SPATIAL INTERPOLATION METHODS ACCURACY AND DATA REQUIREMENTS IN TOPOGRAPHIC ASSESSMENTS

KEYWORDS: KRIGING, INVERSE DISTANCE WEIGHTED, SPLINE, DEM, NATURAL NEIGHBOR, DATA RESOLUTION, SPATIAL INTERPOLATION

MONICA CASSO
DEPARTMENT OF GEOGRAPHY, THE PENNSYLVANIA STATE UNIVERSITY
March 2020

Table of Contents

Spatial Interpolation Methods Accuracy and Data Requirements in Topographic Assessments

Authors: Monica Casso, Department of Geography, The Pennsylvania State University

Topics: Interpolation Methodology, Geomorphology, Data Resolution, Development Planning
Keywords: Kriging, Inverse Distance Weighted, Spline, DEM

Abstract

Topography is widely used across a variety of applications in many industries and disciplines from the Oil and Gas Industry for infrastructure and pipeline route planning, to residential planning, communications planning, flood hazards, and others. Thus, ensuring topography is accurately captured is important when making these assessments. Although, many interpolation methods are available across the many Geographic Information System (GIS), no golden standard of which method to use for development planning exists. That likely has to do with the fact that the Earth's surface varies depending on your geographic location thus there may not be a one size fits all. This study will assess the accuracy of results for the common interpolation methods Kriging, Spline, Natural Neighbor, and IDW. Assessments will be made across different terrain types (little variability to a highly variable complex terrain) and evaluate down sampled data sets. Comparisons between methods and sample sizes will allow for better understanding of how well or poorly the interpolation methods at the various spatial resolutions did at an overall scale giving a percentage of error but also providing awareness of where interpolation methods struggled the most. These insights are important as it allows one to tailor the interpolation method chosen to the terrain type being developed as well as highlights how much or how little spatial information is needed and thus provide for cost saving opportunities.

Background

Interpolation is the process in which an estimate is made between a given set of point values to fill the gaps in between. There are a variety of different interpolation methods available and one of the common uses for interpolation is to depict the Earth's topography. The product that results from interpolating data points to achieve a continuous surface that represents the Earth's terrain is called a digital elevation model (DEM) [5]. A DEM is a raster file in which each cell has a representative elevation value [6]. A DEM's are used across many disciplines such as infrastructure and pipeline route planning in the Oil and Gas Industry, Soils Mapping, Engineering, Agriculture, and other applications. The USDA – NRCS identifies the importance of a DEM in the NRCS Digital Elevation Model (DEM) Whitepaper: NRCS High Resolution Elevation Data [7]. The whitepaper lightly touches on interpolation but does not detail the impact different interpolation methods with respect to getting an accurate representation of the Earth's topography across a different terrain type. Spatial resolution and its impact on DEM production is more heavily investigated in the whitepaper by attaining feedback across different disciplines yet more concrete evidence would be beneficial to support the need of finer resolution scales.

To acquire a fit for purpose DEM, knowledge on the terrain of the study area and what learnings are wanted to be achieved will better dictate the required accuracy and resolution needed [3]. The Earth's surface is irregular but is more so in some regions than others [4]. Some regions also very rapidly change topography motifs, such as mountainous or channelized terrain [4]. Assuming all interpolation methods will aid in achieving true representations would lead to incorrect results. Thus, it is essential to understand which interpolation method will work best for the topographic setting.

Aside from having multiple interpolation methods for production, the quality of a DEM also relies on the input resolution [5]. Because the spatial resolution of data acquisition has a direct implication on cost, it is often sacrificed in an attempt for savings. Thus, the framework of post spacing sample points is placed at greater distances which means larger void areas need to be interpolated. However, too far apart of a data acquisition spacing could lead one to miss subtle or minor changes in topography that could be critical in projects such as identifying flood hazards. It is difficult to understand how little data resolution spacing one could get away with, yet there is an assumption that it may vary more in some terrains than others. For example, a gently dipping slope with little variation may be less prone to inconsistencies of a coarser data resolution spacing acquisition than a more varying terrain such as a mountainous area or meandering river.

As previously mentioned, topography is an important piece of data across various disciplines. An inaccurate produced representation (DEM) can have negative implications on economics, feasibility, and site selection assessments. For example, if highly varying terrains are oversimplified, this could lead to underestimating development/planning costs for items such as laying down pipelines. Oversimplification of topography or other inaccurate productions of topography can also lead to poor communications facilities planning as site selection of this sort requires unobstructed locations to be able to transmit and receive signals. Topography also plays a large role on properly assessing geohazards such as flood zones or instability areas prone to debris flows due to Earthquakes.

Goals and Objectives

The intent of this analysis is to understand whether selecting the correct spatial interpolation method and/or a finer spatial resolution acquisition for the terrain of interest makes a difference on the resulting accuracy in attempts to identify cost saving opportunities across industries.

Methodology

Data Resolution Selection

The surveys resulting from the NRCS whitepaper across the different disciplines show a strong preference for 1m post spacing, an average of 2m post spacing, and less indicated that a 5m resolution could suffice their accuracy needs. These three different meter spacing variations will be tested as they cover a broad range of disciplines, from the ones that require more precision to those who need less so.

Site Selection

To test the accuracy, a ground truth dataset is needed to compare the results to. To identify a good ground truth candidate, a requirement of a 1m spatial resolution dataset is selected as that is the minimum resolution spacing being test. The next requirement would be to understand the effects on varying terrains; thus, more than one terrain type is needed resulting in multiple ground truth datasets. In this case, a total of three ground truth datasets are chosen, a terrain with minimal topographic variation (a relatively flat surface area), a terrain with some variation (a relatively flat area with some minor abrupt topographic change, a channelized area), and a terrain with diaristic variation (a highly varying topographic profile such as a mountainous area). The final requirement has less to do with the technical data itself rather to do with data availability and cost, for the purposes of this study only publicly available data is used. The data used for this study can be found in Table 1.

Table 1Datasets and Sources

Dataset	Description of Dataset	Terrain Type
1-meter DEM Citation: Citation_Information: Originator: U.S. Geological Survey Publication_Date: 20181018 Title: USGS NED one meter x73y438 CO Central Western 2016 IMG 2018 Geospatial_Data_Presentation_Form: raster digital data Publication_Information: Publication_Place: Reston, VA Publisher: U.S. Geological Survey Online_Linkage: http://nationalmap.gov/elevation.html Online_Linkage: http://nationalmap.gov/viewer.html	AOI Garflied Colorado Spatial_Reference_Information: Horizontal_Coordinate_System_Definition: Planar: Grid_Coordinate_System_Name: Transverse_Mercator Universal_Transverse_Mercator: UTM_Zone_Number: 30 Transverse_Mercator: Scale_Factor_at_Central_Meridian: 0.9996 Longitude_of_Central_Meridian: 0.0 Latitude_of_Projection_Origin: 0.0 False_Easting: 500000.0 False_Northing: 0.0	Terrain with diaristic variation (a highly varying topographic profile such as a mountainous area).
1-meter DEM Citation: Citation_Information: Originator: U.S. Geological Survey Publication_Date: 20180724 Title: USGS NED one meter x66y368 AR Ouachita B1 2016 IMG 2018 Geospatial_Data_Presentation_Form: raster digital data Publication_Information: Publication_Place: Reston, VA Publisher: U.S. Geological Survey Online_Linkage: http://nationalmap.gov/elevation.html Online_Linkage: http://nationalmap.gov/viewer.html	AOI Ouachita, Arkansas Spatial_Reference_Information: Horizontal_Coordinate_System_Definition: Planar: Grid_Coordinate_System_Name: Transverse_Mercator Universal_Transverse_Mercator: UTM_Zone_Number: Transverse_Mercator: Scale_Factor_at_Central_Meridian: 0.9996 Longitude_of_Central_Meridian: 0.0 Latitude_of_Projection_Origin: 0.0 False_Easting: 500000.0 False_Northing: 0.0	Terrain with some variation (a relatively flat area with some minor abrupt topographic change, a channelized area)
1-meter DEM Citation: Citation_Information: Originator: U.S. Geological Survey Publication_Date: 20191018 Title: USGS NED one meter x23y353 TX West Central B7 2018 IMG 2019 Geospatial_Data_Presentation_Form: raster digital data Publication_Information: Publication_Place: Reston, VA Publisher: U.S. Geological Survey Online_Linkage: http://nationalmap.gov/elevation.html Online_Linkage: http://nationalmap.gov/viewer.html	AOI West Texas (Near Midland and Odessa) Spatial_Reference_Information: Horizontal_Coordinate_System_Definition: Planar: Grid_Coordinate_System_Name: Transverse_Mercator Universal_Transverse_Mercator: UTM_Zone_Number: 30 Transverse_Mercator: Scale_Factor_at_Central_Meridian: 0.9996 Longitude_of_Central_Meridian: 0.0 Latitude_of_Projection_Origin: 0.0 False_Easting: 500000.0 False_Northing: 0.0	Terrain with minimal topographic variation (a relatively flat surface area)

Data Preparation

The ground truth datasets chosen are meant to keep consistency and lessen the errors of post processing effects between different data acquisition methods. The three datasets are cropped to a smaller AOI that keeps characteristics of the terrain to reduce processing time during interpolation. From the cropped datasets a subset of sample points is taken to mimic the data resolution sizes desired. This is achieved via the batch Fishnet Geoprocessing tool where a standard post spacing size is set to the 1x1m, 2x2m, and 5x5m as discussed previously. The values are then extracted from the ground truth data set per terrain type at each point using the Extract Value Geoprocessing tool. These are the nine base datasets from which the interpolation method will be tested from, thus four copies are obtained to undergo each interpolation method. The interpolation is run from interpolation tools offered by Esri in ArcGIS Pro, IDW, Kriging, Natural Neighbor, and Spline from which a total of 36 topographic realizations are produced. Table 2 briefly overviews Esri's description for each interpolation method. In this case study, all the defaults are taken for the interpolation method options in IDW, Natural Neighbor, and Spline interpolators. For the Kriging interpolator, the optimized settings from the Geostatistical Analyst were used to give a better result rather than blindly making assumptions on what the parameters should be. The optimized settings were estimated using the ground truth 1-meter resolution values for each respective terrain setting.

Table 2 Interpolation Methods as Described by Esri. More information can be found at https://pro.arcgis.com/en/pro-app/tool-reference/3d-analyst/an-overview-of-the-raster-interpolation-toolset.htm

Interpolation Method	Description
NATURAL NEIGHBOR	Finds the closest subset of input samples to a query point and applies weights to them based on proportionate areas to interpolate a value (Sibson 1981)
SPLINE	Estimates values using a mathematical function that minimizes overall surface curvature, resulting in a smooth surface that passes exactly through the input points.
INVERSE-DISTANCE WEIGHTED (IDW)	Determines cell values using a linearly weighted combination of a set of sample points.
KRIGING	Kriging is an advanced geostatistical procedure that generates an estimated surface from a scattered set of points with z-values.

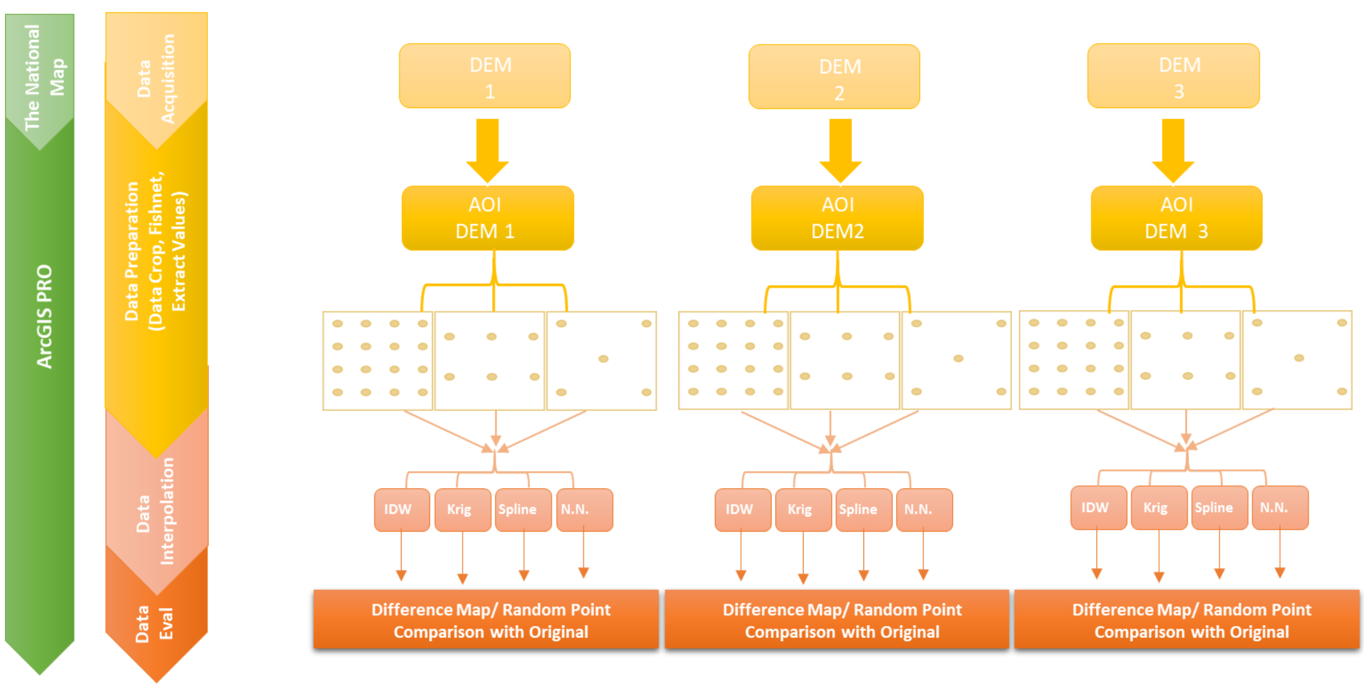

Figure 1 Workflow Breakdown

Data Evaluation

The resulting surfaces then undergo comparative analysis against the original DEM. The original 1m fishnet points will be used to sample each surface respectively (the original DEM plus the newly interpolated surfaces for the three different terrain profiles at three different resolutions). These extracted values from the interpolated surface points will be plotted against the original DEM random point set extracted values to understand which interpolation methodology and their coupled different resolution sizes give the best correlation. In addition, difference maps of the interpolated surfaces from the original DEM are used as a visual aid to give ideas if and where the interpolation methods may have struggled.

Anticipated Results

It was expected that the point set with higher resolution will result in a closer match to the original input as there are more data points available to keep fidelity to the true structured surface. Meaning, it would be less likely to produce an artificial bullseye effect that would be more commonly seen in more sparse datasets. As for data interpolation, kriging tends to be a commonly used method when it comes to geology as spatially correlated distances/ directional bias are often a key player in this discipline [1]. The natural neighbor methodology is the recommended ESRI approach for building DEM's from LAS and terrain datasets as it provides a smoother generated surface compared to a linearly derived surface. However as stated previously, Earth's topography is not smooth and therefore this may result in an overly simplified result.

Findings

Max Value

Minimal Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		max value				Percent Difference		
Interpolator	Idw	826.0073	826.0073	825.9214	825.8279	0.0000%	0.0104%	0.0217%
	Kriging	826.0073	825.9559	825.8876	825.8120	0.0062%	0.0145%	0.0237%
	Natural Neighbors	826.0073	826.0073	825.9214	825.8279	0.0000%	0.0104%	0.0217%
	Spline	826.0073	826.0073	825.9214	825.8291	0.0000%	0.0104%	0.0216%

Table 3 This table defines the maximum value found for each interpolated raster dataset at the various tested resolutions in a minimally changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Slight Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		max value				Percent Difference		
Interpolator	Idw	37.2332	37.2332	37.2332	36.9612	0.0000%	0.0000%	0.7331%
	Kriging	37.2332	37.2332	37.2332	36.9612	0.0000%	0.0000%	0.7331%
	Natural Neighbors	37.2332	37.2332	37.2332	36.9612	0.0000%	0.0000%	0.7331%
	Spline	37.2332	37.2324	37.2332	37.0338	0.0021%	0.0000%	0.5368%

Table 4 This table defines the maximum value found for each interpolated raster dataset at the various tested resolutions in a slightly changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Drastic Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		max value				Percent Difference		
Interpolator	Idw	2499.4253	2499.4253	2499.0884	2499.0508	0.0000%	0.0135%	0.0150%
	Kriging	2499.4253	2499.4253	2499.1431	2499.0508	0.0000%	0.0113%	0.0150%
	Natural Neighbors	2499.4253	2499.4253	2499.0750	2499.0508	0.0000%	0.0140%	0.0150%
	Spline	2499.4253	2499.4253	2499.3254	2499.5300	0.0000%	0.0040%	0.0042%

Table 5 This table defines the maximum value found for each interpolated raster dataset at the various tested resolutions in a drastically changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Only the Spline interpolator at 5-meters ever goes above the Max Value (height) of the Original dataset and is only seen in the drastic terrain variations (Tables 3-5 & charts 1-3). Otherwise, the 5-meters resolution underestimates the Max Values in all terrains and interpolator combinations (Tables 3-5 & charts 1-3). The Max Value can be attained in all terrain environments consistently with the IDW and Natural Neighbors interpolators at a 1-meter resolution (Tables 3-5 & charts 1-3). In slight terrain variations, the max value can be achieved closely at both 1- and 2-meter resolutions for all interpolators tested, except for the spline interpolator at 1-m resolution (Table 4 & chart 2).

Minimum Value

Minimal Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		min value				Percent Difference		
Interpolator	Idw	797.6309	797.6650	797.6650	797.7391	0.0043%	0.0043%	0.0136%
	Kriging	797.6309	797.6734	797.6650	797.7487	0.0053%	0.0043%	0.0148%
	Natural Neighbors	797.6309	797.6650	797.6650	797.7391	0.0043%	0.0043%	0.0136%
	Spline	797.6309	797.6370	797.6650	797.7391	0.0008%	0.0043%	0.0136%

Table 6 This table defines the minimum value found for each interpolated raster dataset at the various tested resolutions in a minimally changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Slight Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		min value				Percent Difference		
Interpolator	Idw	28.1465	28.1465	28.1465	28.1648	0.0000%	0.0000%	0.0650%
	Kriging	28.1465	28.1465	28.1465	28.1648	0.0000%	0.0000%	0.0650%
	Natural Neighbors	28.1465	28.1465	28.1465	28.1648	0.0000%	0.0000%	0.0650%
	Spline	28.1465	28.1465	28.1465	27.9969	0.0000%	0.0000%	0.5329%

Table 7 This table defines the minimum value found for each interpolated raster dataset at the various tested resolutions in a slightly changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Drastic Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		min value				Percent Difference		
Interpolator	Idw	1706.0232	1706.0232	1706.0574	1706.0593	0.0000%	0.0020%	0.0021%
	Kriging	1706.0232	1706.0232	1705.3094	1706.0593	0.0000%	0.0418%	0.0021%
	Natural Neighbors	1706.0232	1706.0232	1706.0277	1706.0593	0.0000%	0.0003%	0.0021%
	Spline	1706.0232	1706.0232	1705.5210	1705.5100	0.0000%	0.0294%	0.0301%

Table 8 This table defines the minimum value found for each interpolated raster dataset at the various tested resolutions in a drastically changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

In a minimal terrain variation environment, all interpolators overestimate the Min Value (height) (Table 6 & chart 4). In slight terrain variations, the Min Value can be achieved closely at 1- and 2-meter resolutions for all interpolators tested (Table 7 & chart 5). In this setting, a 5-meter spacing is only off by approximately 0.065% in all interpolators except spline where the difference of the Min Values can off by a 0.533% difference (Table 7). The drastic terrain variations environments follow the norm where the 1-meter resolution produced from any of the interpolators provide the best results (Table 8 & Chart 6). However, a close Min Value of no more than a 0.0021% difference from the original dataset can be attained by both IDW and Natural Neighbors interpolators at the coarser resolutions of 2- and 5-meters (Table 8 & Chart 6).

