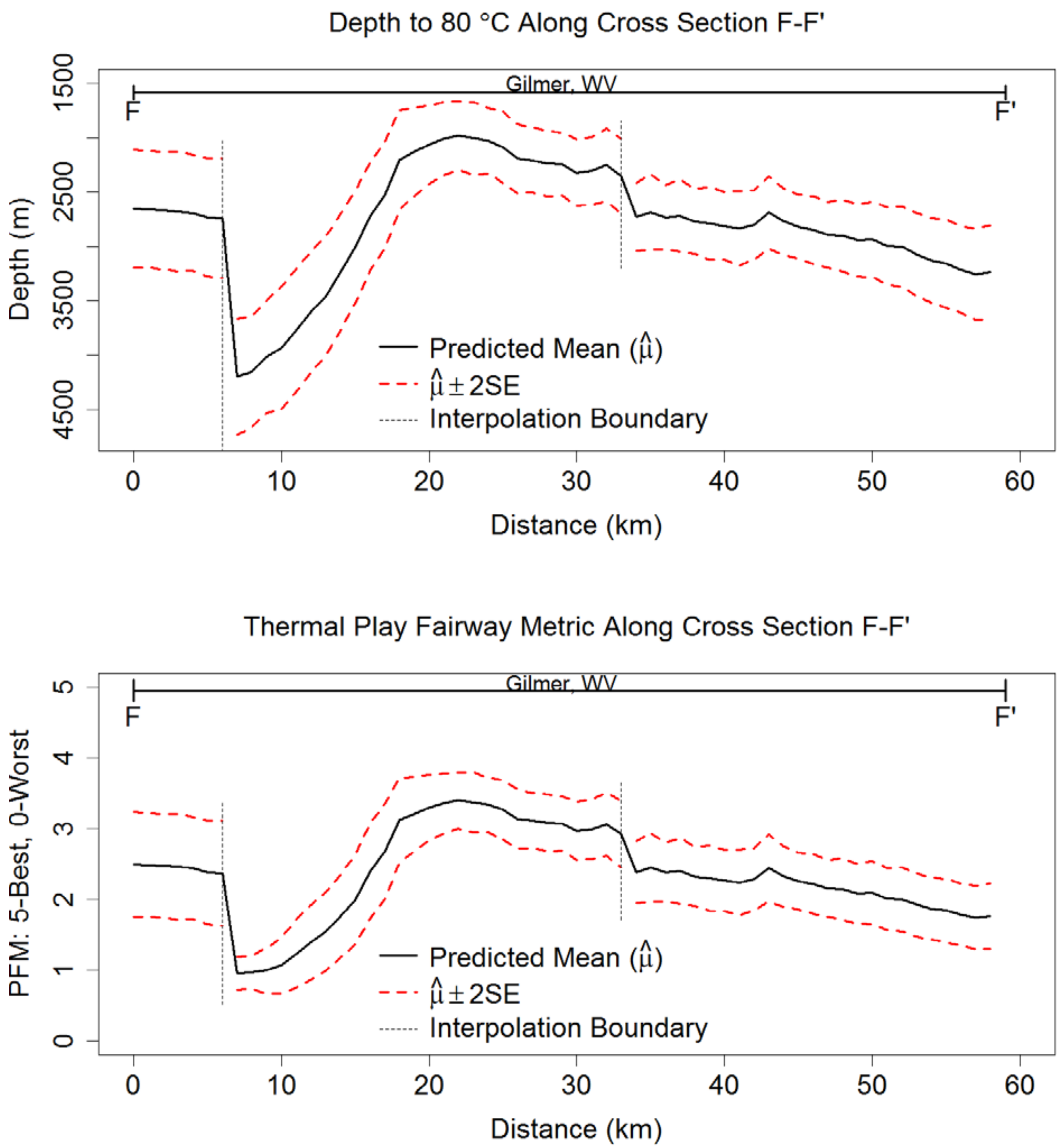


Figure 9. *Variation in depth to 80 °C and the thermal Play Fairway Metric along cross section F-F' in Figure 8. The most certain shallowest location is about 2300 m.*