Mean Value

Minimal Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		mean value				Percent Difference		
Interpolator	Idw	811.612	811.621	811.611	811.618	0.00121%	0.00002%	0.00078%
	Kriging	811.612	811.621	811.611	811.618	0.00121%	0.00002%	0.00078%
	Natural Neighbors	811.612	811.621	811.611	811.612	0.00121%	0.00002%	0.00006%
	Spline	811.612	811.621	811.611	811.618	0.00121%	0.00002%	0.00078%

Table 9 This table defines the mean value found for each interpolated raster dataset at the various tested resolutions in a minimally changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Slight Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		mean value				Percent Difference		
Interpolator	Idw	34.1150	34.1150	34.1135	34.1054	0.0000%	0.0046%	0.0283%
	Kriging	34.1150	34.1150	34.1135	34.1053	0.0000%	0.0046%	0.0284%
	Natural Neighbors	34.1150	34.1150	34.1135	34.1054	0.0000%	0.0046%	0.0283%
	Spline	34.1150	34.1150	34.1135	34.1055	0.0000%	0.0046%	0.0279%

Table 10 This table defines the mean value found for each interpolated raster dataset at the various tested resolutions in a slightly changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Drastic Terrain Changes		Resolution						
		Original	1m	2m	5m	1m	2m	5m
		mean value				Percent Difference		
Interpolator	Idw	1990.9501	1990.8951	1990.8837	1990.8950	0.0028%	0.0033%	0.0028%
	Kriging	1990.9501	1990.8951	1990.7815	1990.6594	0.0028%	0.0085%	0.0146%
	Natural Neighbors	1990.9501	1990.8951	1990.7815	1990.6594	0.0028%	0.0085%	0.0146%
	Spline	1990.9501	1990.8951	1990.8838	1990.8952	0.0028%	0.0033%	0.0028%

Table 11 This table defines the mean value found for each interpolated raster dataset at the various tested resolutions in a drastically changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

In a minimal terrain variation environment, all interpolators provide the least percent difference for the Mean Value at a 2-meter resolution in comparison to the original dataset (Table 12 & Chart 7). This statement is invalid for the other variations of terrain tested where the least percent difference is no greater than 0.002% and can be commonly produced using a 1-meter resolution and any of the test interpolators (Tables 12-14 & Charts 7-9). In drastic variation of terrain environment, the percent difference at 2-meter resolution for Kriging and Natural Neighbor interpolates are higher than at the same resolution in the other tested less variable terrain settings.

Standard Deviation

Interpolator	Minimal Terrain Changes	Original	1m	2m	5m	1m	2m	5m
		Resolution						
		Standard Deviation value				Percent Difference		
Interpolator	Idw	6.10277	6.09856	6.09455	6.07916	0.0689%	0.1347%	0.3876%
	Kriging	6.10277	6.09853	6.09460	6.07954	0.0695%	0.1340%	0.3812%
	Natural Neighbors	6.10277	6.09856	6.09461	6.07675	0.0689%	0.1338%	0.4272%
	Spline	6.10277	6.09856	6.09464	6.07953	0.0689%	0.1333%	0.3815%

Table 12 This table defines the Standard Deviation value found for each interpolated raster dataset at the various tested resolutions in a minimally changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Interpolator	Slight Terrain Changes	Original	1m	2m	5m	1m	2m	5m
		Resolution						
		Standard Deviation value				Percent Difference		
Interpolator	Idw	1.87048	1.87048	1.86399	1.84308	0.0000%	0.3476%	1.4760%
	Kriging	1.87048	1.87048	1.85855	1.80835	0.0000%	0.6398%	3.3776%
	Natural Neighbors	1.87048	1.87048	1.86759	1.85704	0.0000%	0.1550%	0.7211%
	Spline	1.87048	1.87049	1.87031	1.87108	0.0002%	0.0093%	0.0321%

Table 13 This table defines the Standard Deviation value found for each interpolated raster dataset at the various tested resolutions in a slightly changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

Interpolator	Drastic Terrain Changes	Original	1m	2m	5m	1m	2m	5m
		Resolution						
		Standard Deviation value				Percent Difference		
Interpolator	Idw	197.2571	197.2385	197.2274	197.2274	0.009%	0.015%	0.015%
	Kriging	197.2571	197.2385	197.1776	197.1085	0.009%	0.040%	0.075%
	Natural Neighbors	197.2571	197.2385	197.1763	197.1128	0.009%	0.041%	0.073%
	Spline	197.2571	197.2385	197.2294	197.2381	0.009%	0.014%	0.010%

Table 14 This table defines the Standard Deviation value found for each interpolated raster dataset at the various tested resolutions in a drastically changing terrain setting. The original values shown are not interpolated, they are the values from the uncompromised original dataset for comparison. The last three columns highlight the calculated percent difference between the interpolated and original dataset respectively.

In all terrain environments and interpolators, the variance of the interpolated datasets is never greater than the original dataset (Tables 9-11 & Charts 10-12). However, they all generally have a smaller spread (Tables 9-11 & Charts 10-12). The only terrain type that provides results that are the same spread as the original dataset is that of a slight variation in terrain (Table 10 and Chart 11). The only interpolator in this environment capturing the equivalent spread to the tenths measurement at every resolution is the spline interpolator (Chart 11).

Correlation Coefficient

Correlation Coefficient		Resolution		
		1m	2m	5m
Minimal Terrain Changes	Idw	1	0.9999937350	0.9999799993
	Kriging	0.9999963316	0.9999925539	0.9999810272
	Natural Neighbors	1	0.9999948932	0.9999825781
	Spline	0.9999999971	0.9999947971	0.9999807095
Slight Terrain Changes	Idw	1	0.9945411986	0.9945891186
	Kriging	1	0.9939485488	0.9825514100
	Natural Neighbors	1	0.9946075687	0.9972785962
	Spline	0.9999999402	0.9944898389	0.9981998437
Drastic Terrain Changes	Idw	0.9999999998	0.9999961671	0.9999942318
	Kriging	0.9999999998	0.9999937209	0.9999947924
	Natural Neighbors	0.9999999998	0.9999938168	0.9999971358
	Spline	0.9999977908	0.9999967814	0.9999977908

Table 15 This table summarizes the correlation coefficient values attained from plotting the interpolated values against the original dataset values. The closer a number is to the value of 1 indicates a stronger match. Where the value is 1, it is a like for like value, in other words you have an exact match.

Only the IDW and Natural Neighbors interpolators at a 1-meter resolution in a minimal and slight terrain variation setting were able to create a like for like match of the original dataset (Table 15 and Figures 14 – 16). The Kriging interpolator was able to reach a correlation coefficient of 1 at a 1-meter resolution but only in a slight terrain variation setting (Table 15 and Chart 15). All other interpolators were able to reach a correlation coefficient of at least 0.99 to varying degrees (Table 15 and Figures 14 – 16). None of the interpolators at any resolution scale were able to attain a correlation coefficient of 1 at a drastic terrain variation (Table 15 and Figure 16). However, the least amount of variation in comparison to the original was found in the 1-meter resolutions with the IDW, Natural Neighbors, and Kriging interpolators at a drastic terrain variation (Table 15).

Discussion

It was thought that a drastic change in elevation setting would be the most challenging for an interpolator to achieve an accurate representation as there would be more dissimilarities across the area of interest. Looking the scatter plots indicates otherwise. In comparing the different settings in Figures 14-16, there is much more variance in the slight terrain settings at the 5- and 2-meter spacings. This may be attributed because the terrain setting chosen is more of an abrupt change rather than a gradual or continuous change in elevation. Even though a mountainous area consists of very dramatic changes over the landscape, it's not abrupt and follows a pattern. Whereas a river, a meandering river in this case isn't following a straight path thus the points surrounding it may not be as close to the elevation value as extracted point as you would expect.

In the minimal terrain variation dataset, there seemed to be a commonality of where most of the interpolators were struggling. Looking at Figures 17-19 it is evident that where the data values would start to converge such as forming a very subtle valley or wherever the surface is not linearly evenly graded, the errors of the interpolators tend to stand out more, either underestimating as shown in reds or overestimating as shown by the variations of blue.

Recommendations

As expected, all interpolators do relatively better at a 1-meter resolution. Overall across different aspects (min/max/mean/std. dev values and correlation coefficient), the IDW and Natural Neighbors approach seem to provide consistently more accurate results across all terrains at 1-meter resolution. The proposed USDA – NRCS data collection method for this resolution scale is via Lidar and the recommended DEM production methodology by Esri uses Natural Neighbors, thus this study further endorses that choice of interpolation at a 1-meter resolution. Note that Natural Neighbors does have a longer associated run time and as Esri states, the resulting surface is smoother than what a linear interpolation would provide and is less vulnerable to small changes in the triangulation [2]. Thus, if you are on a time crunch using IDW may be better for your needs, as discussed previously both methods provide similarly accurate level results so you aren't compromising your results like you would if you used a Spline interpolation method. The kriging approach is often treated as a good candidate but since you are invoking some sort of directional biasing to the product this leaves more room for error when your terrain is starting to vary as seen in Figures14-16.

To save costs, one could even use a 2- or 5-meter spacing, even though the accuracy begins to diminish it's not giving results less than a 0.9999 correlation coefficient, this is only true for minimal and drastic variation environments. For areas that have slight variation where those changes are abrupt such as channelized settings, a 1-meter resolution is most likely preferred if precision is key in studies such as flood hazard analysis. Other general studies may be able to get away with the 2- or 5-m resolution but would need to keep in mind the uncertainty they will have in their DEM.

One caveat to this analysis is that the ground truth data resolution could only be sampled down to a 1-meter spacing due to the publicly data available for this analysis. To get a better/more realistic feel of how well the spatial interpolators do at a 1-meter resolution, it would be ideal of have a finer resolution that is then down sampled into 1-meter resolution and interpolated. Currently, the 1-meter interpolations are more of a like for like example and you would expect as seen that your values due to choosing a 1-meter cell by 1-meter grid cell are producing very accurate if not the same results.

Technical difficulties were initially encountered due to the large amount of data captured in the original extent of the AOI for each DEM. The datasets had to be significantly reduced by cropping the data down to a more manageable size to get a decent runtime for the different interpolations. A suggestion for production of a DEM is to ensure access to a windows blade or other mechanism to allow for such high computing is available.

References

[1] Childs, Colin. 2004. “Interpolating Surfaces in ArcGIS Spatial Analyst.” ArcUser. Redlands, CA: ESRI Press.

[2] Esri. Creating raster DEMs and DSMs from large lidar point collections. Retrieved January 12, 2020, from https://desktop.arcgis.com/en/arcmap/10.3/manage-data/las-dataset/lidar-solutions-creating-raster-dems-and-dsms-from-large-lidar-point-collections.htm

[3] Florinsky, Igor V. “Digital Elevation Models.” Digital Elevation Models - an Overview | ScienceDirect Topics, 2012, www.sciencedirect.com/topics/earth-and-planetary-sciences/digital-elevation-models.

[4] Lisle, Richard J. Geological Structures and Maps: a Practical Guide. Elsevier Butterworth-Heinemann, 2004.

[5] Miller, Robert Lee, 1920-, and James Steven Kahn. Statistical Analysis in the Geological Sciences. New York: Wiley, 1962.

[6] USGS. What Are Digital Elevation Models (DEMs)?, www.usgs.gov/faqs/what-are-digital-elevation-models-dems.

[7] NRCS. Digital Elevation Model (DEM) Whitepaper NRCS High Resolution Elevation Data, 2011. https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1047930.pdf

Charts and Figures

See following pages.

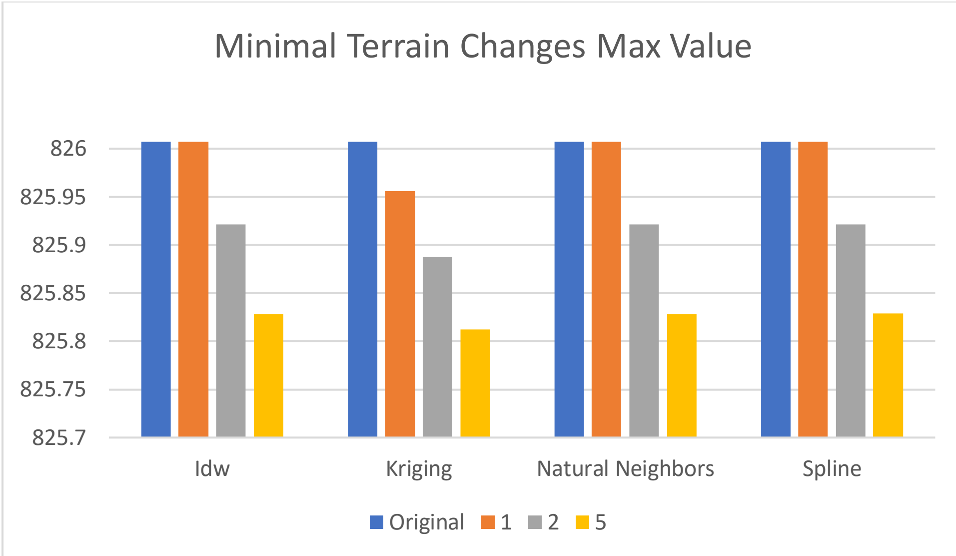

Chart 1 Differences in the interpolators produced maximum values for Minimal Terrain Variation Enviornments

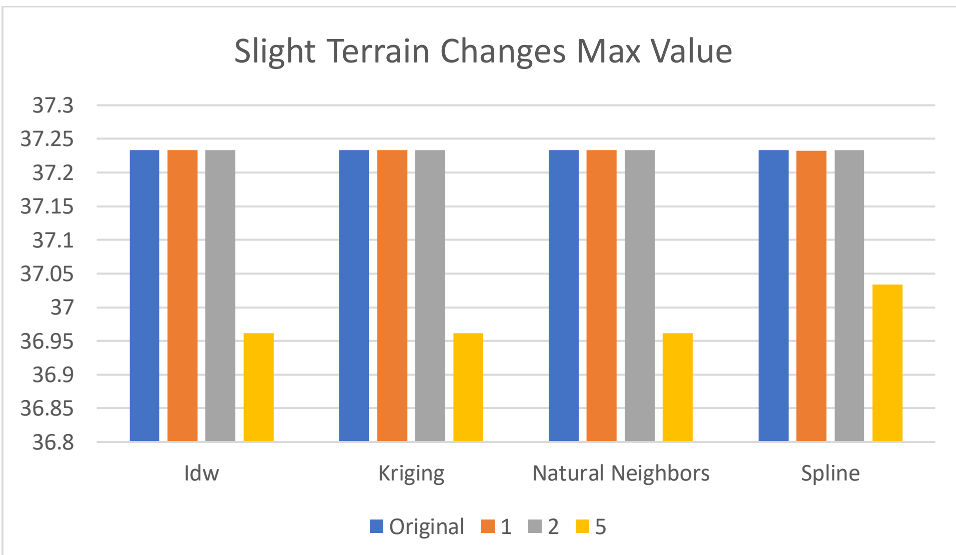

Chart 2 Differences in the interpolators produced maximum values for Slight Terrain Variation Enviornments

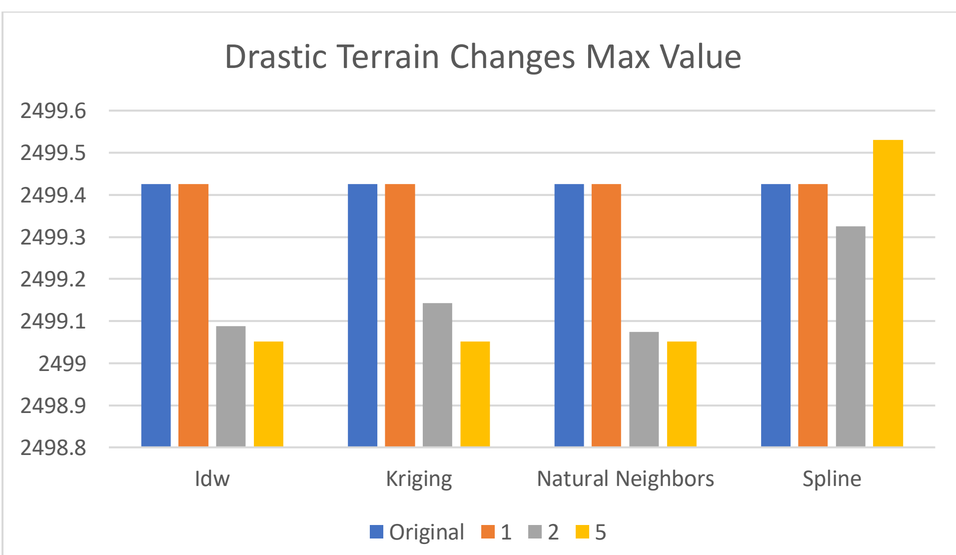

Chart 3 Differences in the interpolators produced maximum values for Drastic Terrain Variation Enviornments

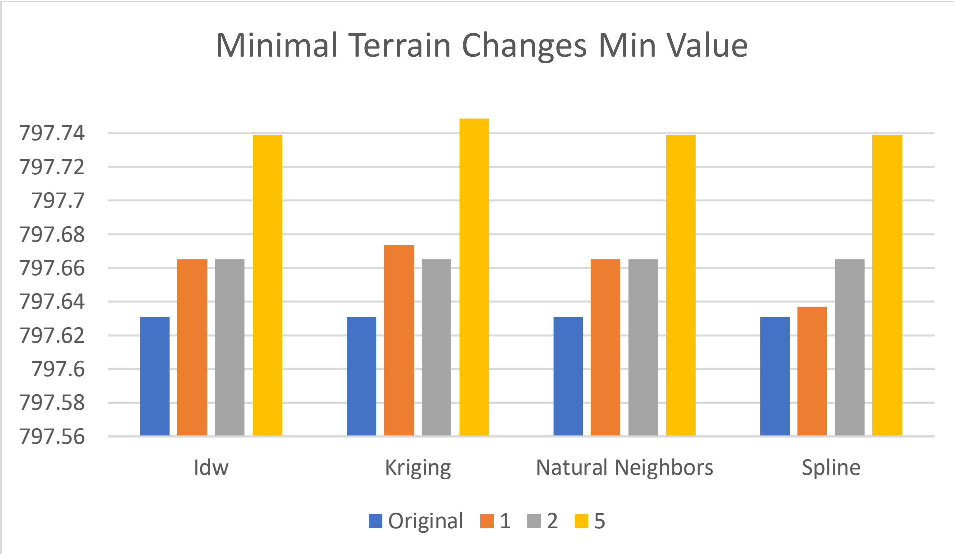

Chart 4 Differences in the interpolators produced minimum values for Minimal Terrain Variation Enviornments

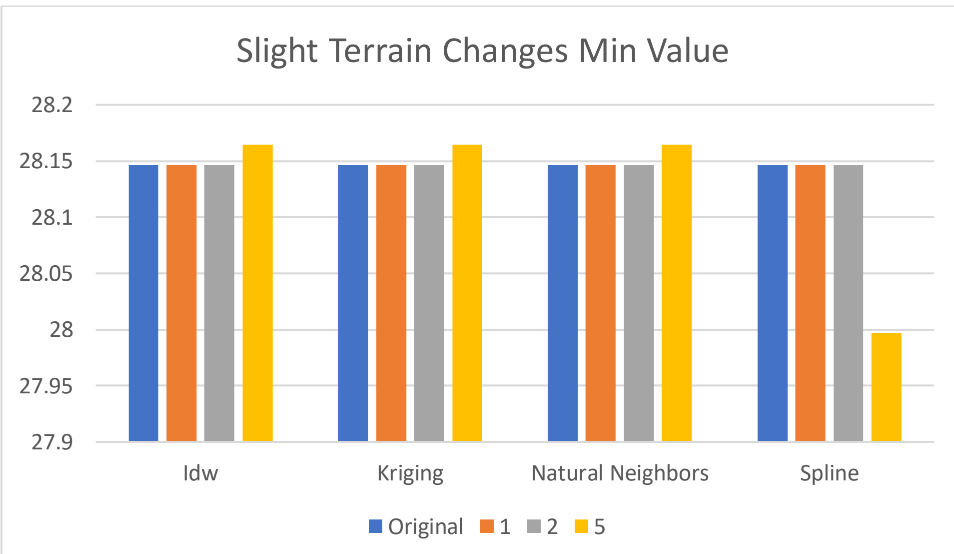

Chart 5 Differences in the interpolators produced minimum values for Slight Terrain Variation Enviornments

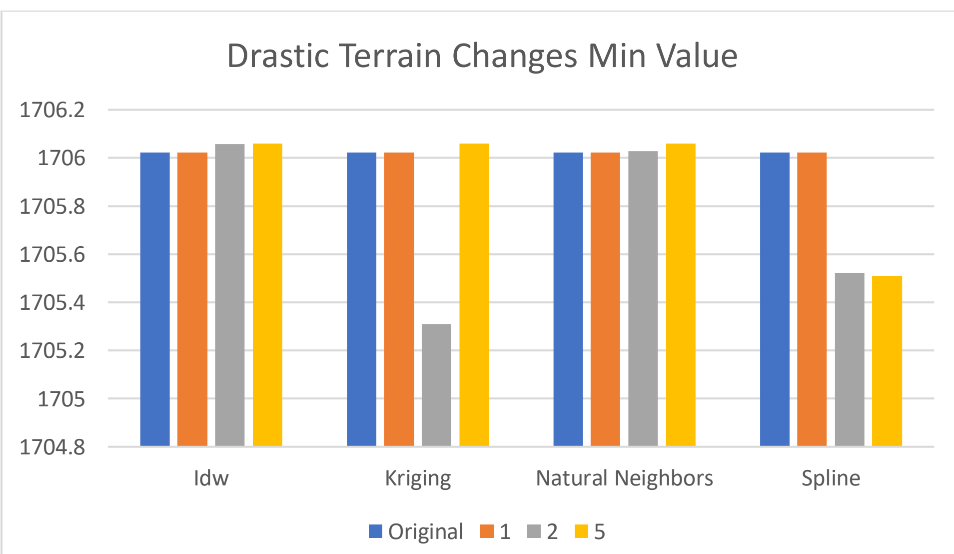

Chart 6 Differences in the interpolators produced minimum values for Drastic Terrain Variation Enviornments

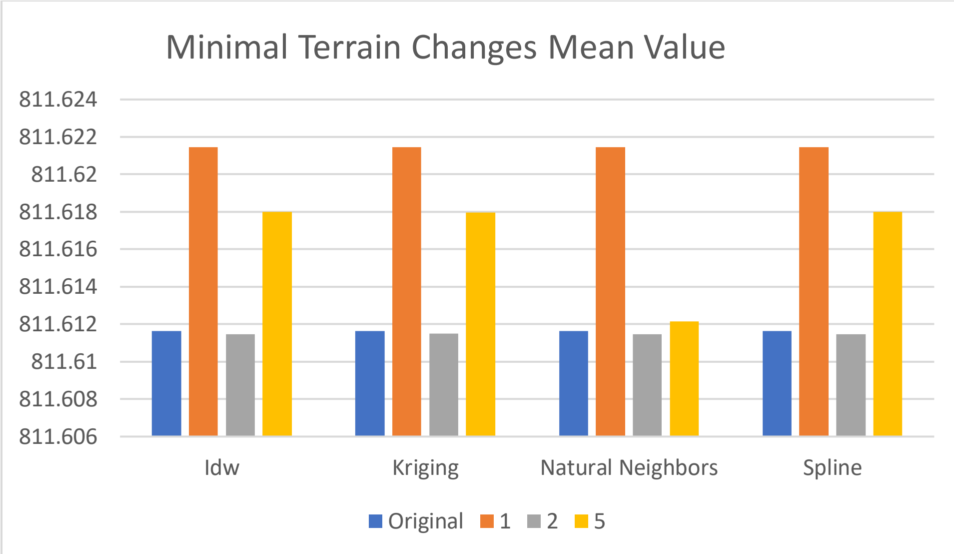

Chart 7 Differences in the interpolators produced mean values for Minimal Terrain Variation Enviornments

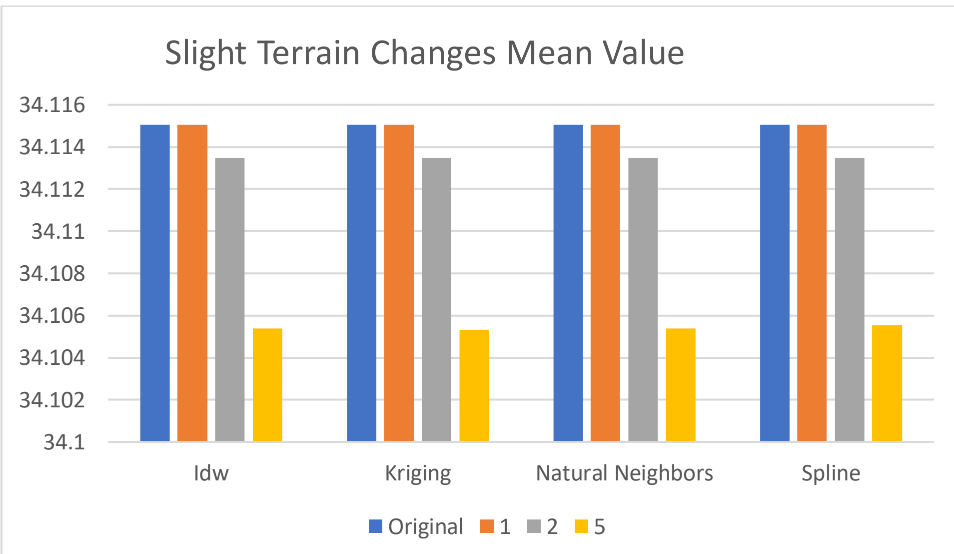

Chart 8 Differences in the interpolators produced mean values for Slight Terrain Variation Enviornments

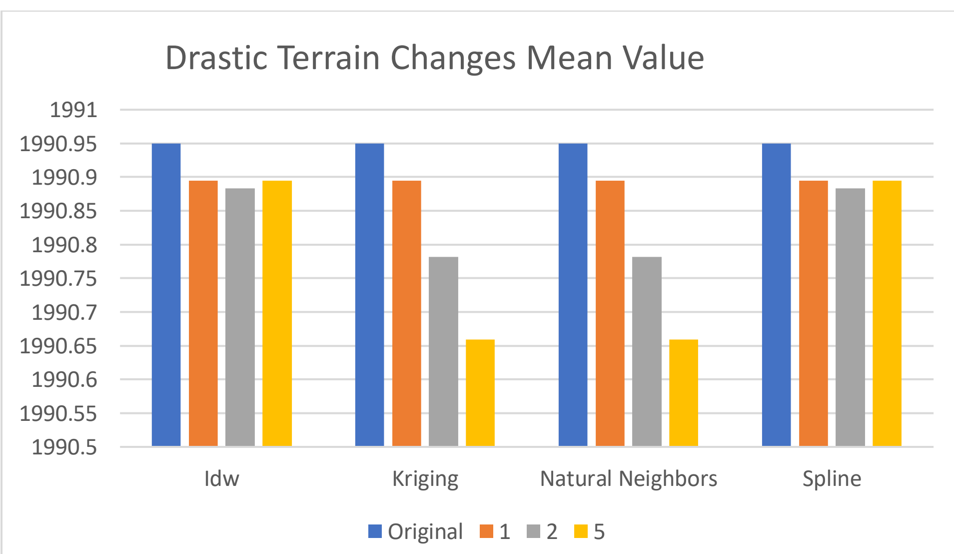

Chart 9 Differences in the interpolators produced mean values for Drastic Terrain Variation Enviornments

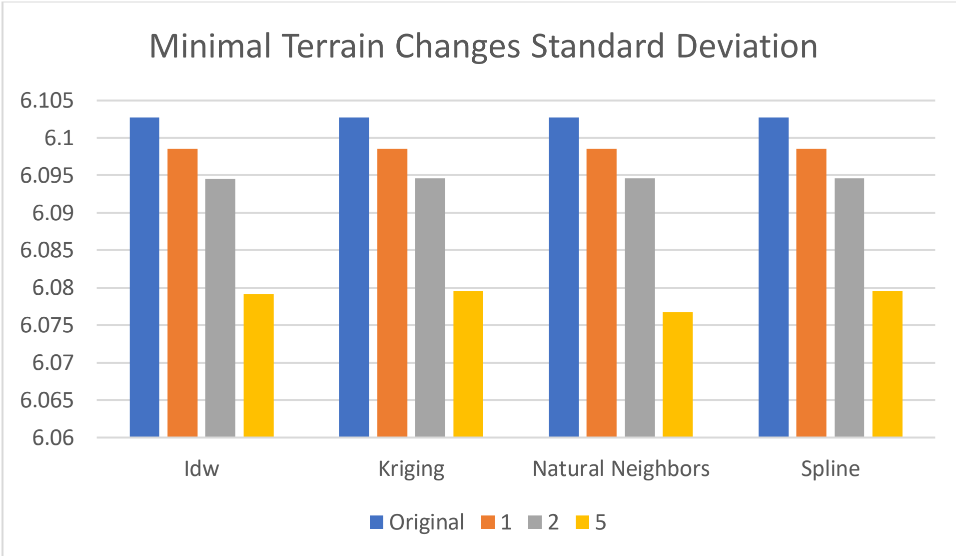

Chart 10 Differences in the interpolators produced Standard Deviation values for Minimal Terrain Variation Enviornments

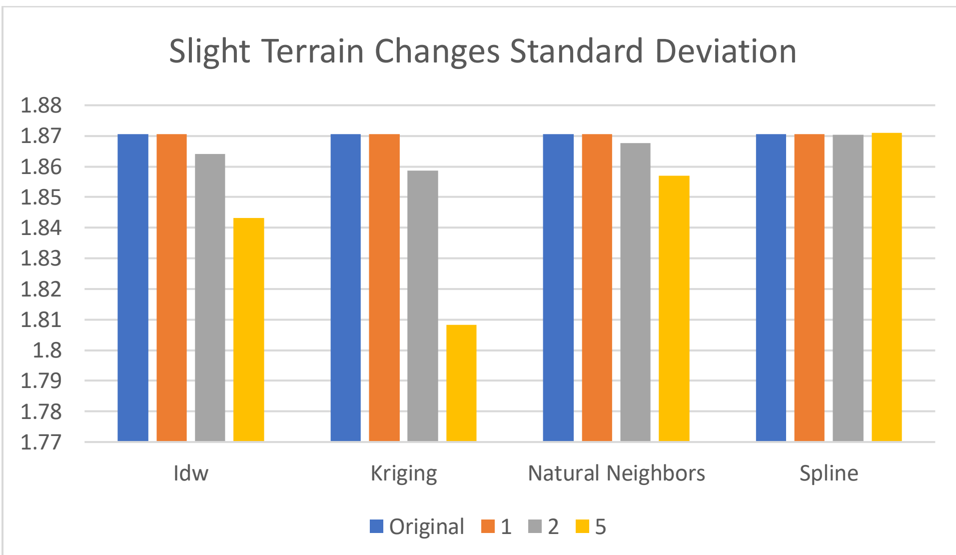

Chart 11 Differences in the interpolators produced Standard Deviation values for Slight Terrain Variation Enviornments

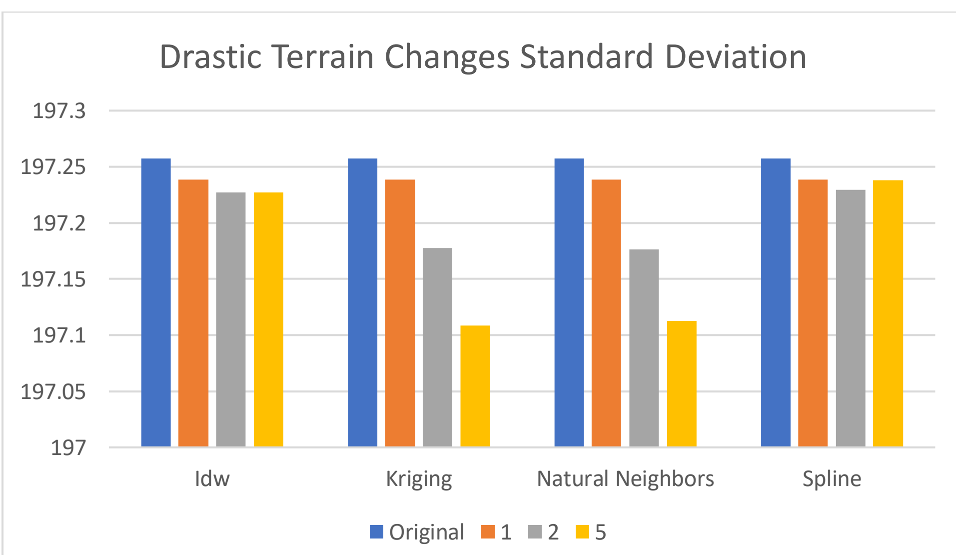

Chart 12 Differences in the interpolators produced Standard Deviation values for Drastic Terrain Variation Enviornments

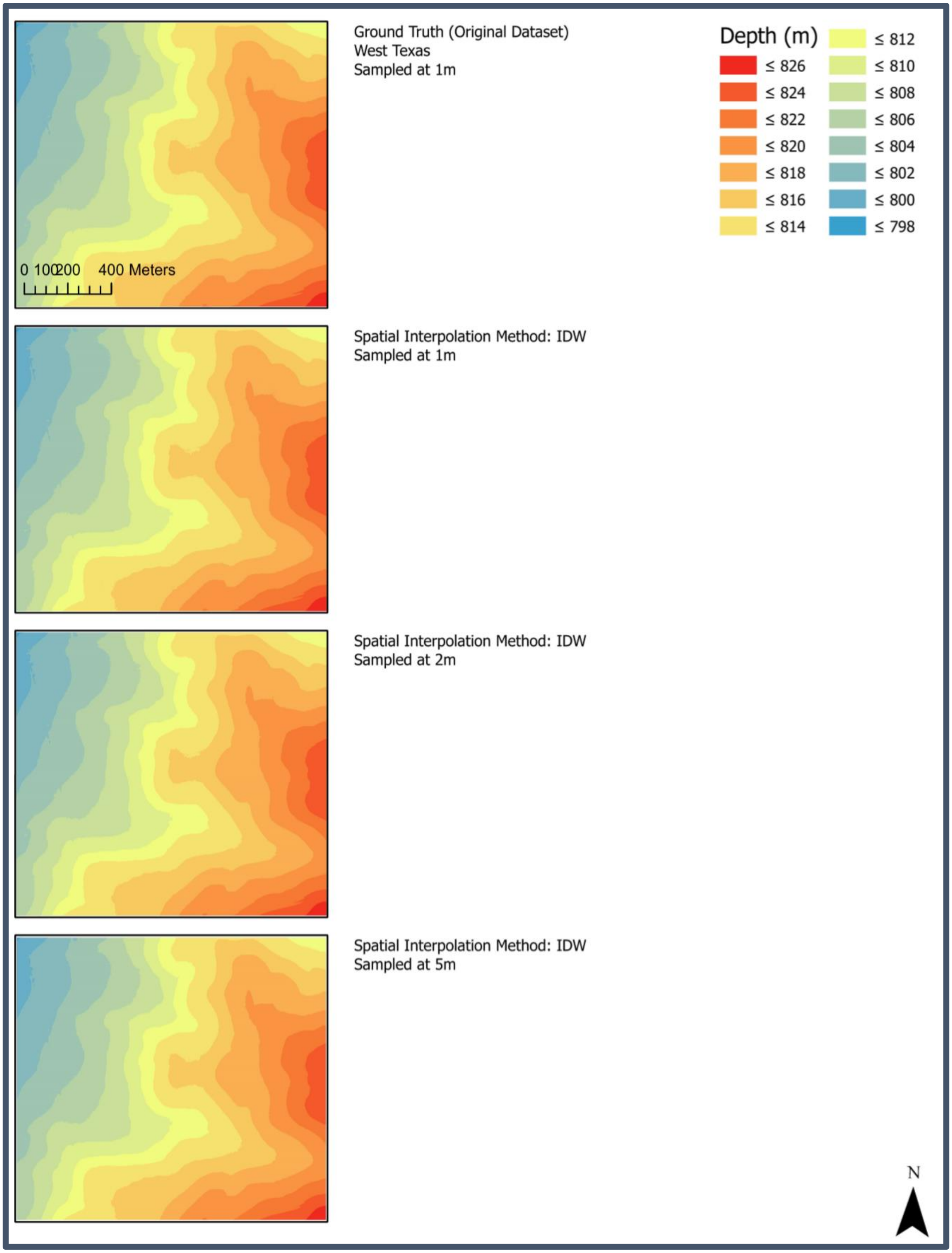

Figure 2 Resulting IDW interpolated surfaces in comparison to their original dataset for Minimal Terrain Variation Enviornments.