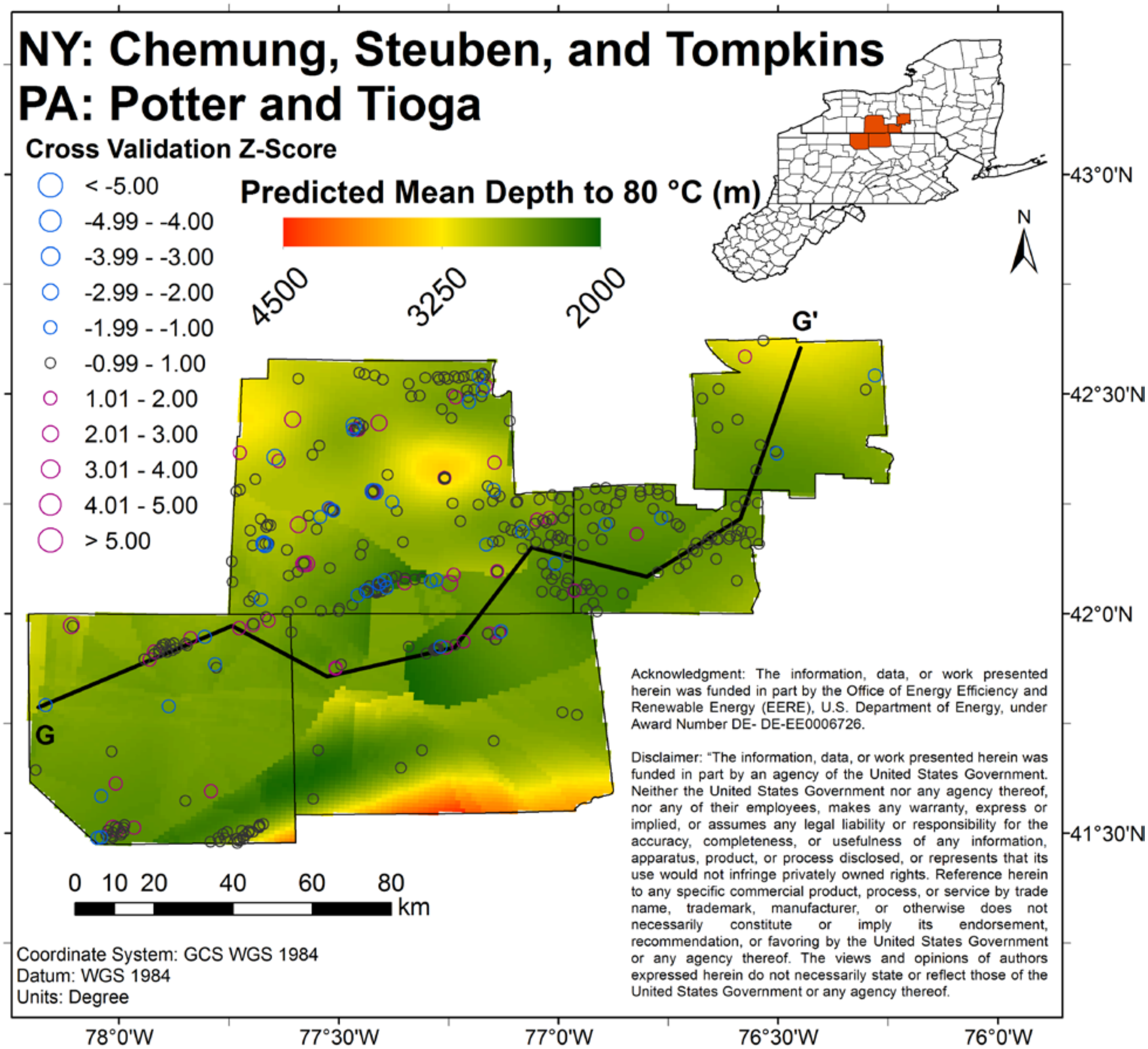


Figure 10. *Depth to 80 °C in Chemung, Steuben, and Tompkins counties in New York, and Potter and Tioga counties in Pennsylvania. Leave one out cross validation results are shown as 'bubbles' that increase in size as the error, measured as a Z-Score, increases in magnitude. Red indicates that the value of the point was greater than the predicted mean at the location (grid cell) of the left-out point. Larger circles may indicate that the point had more influence in the spatial prediction.*