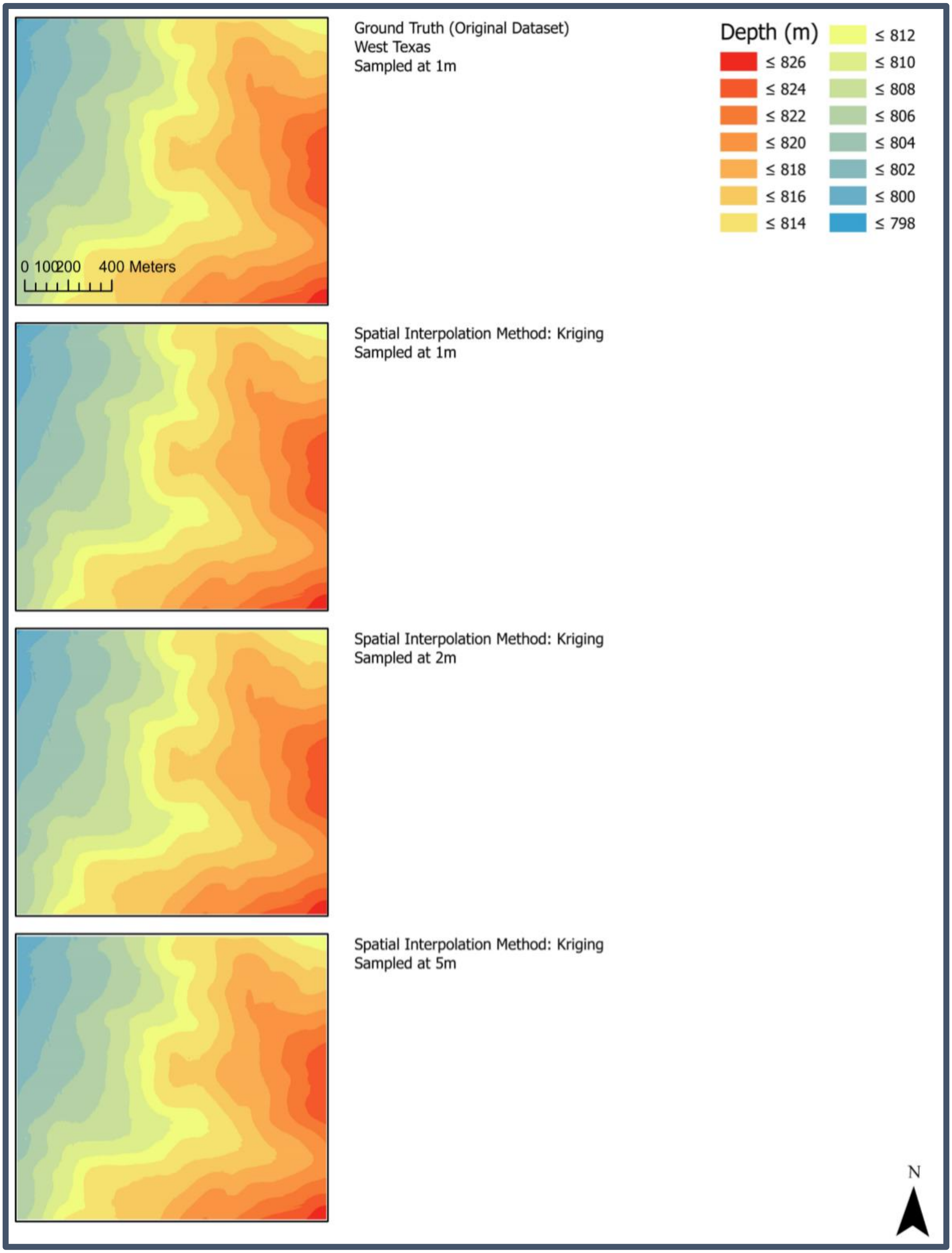

Figure 3 Resulting Kriging interpolated surfaces in comparison to their original dataset for Minimal Terrain Variation Enviornments.

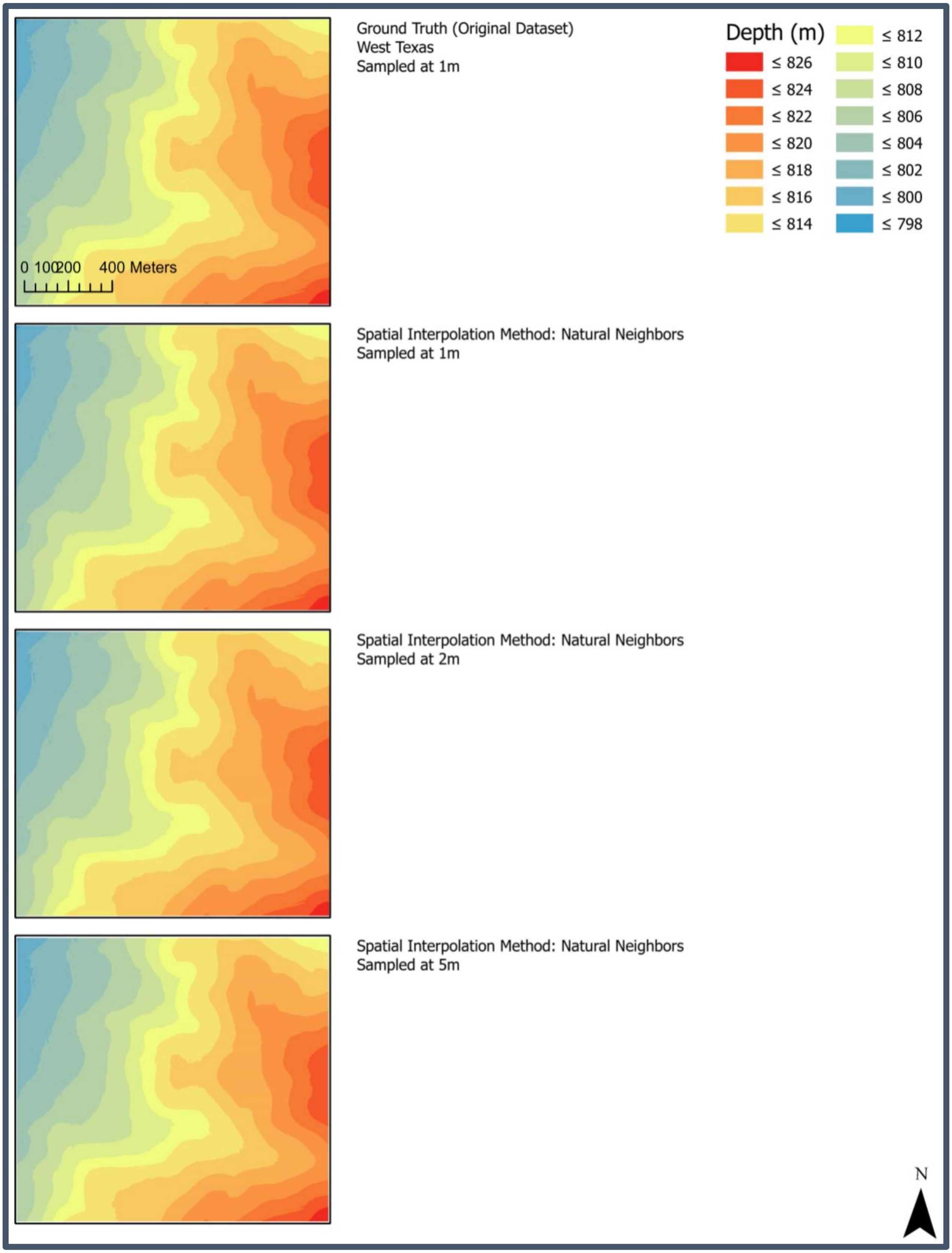

Figure 4 Resulting Natural Neighbors interpolated surfaces in comparison to their original dataset for Minimal Terrain Variation Enviornments.

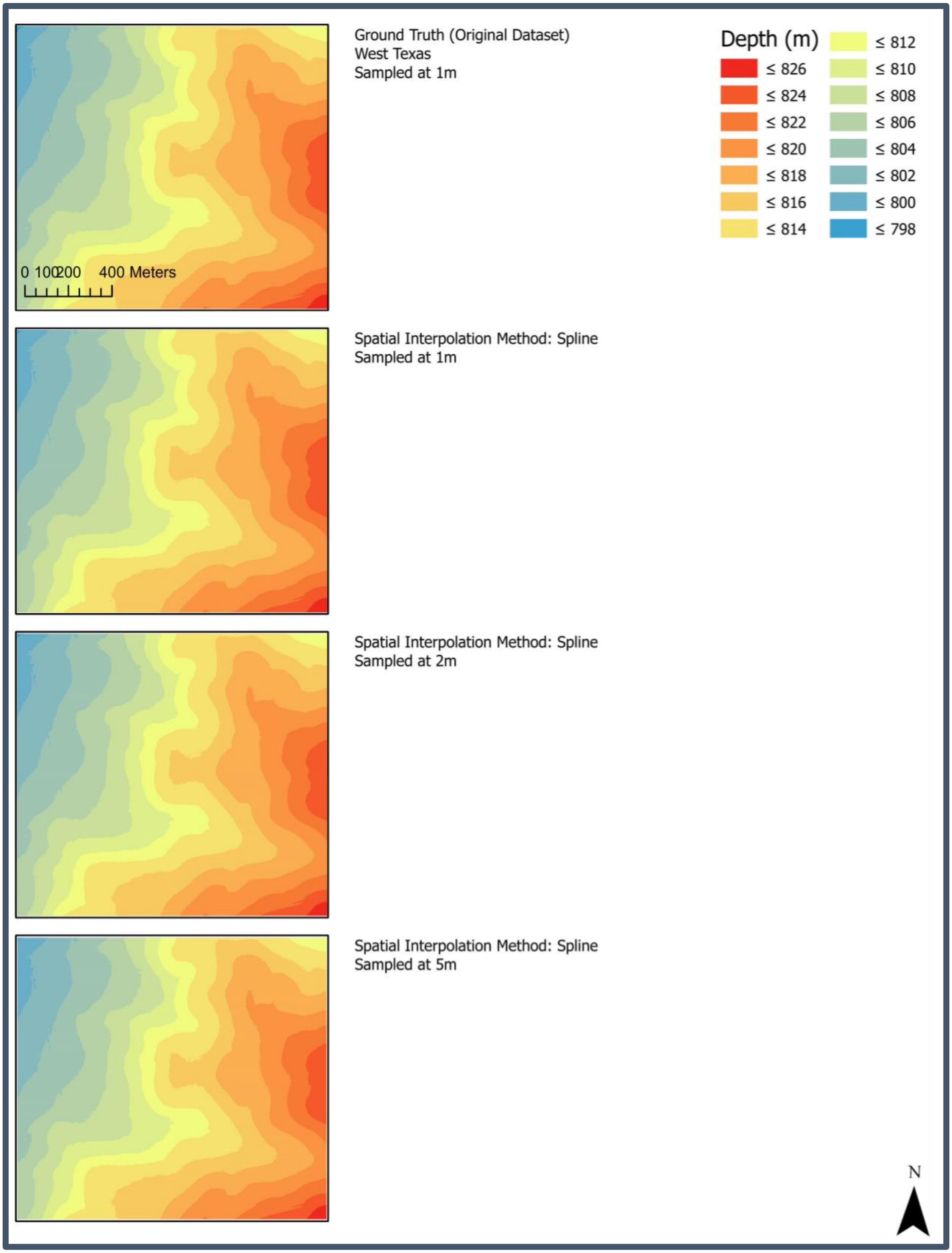

Figure 5 Resulting Spline interpolated surfaces in comparison to their original dataset for Minimal Terrain Variation Enviornments.

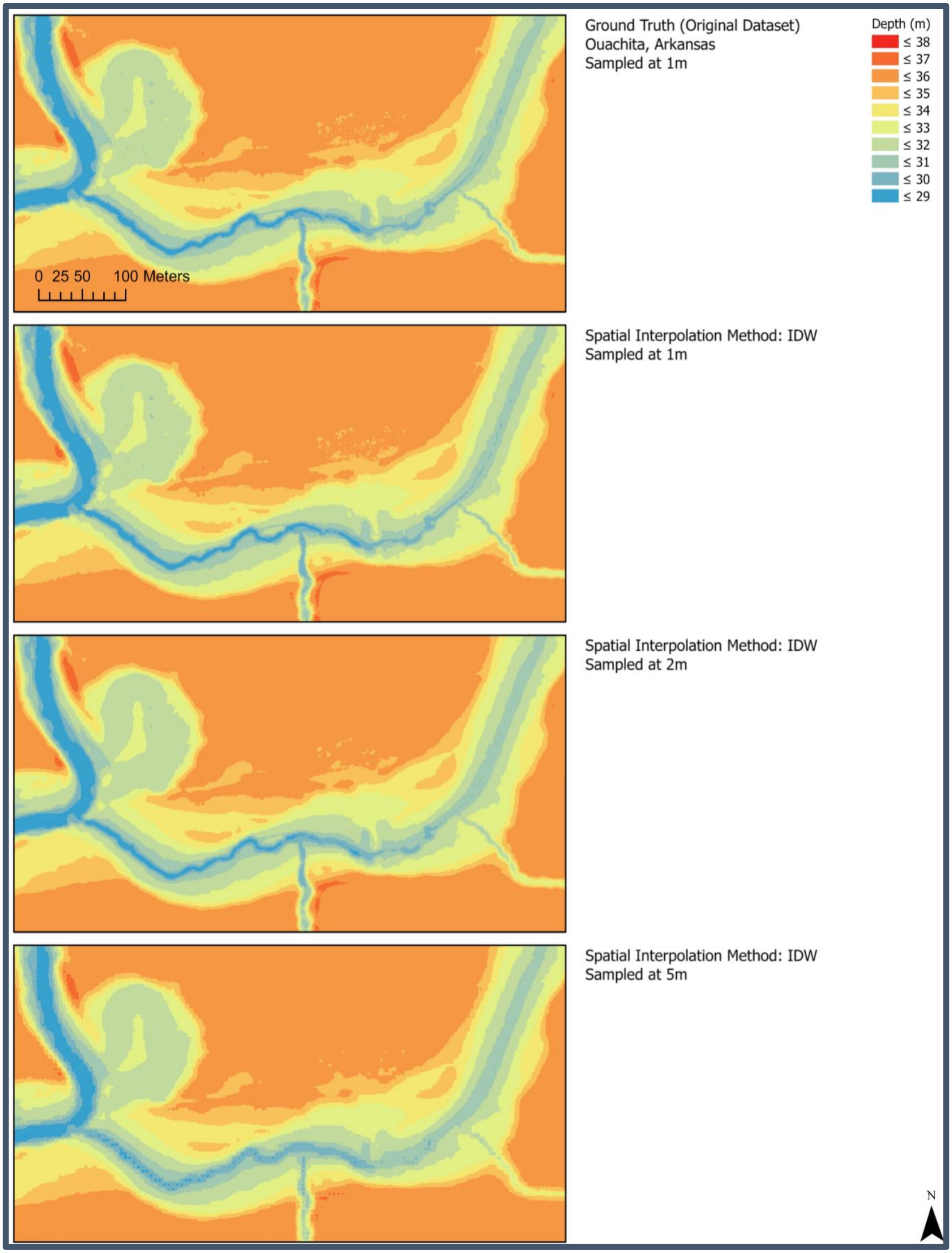

Figure 6 Resulting IDW interpolated surfaces in comparison to their original dataset for Slight Terrain Variation Enviornments.

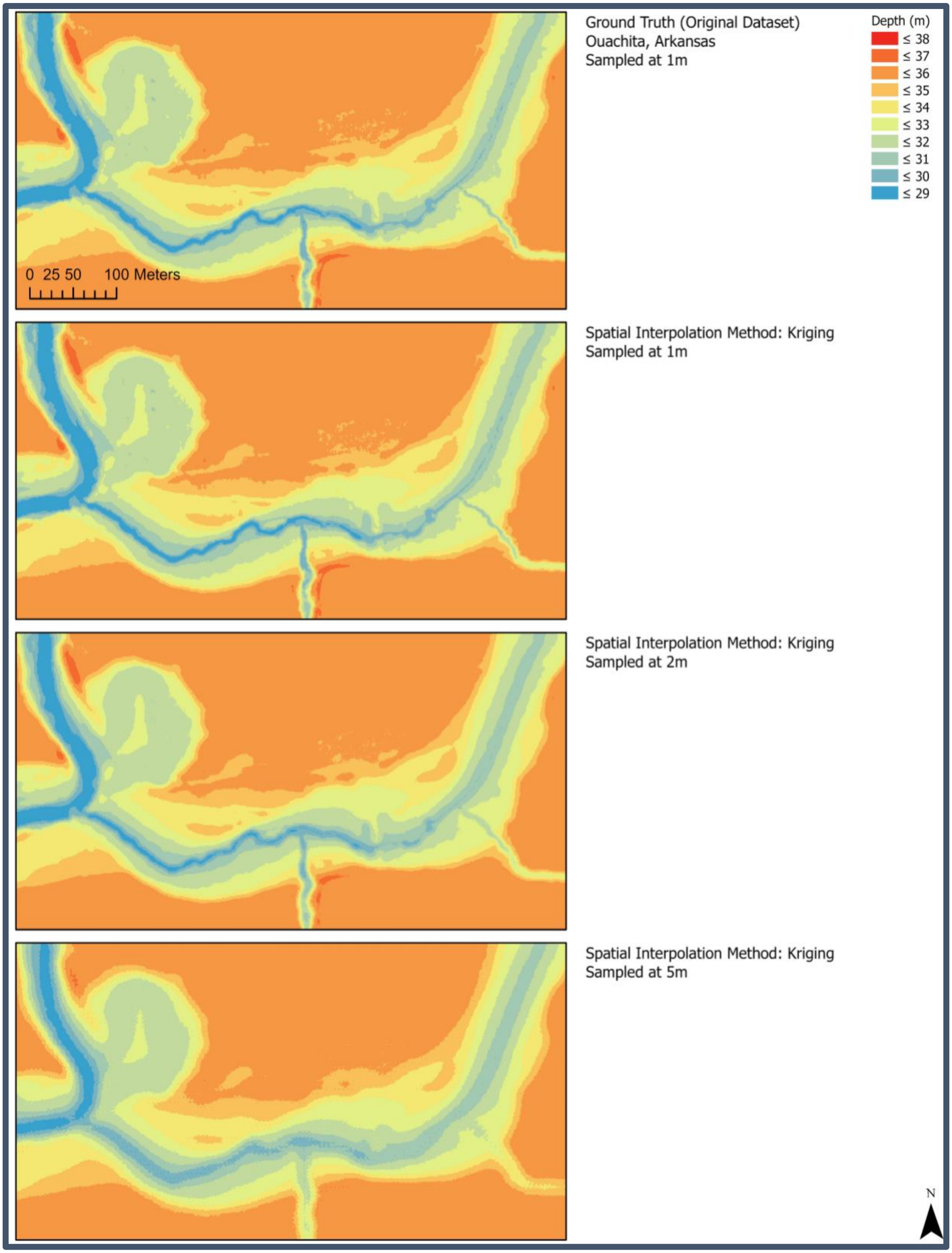

Figure 7 Resulting Kriging interpolated surfaces in comparison to their original dataset for Slight Terrain Variation Enviornments.

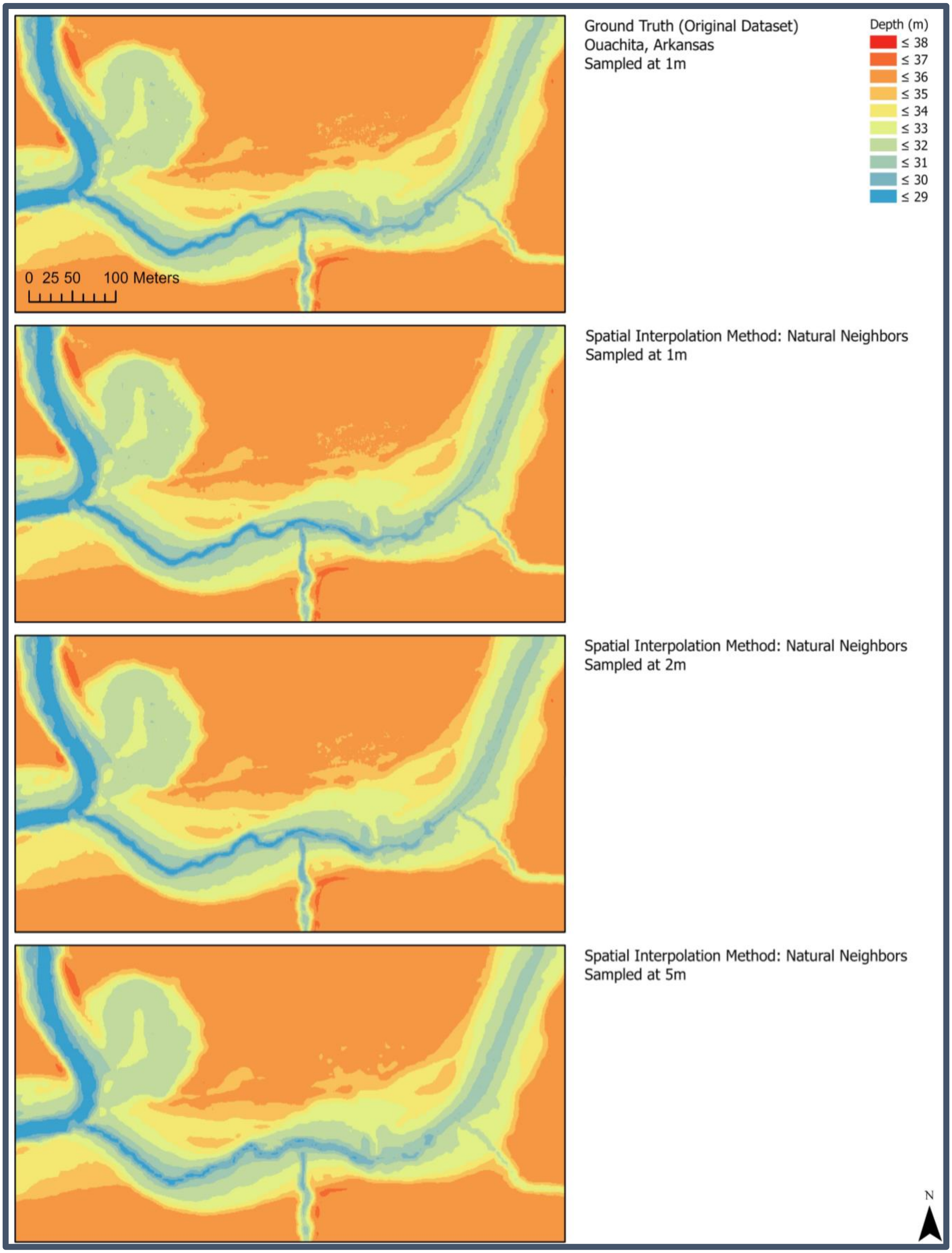

Figure 8 Resulting Natural Neighbors interpolated surfaces in comparison to their original dataset for Slight Terrain Variation Enviornments.

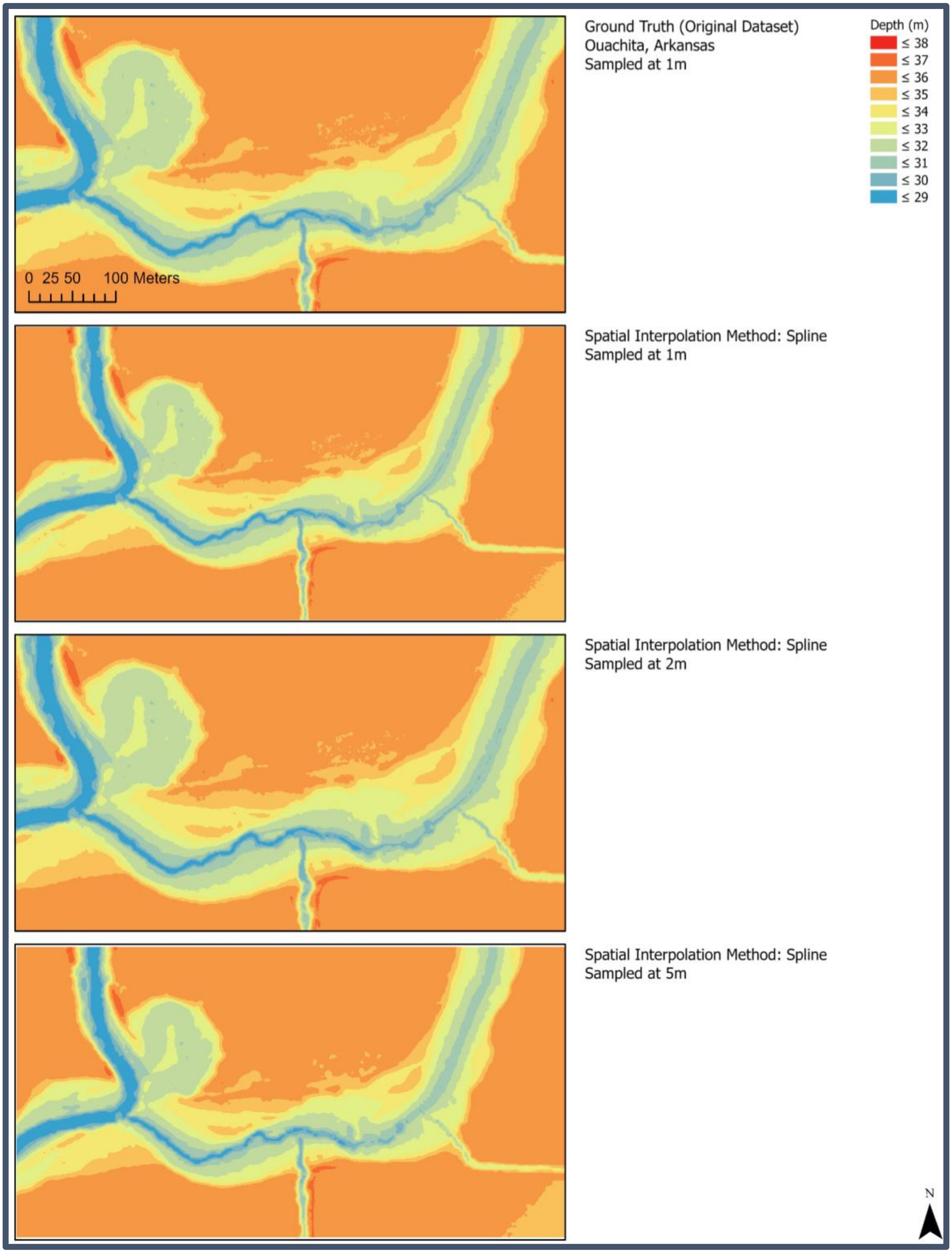

Figure 9 Resulting Spline interpolated surfaces in comparison to their original dataset for Slight Terrain Variation Enviornments.

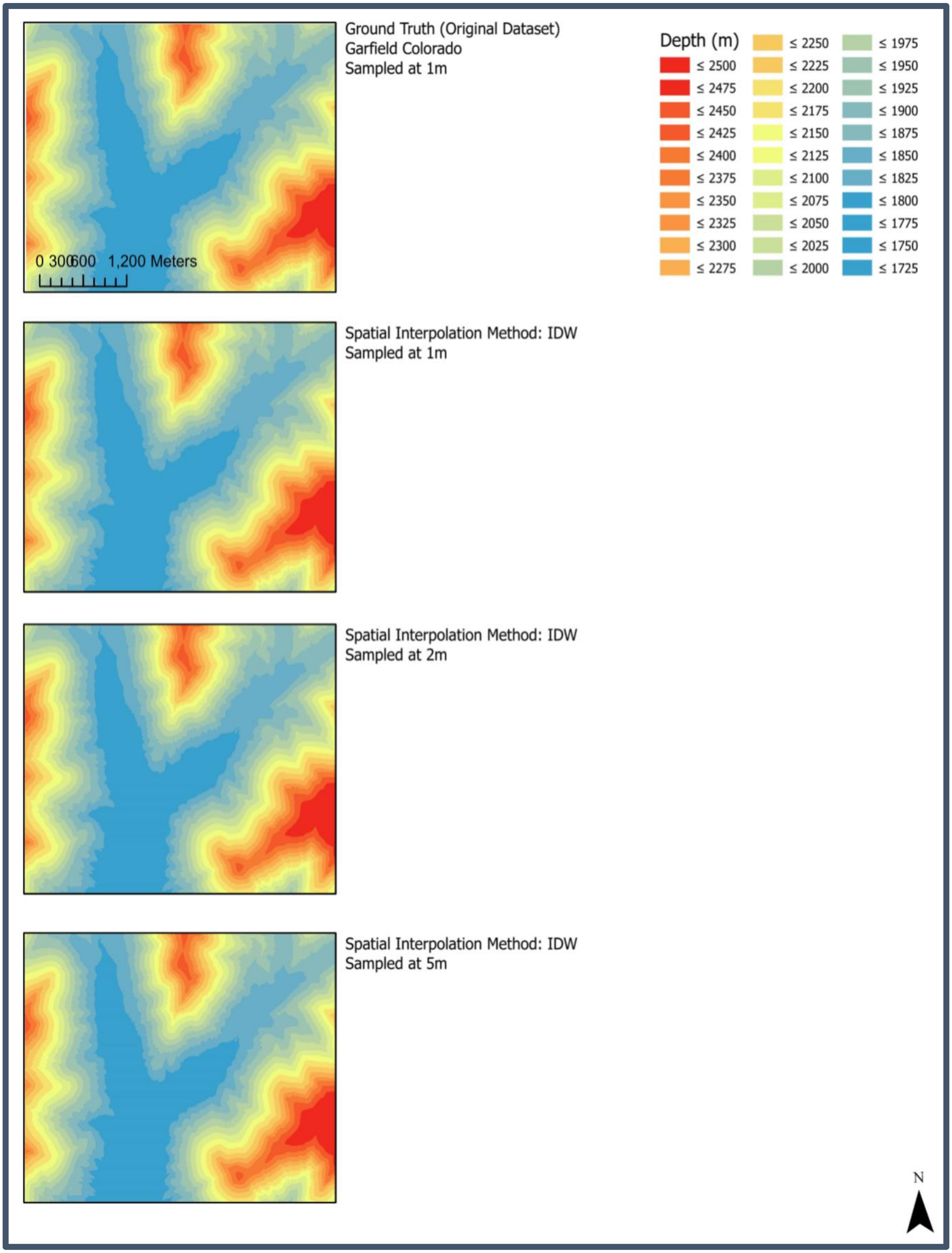

Figure 10 Resulting IDW interpolated surfaces in comparison to their original dataset for Drastic Terrain Variation Enviornments.

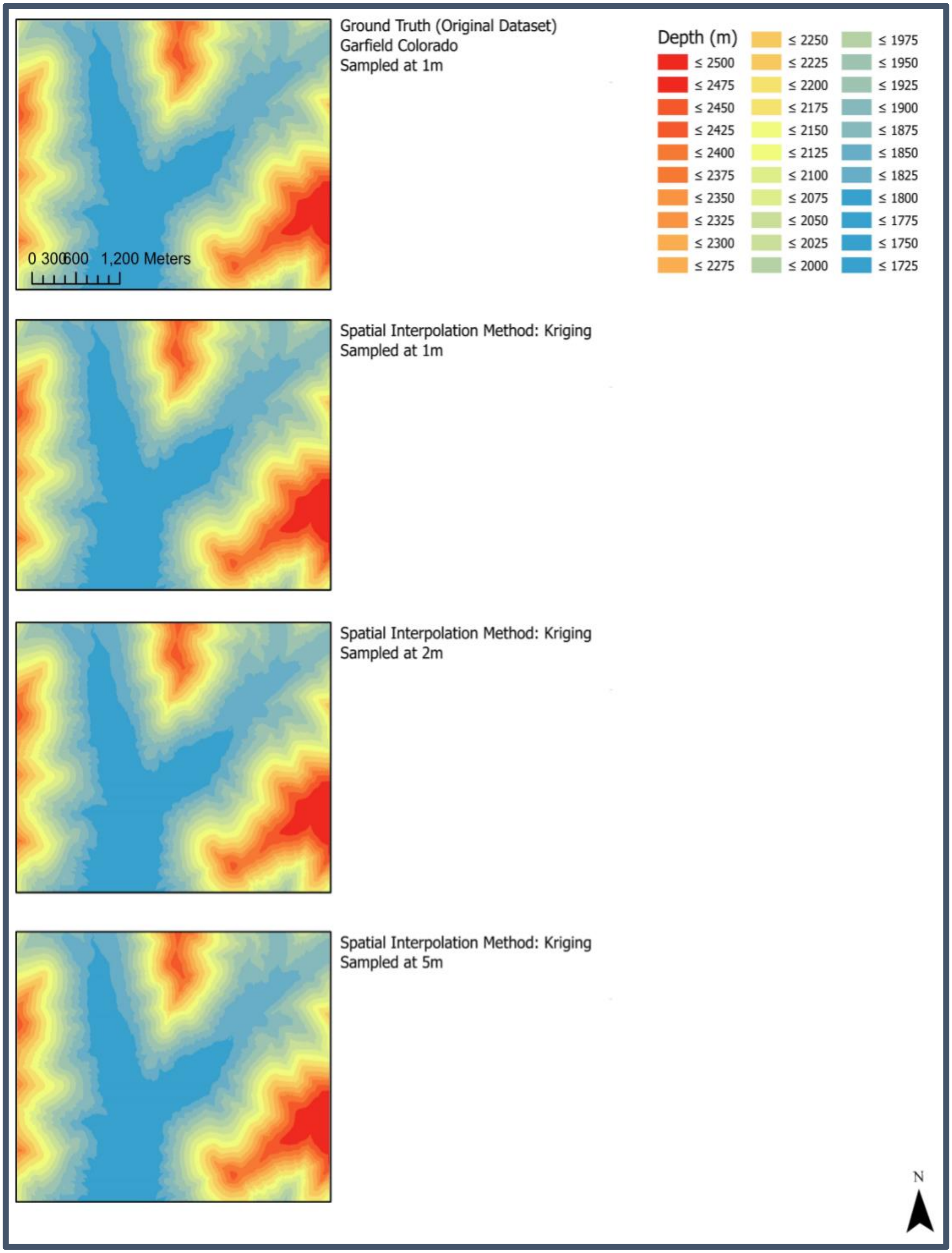

Figure 11 Resulting Kriging interpolated surfaces in comparison to their original dataset for Drastic Terrain Variation Enviornments.

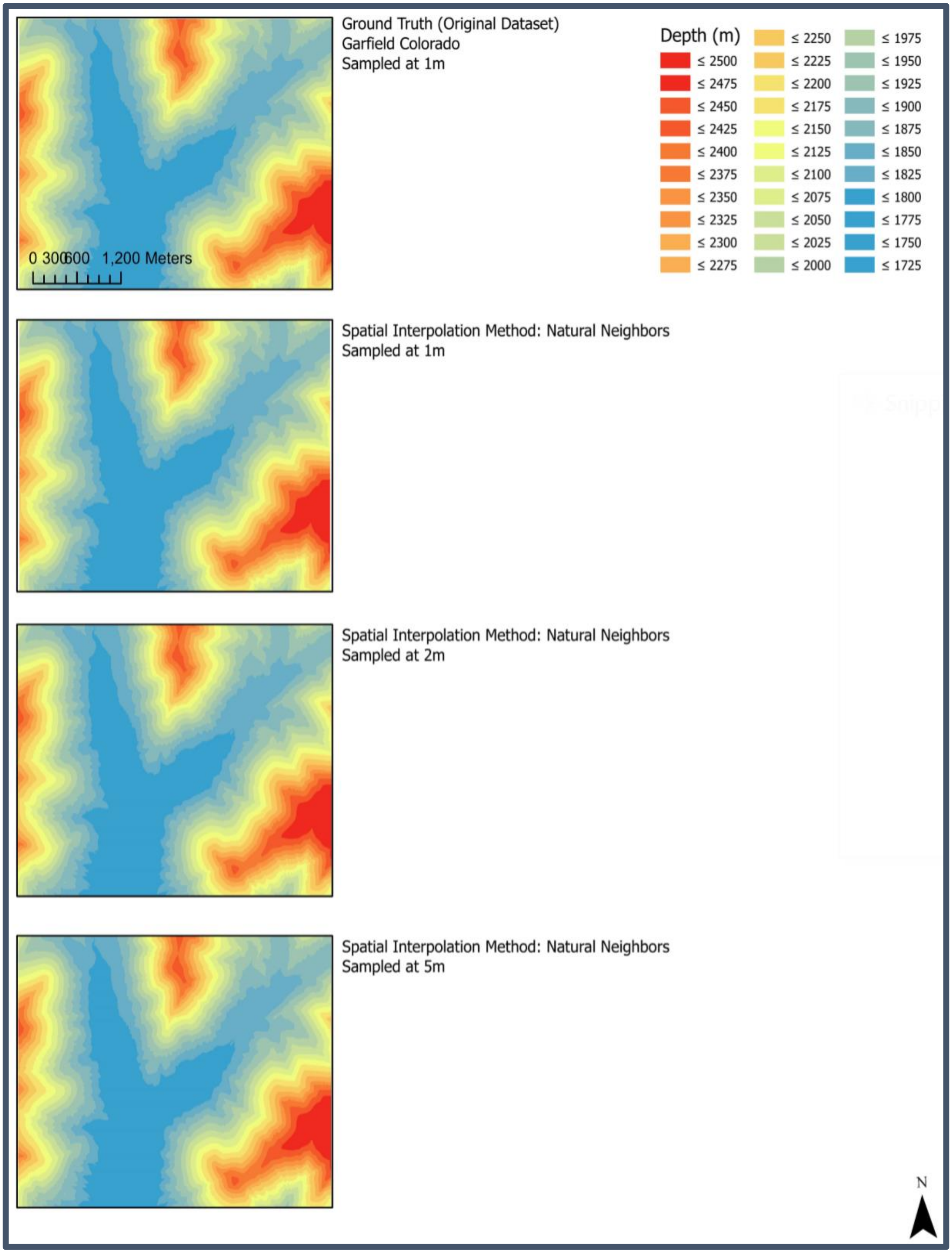

Figure 12 Resulting Natural Neighbors interpolated surfaces in comparison to their original dataset for Drastic Terrain Variation Enviornments.

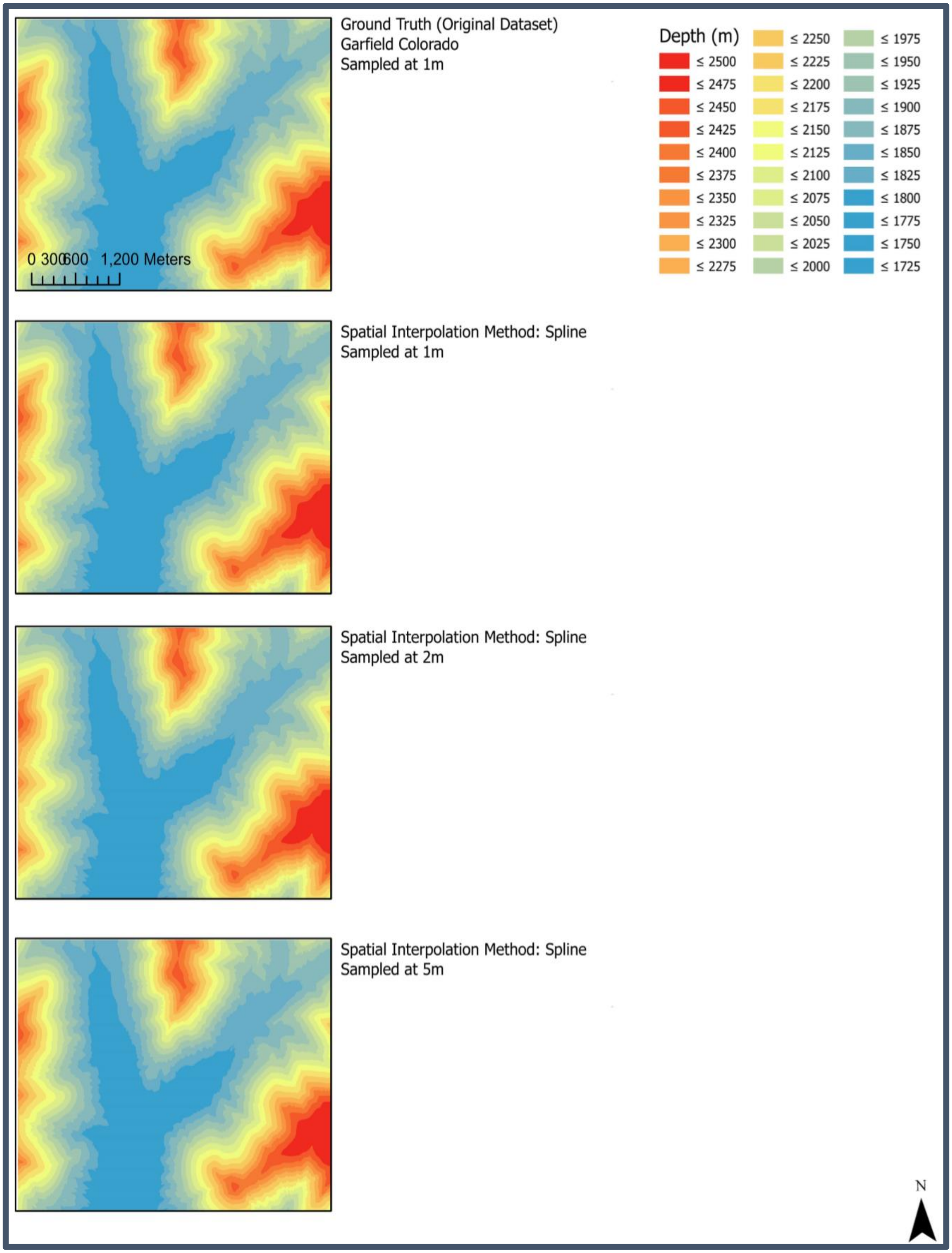

Figure 13 Resulting Spline interpolated surfaces in comparison to their original dataset for Drastic Terrain Variation Enviornments.