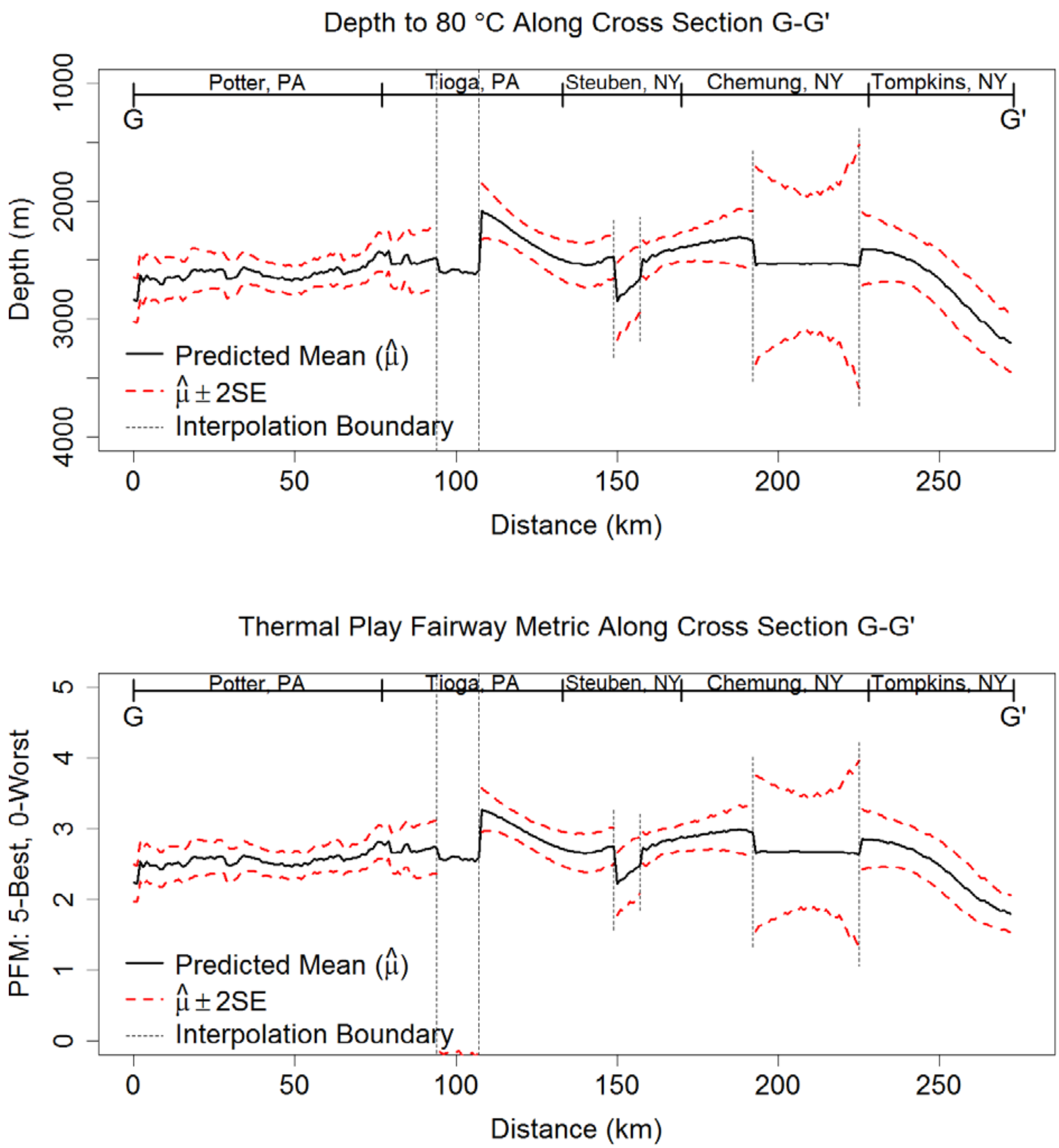


Figure 11. *Variation in depth to 80 °C and the thermal Play Fairway Metric along cross section G-G' in Figure 10. The most certain shallowest location is between Tioga through Chemung counties, which contain reservoirs and two population centers (Elmira, NY and Corning, NY).*