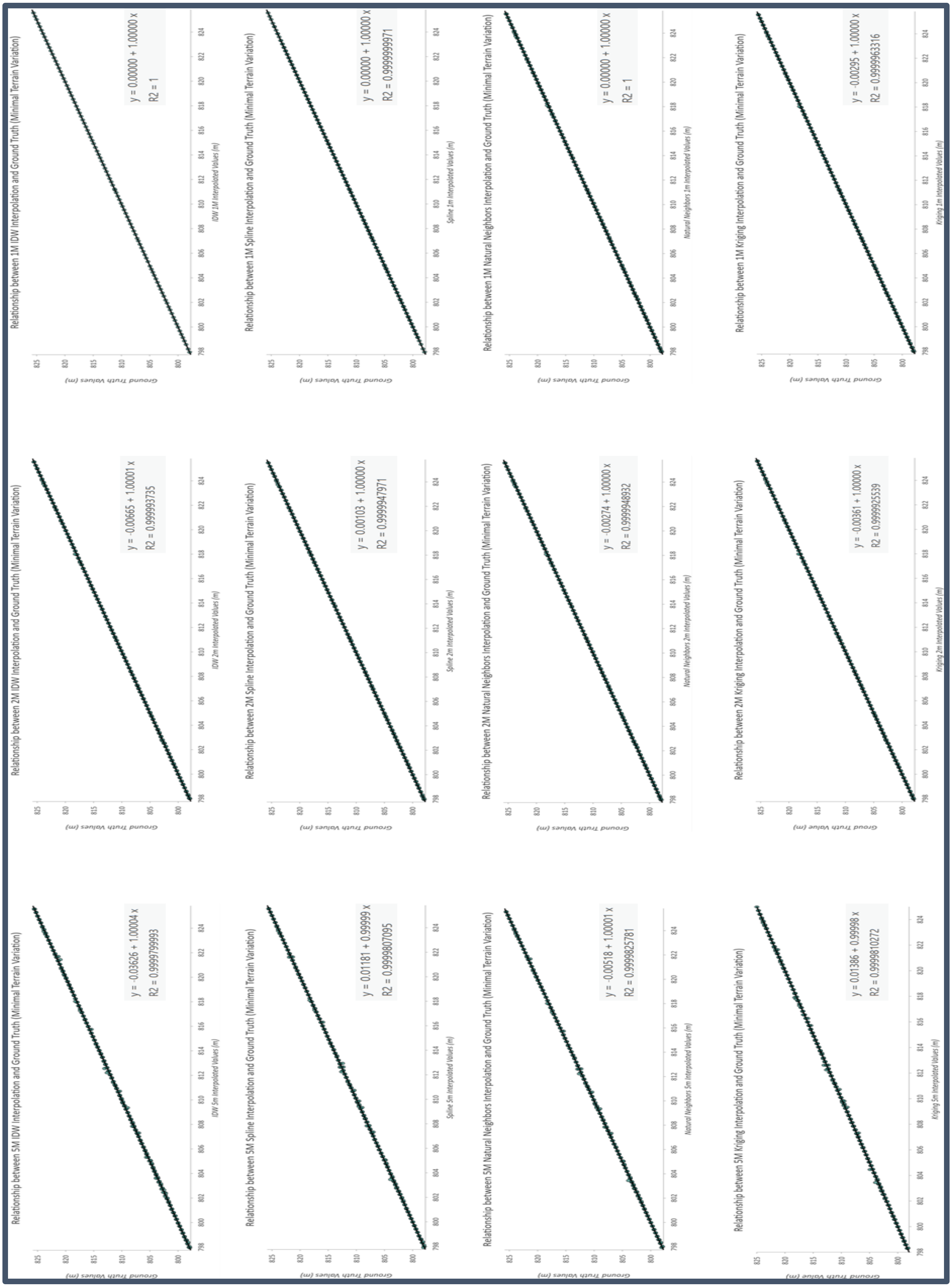

Figure 14 Comparative scatterplots and associated correlation coefficients Interpolated VS Original for Minimal Terrain Variation Environments

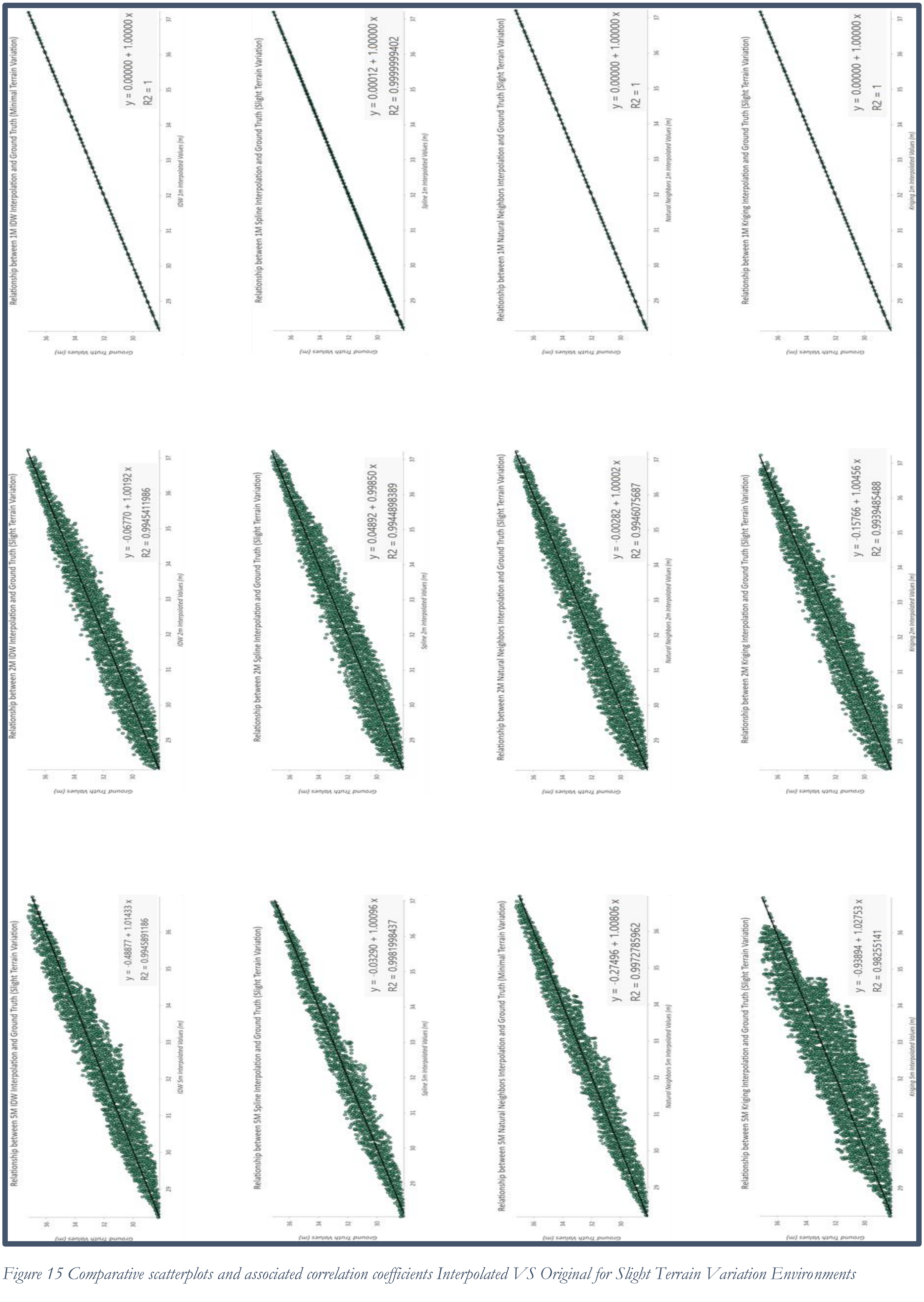

Figure 15 Comparative scatterplots and associated correlation coefficients Interpolated VS Original for Slight Terrain Variation Environments

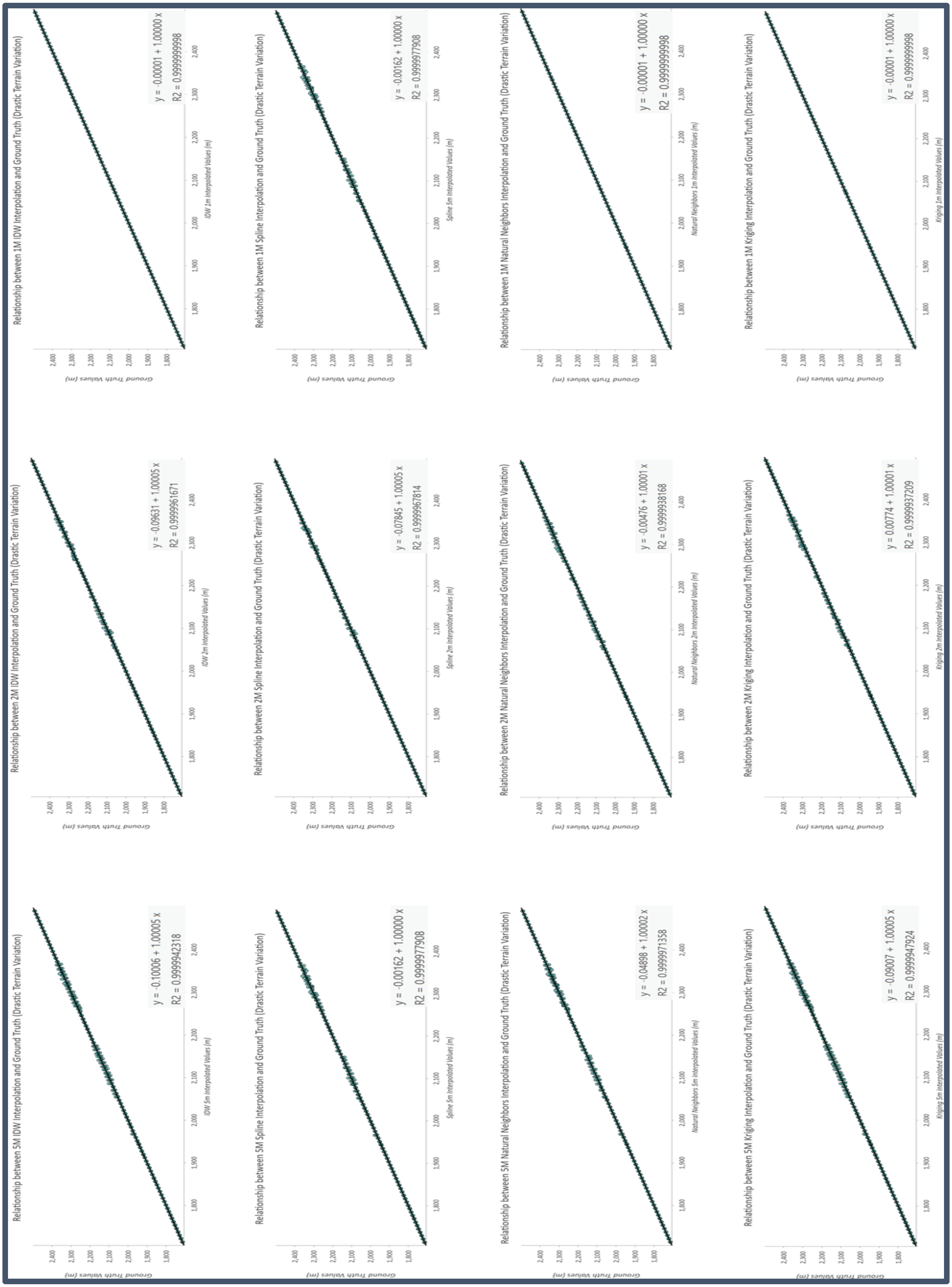

Figure 16 Comparative scatterplots and associated correlation coefficients Interpolated VS Original for Drastic Terrain Variation Environments

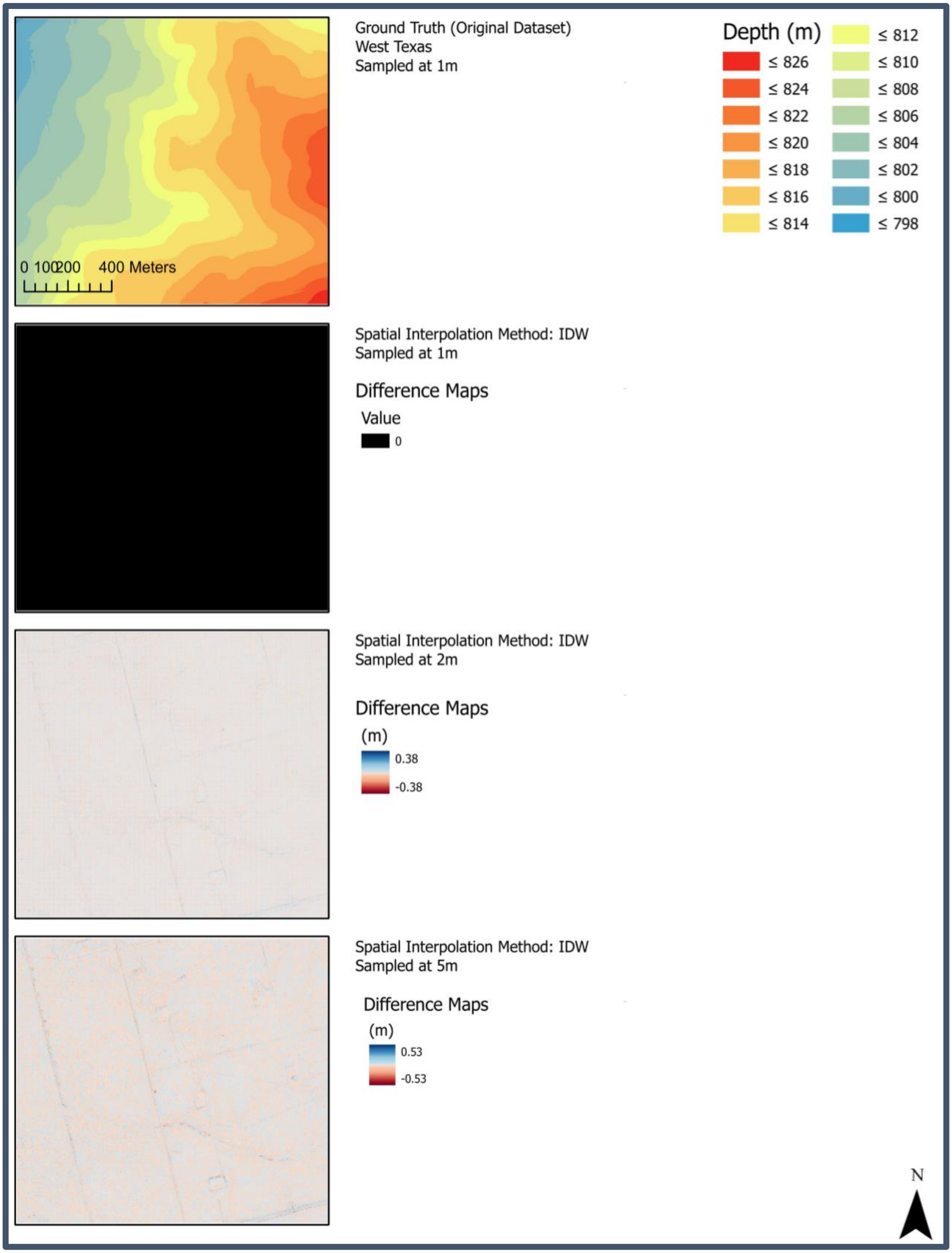

Figure 17 Difference Map Set for IDW Interpolation in Minimal Terrain Variation Environments

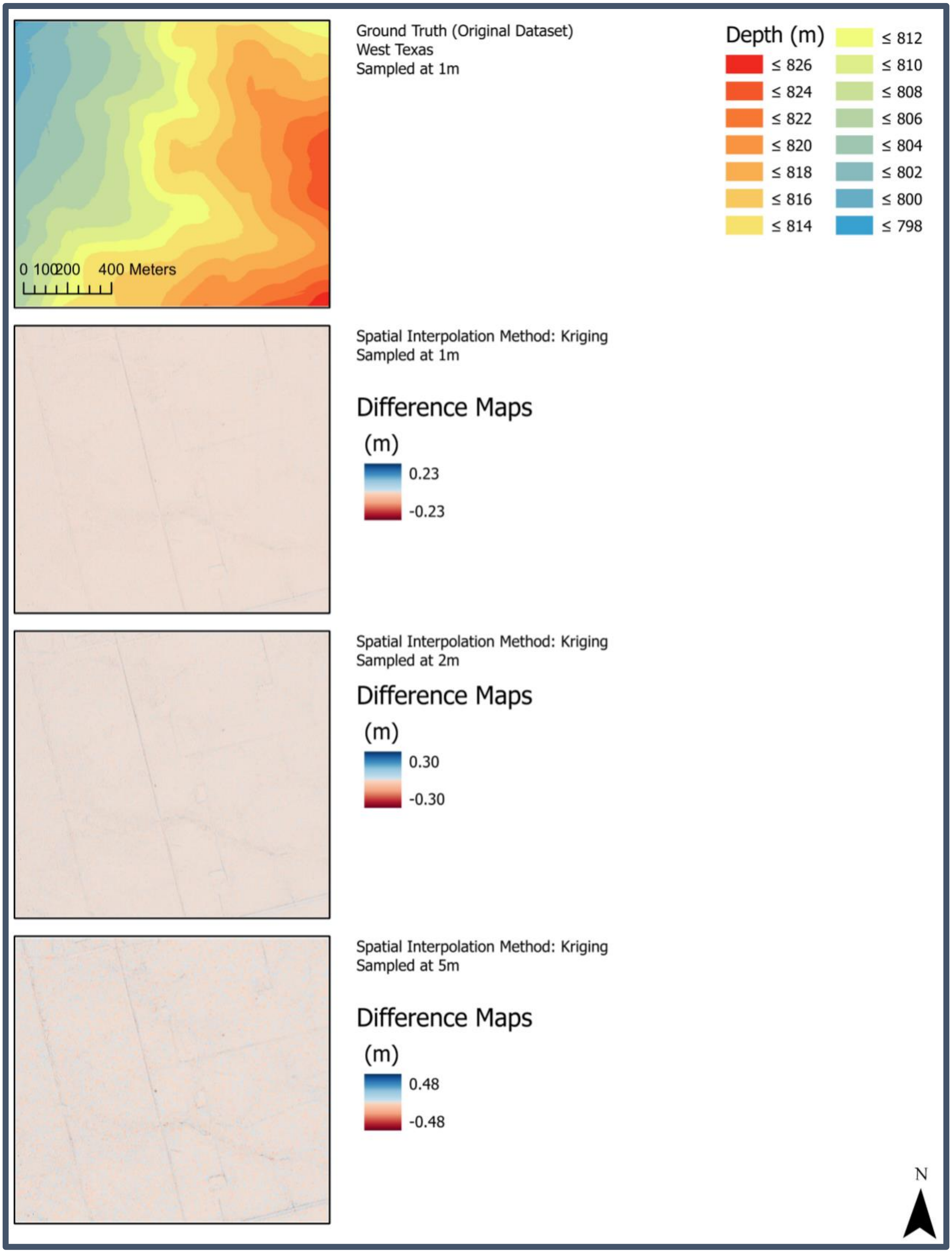

Figure 18 Difference Map Set for Kriging Interpolation in Minimal Terrain Variation Environments

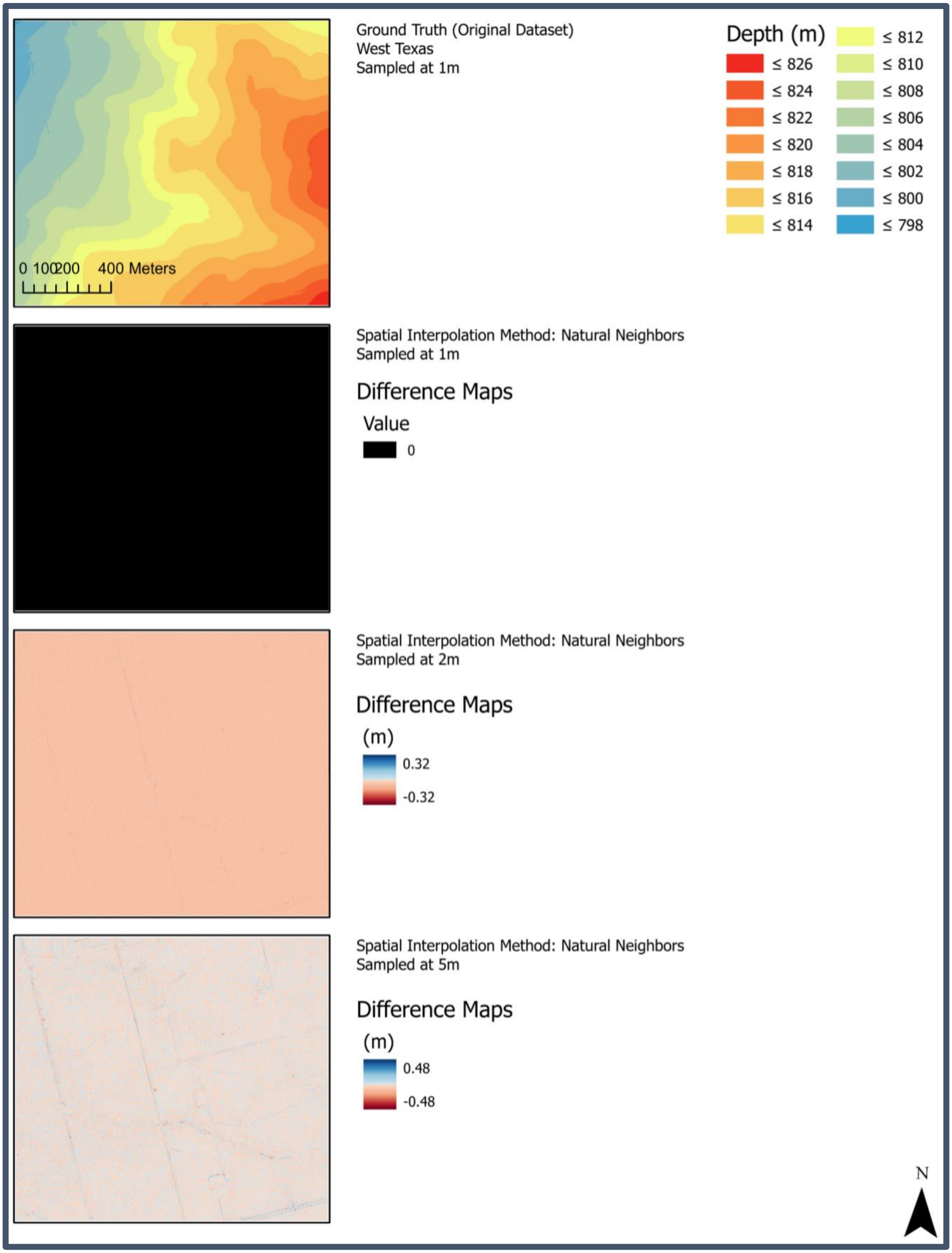

Figure 19 Difference Map Set for Natural Neighbors Interpolation in Minimal Terrain Variation Environments

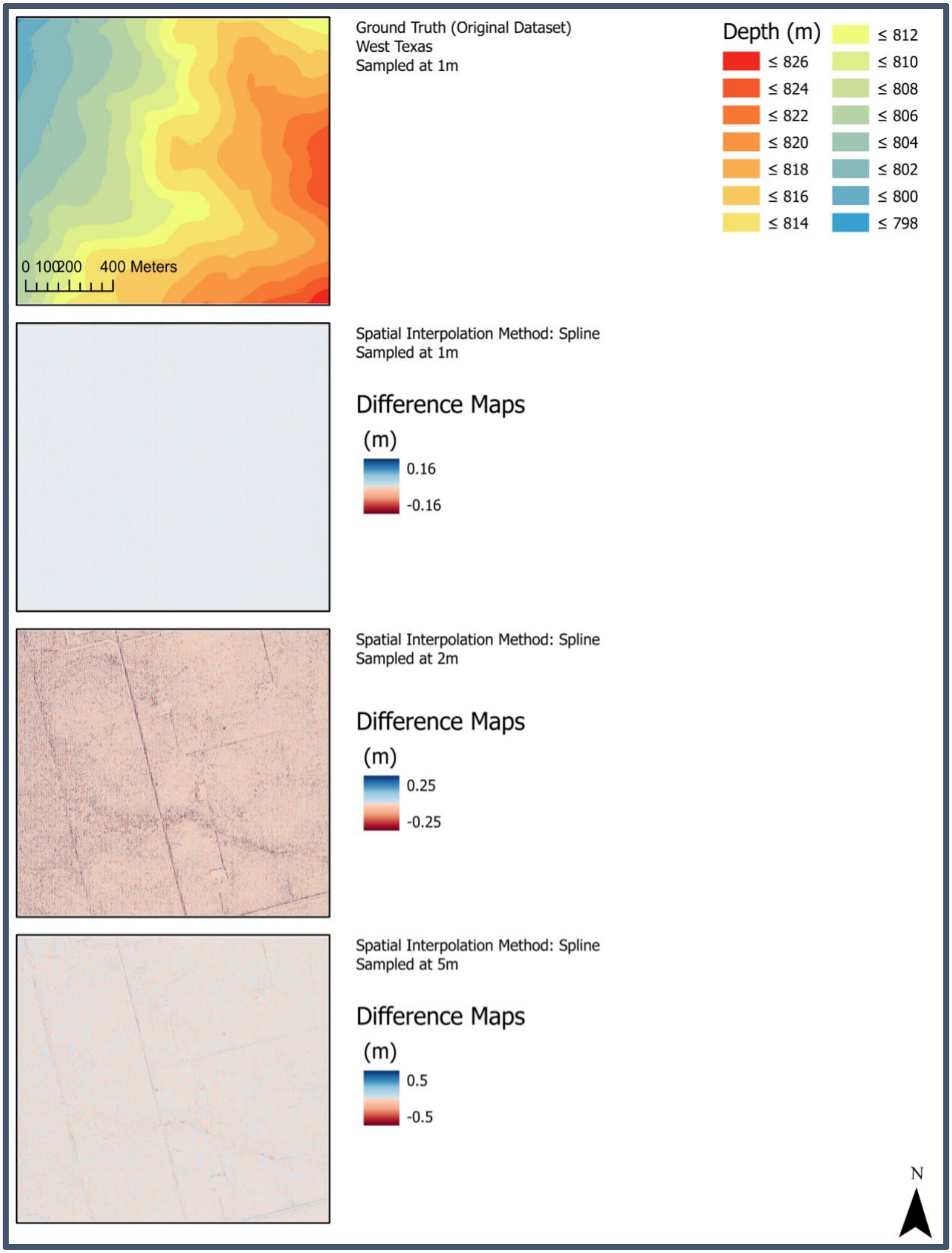

Figure 20 Difference Map Set for Spline Interpolation in Minimal Terrain Variation Environments

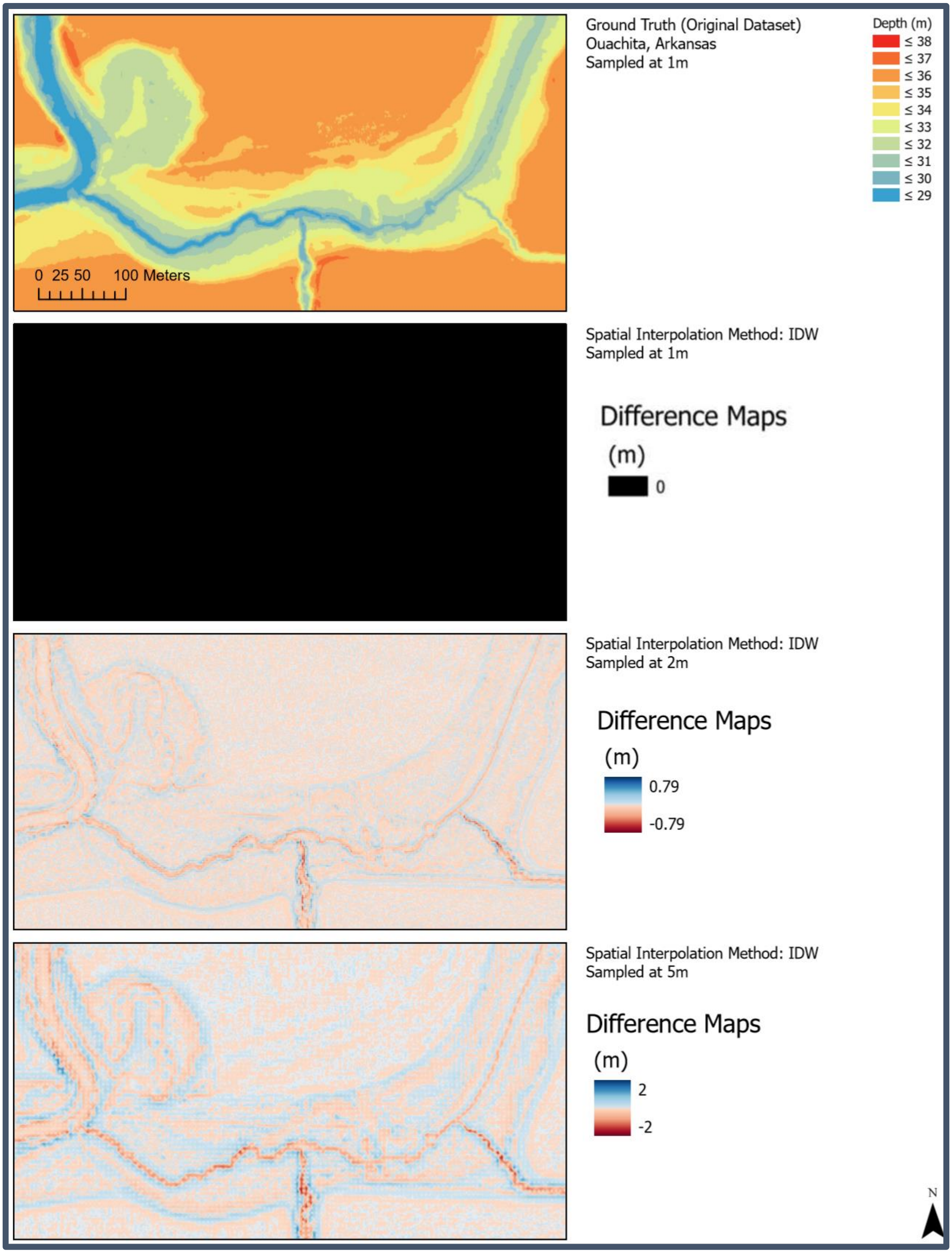

Figure 21 Difference Map Set for IDW Interpolation in Slight Terrain Variation Environments

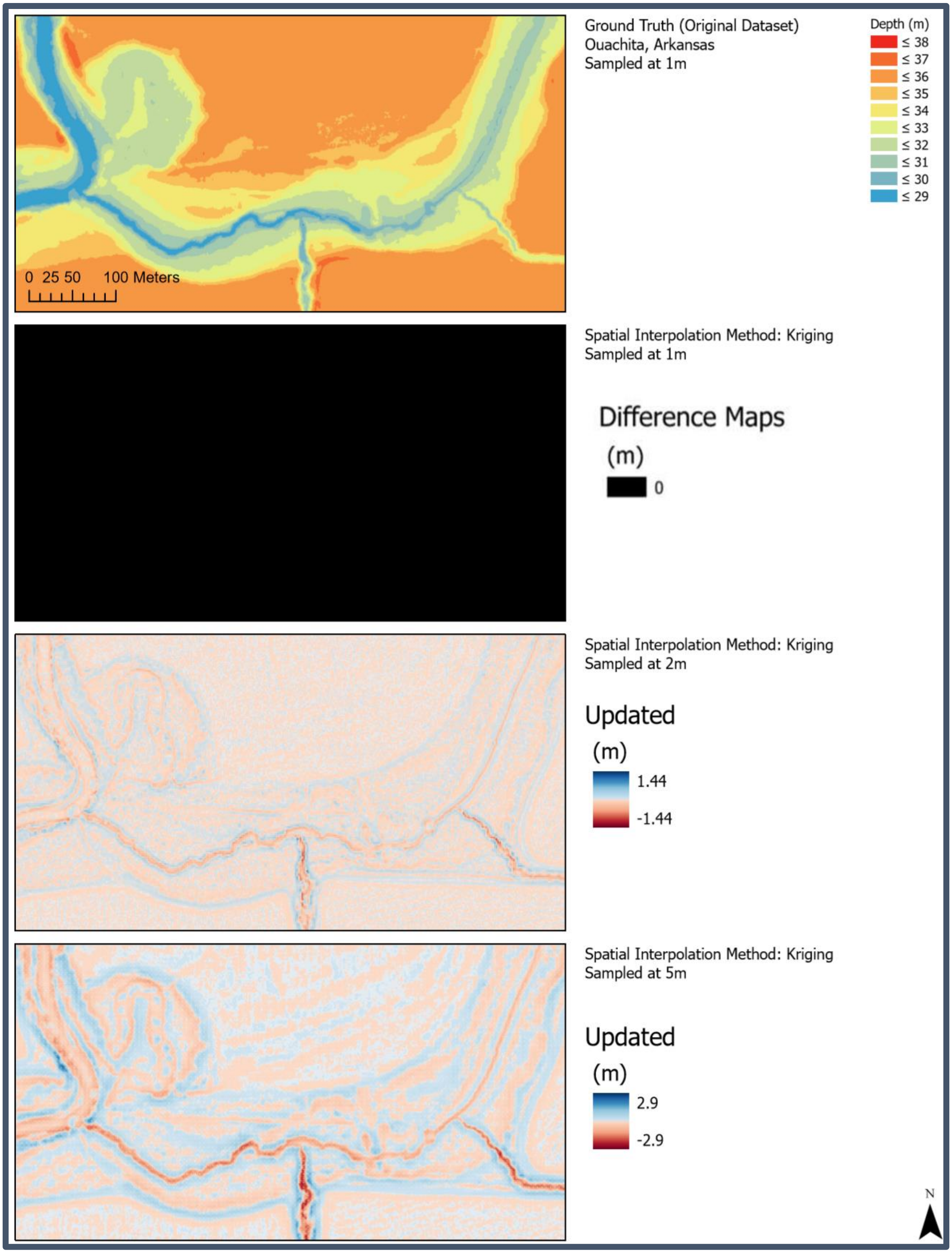

Figure 22 Difference Map Set for Kriging Interpolation in Slight Terrain Variation Environments

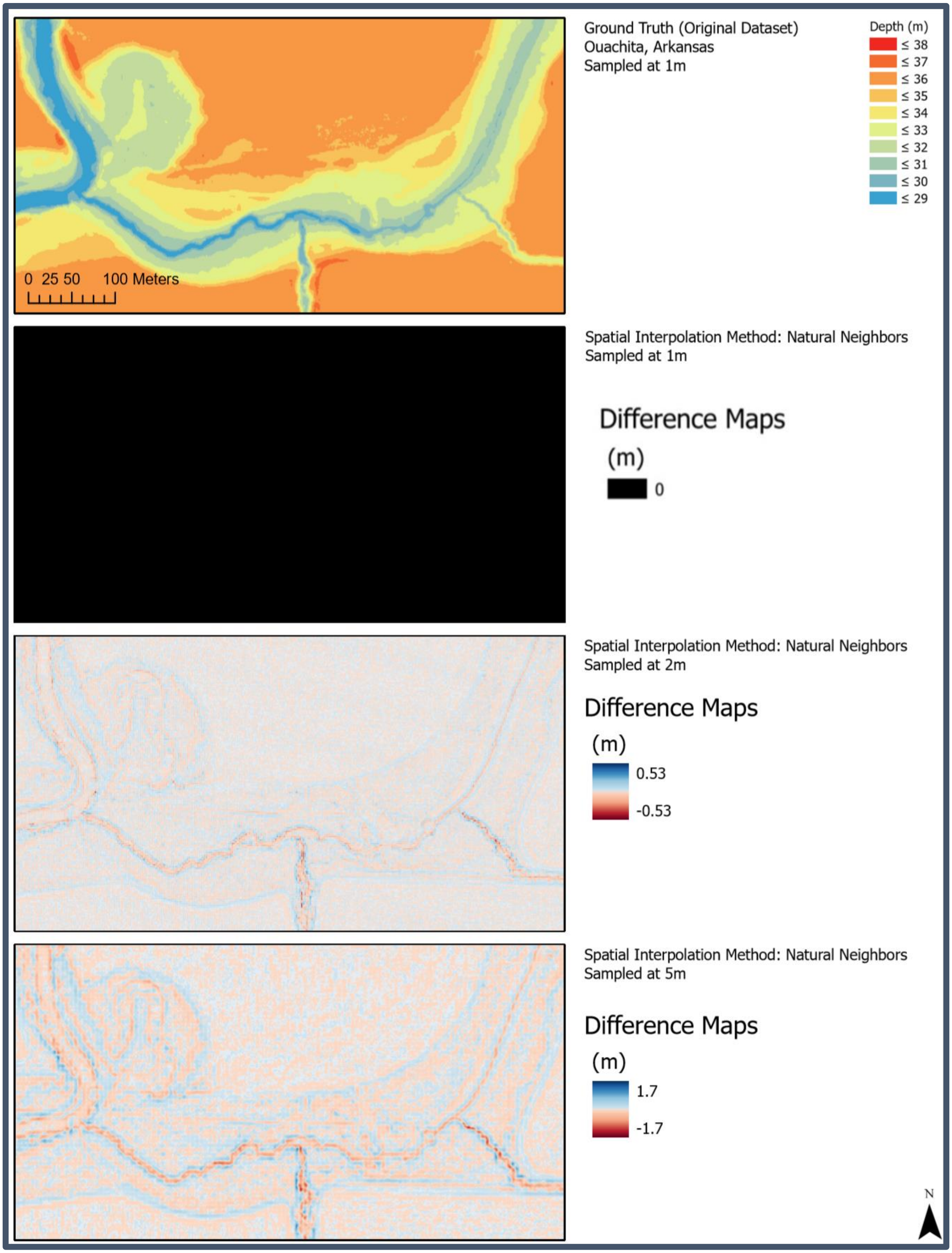

Figure 23 Difference Map Set for Natural Neighbors Interpolation in Slight Terrain Variation Environments

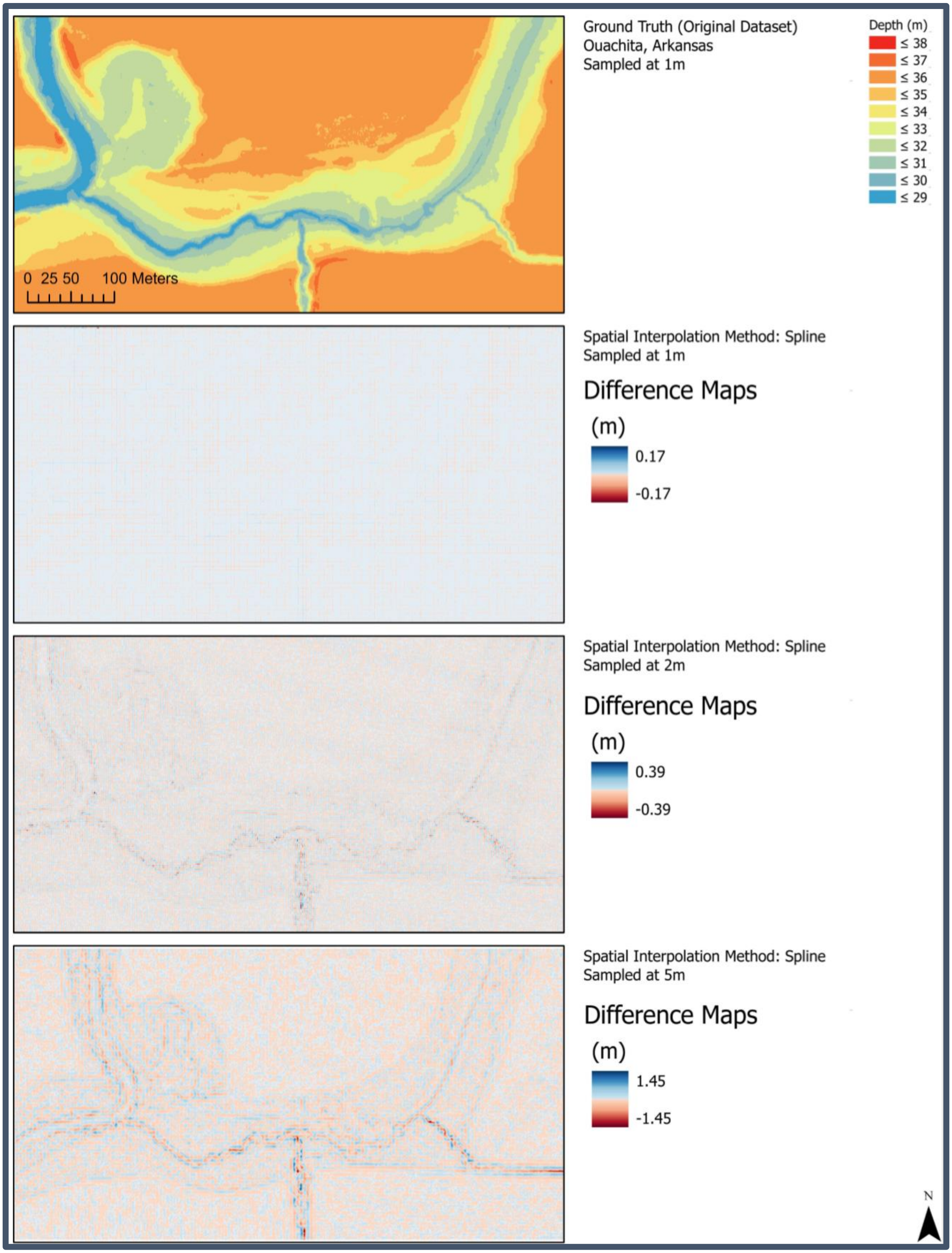

Figure 24 Difference Map Set for Spline Interpolation in Slight Terrain Variation Environments

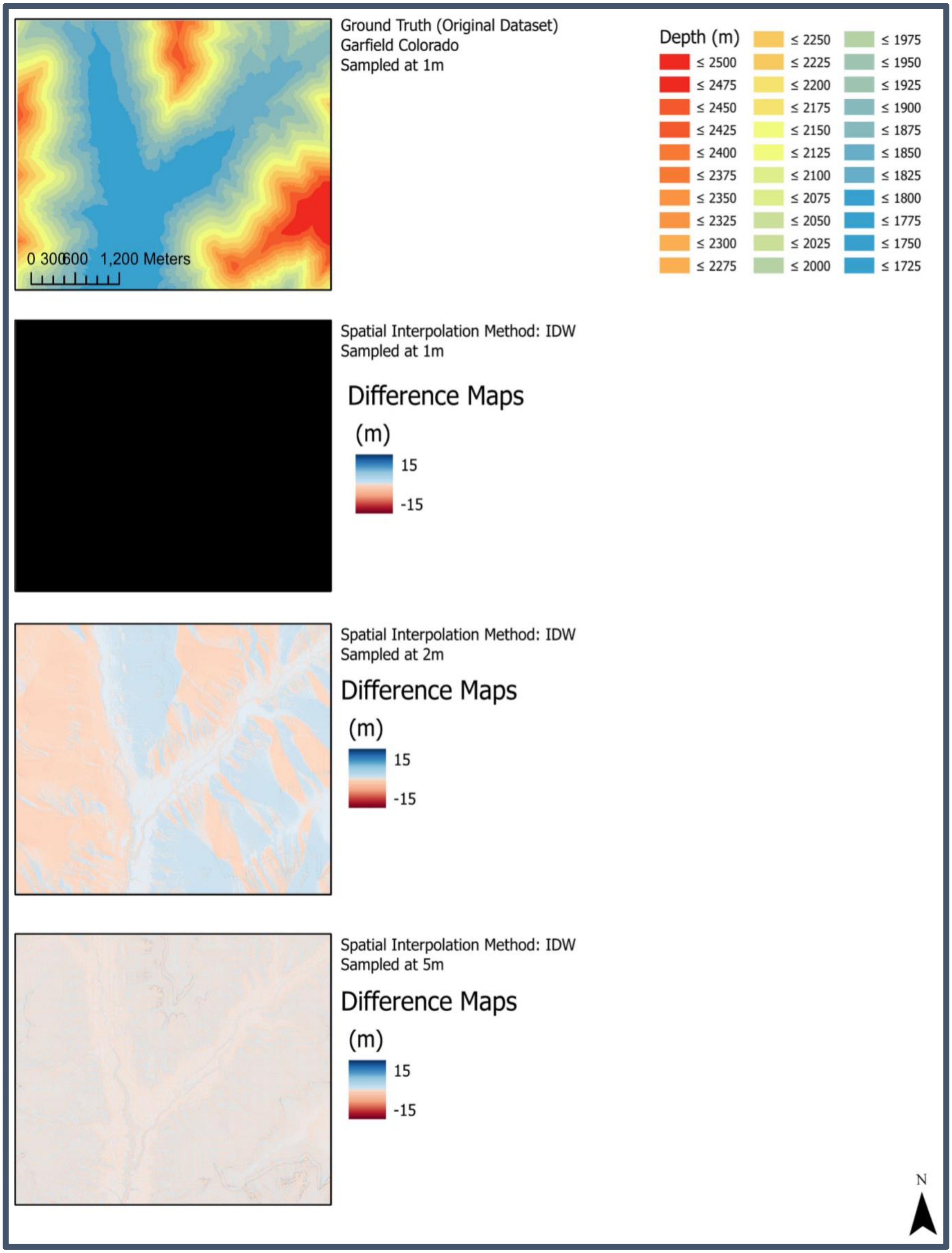

Figure 25 Difference Map Set for IDW Interpolation in Drastic Terrain Variation Environments

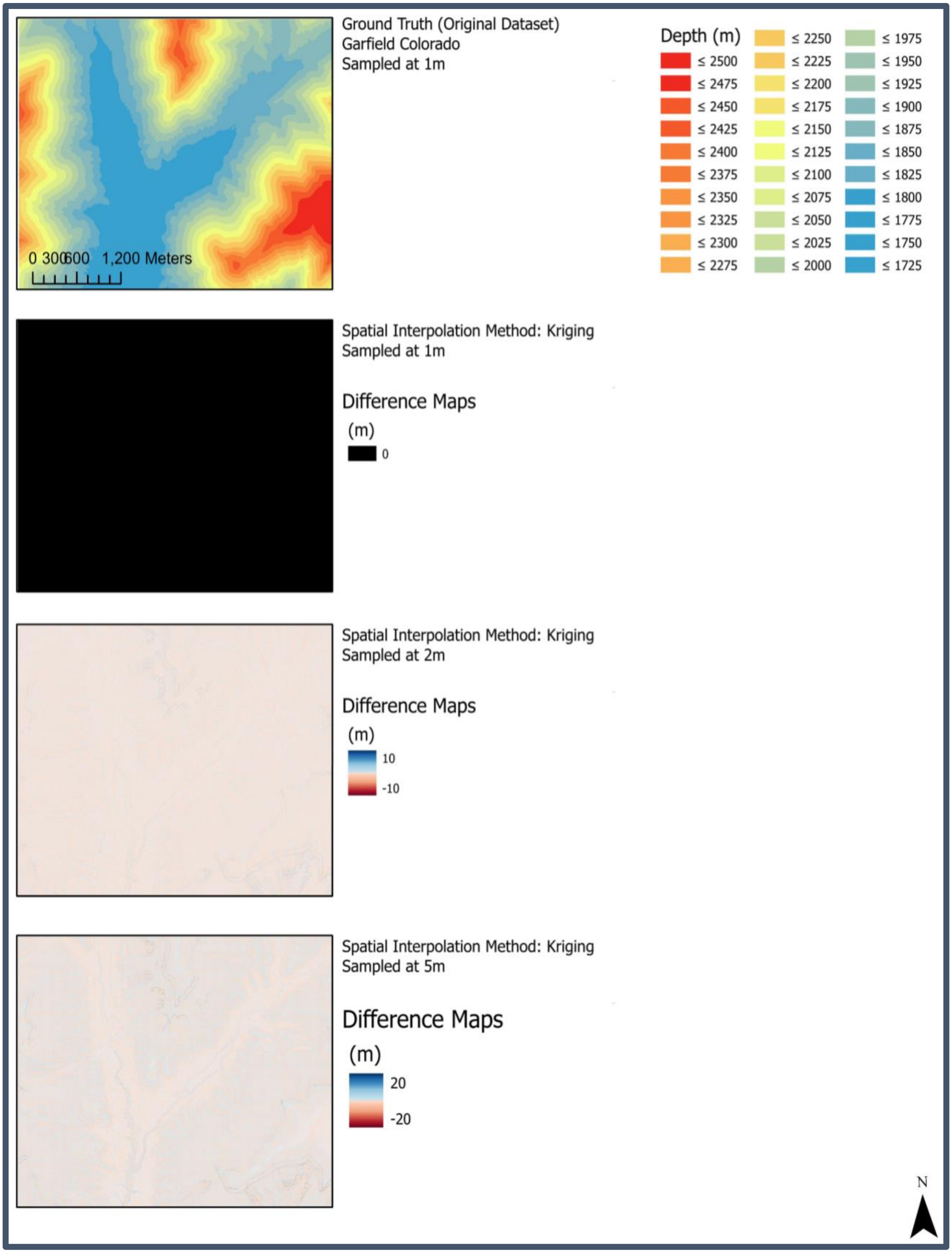

Figure 26 Difference Map Set for Kriging Interpolation in Drastic Terrain Variation Environments

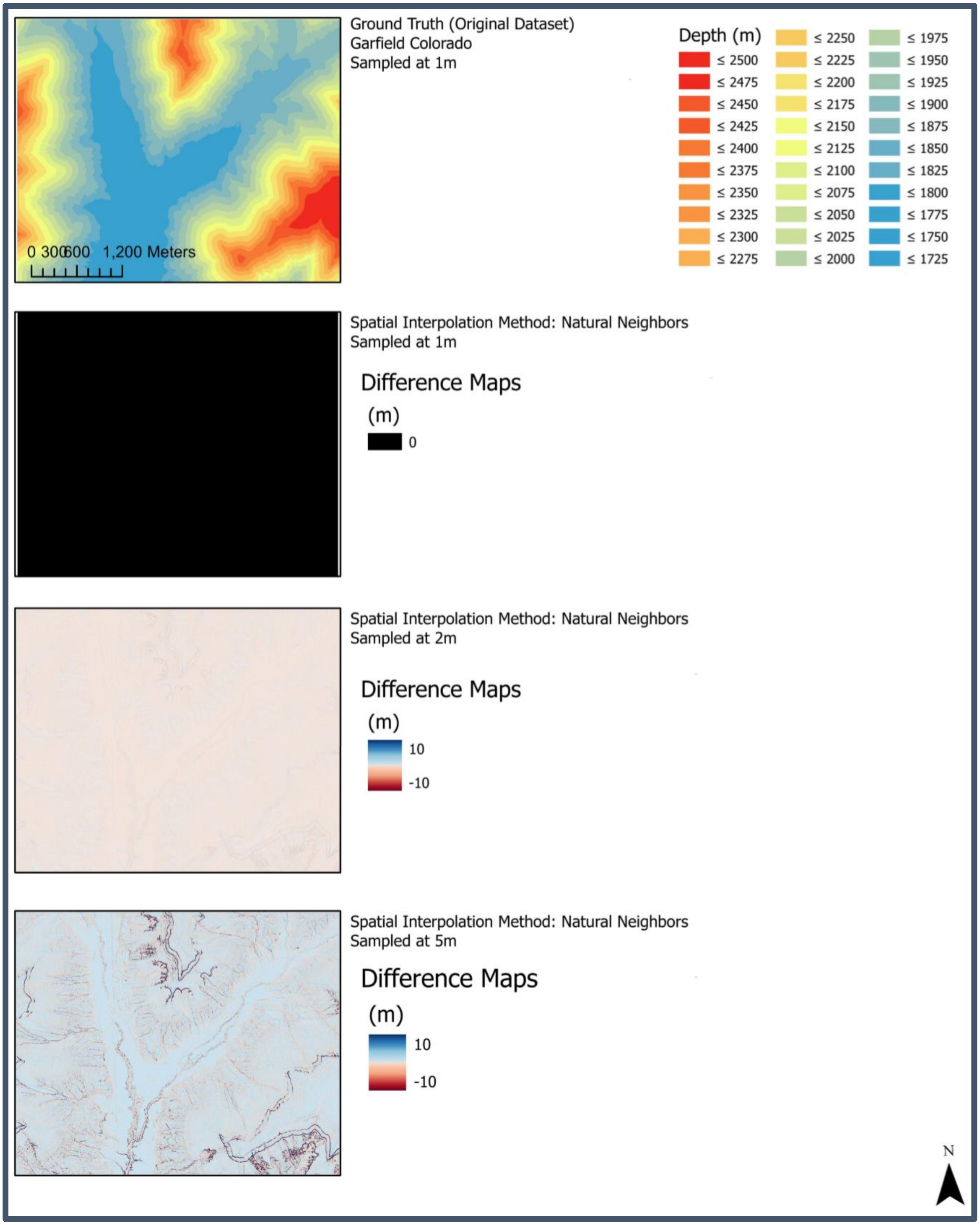

Figure 27 Difference Map Set for Natural Neighbors Interpolation in Drastic Terrain Variation Environments

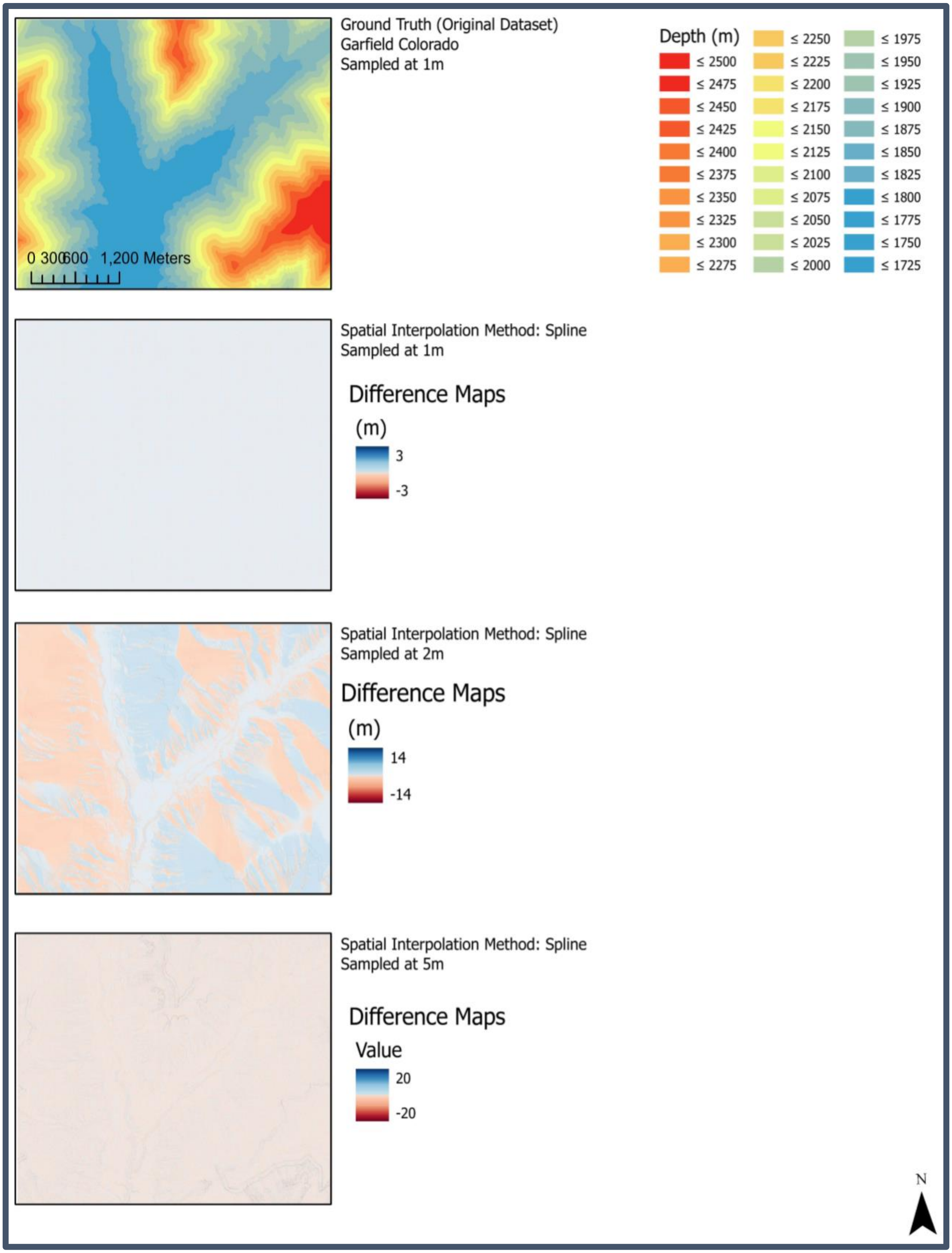

Figure 28 Difference Map Set for Spline Interpolation in Drastic Terrain Variation Environments