Appendix M
Part 4

Assessment of DEM Accuracy on the St. Johns New Madrid Shorebird Habitat Model

U.S. Army Corps of Engineers
Memphis District

Assessment of Digital Elevation Model Accuracy on the St. John's – New Madrid Shorebird Habitat Model

Michael Dougherty, Geographer
US Army Corps of Engineers, Rock Island District, 1500 Rock Island Drive, Rock Island, IL 61299
309-764-5491, Michael.P.Dougherty@usace.army.mil

Problem
The SJNM Shorebird Model was developed using the best available elevation data for the study area; a LIDAR (LIght Detection And Ranging) derived DEM was available for the New Madrid portion of the study area, while a 10m USGS DEM was available for the St. John's portion of the study area. During the review of the model (Battelle, 2011), the question of the appropriateness of the lower quality elevation data (10m USGS DEM) was raised.

Purpose
The purpose of this study is to assess the effect of digital elevation model (DEM) accuracy on the model parameters calculated for the Shorebird Habitat Model developed for the St. John's – New Madrid basin (SJNM), Missouri.

Goal
Although the primary purpose of this study is to determine the effect of DEM accuracy on the model parameters used in the SJNM Shorebird Habitat Model, the findings of this study apply to a wide range of DEM usage within the US Army Corps of Engineers (USACE). Therefore, this study will examine not only the particular DEM used in the development of the SJNM Shorebird Habitat Model, but will also examine a range of common DEM products and resolutions with the goal of developing practical guidelines to better inform practitioners in the application of DEMs to common USACE tasks.

Objectives
1. Determine the vertical accuracy of the 10m USGS DEM compared to the LIDAR derived DEM.
2. Conduct a sensitivity analysis to measure the effect of DEM inaccuracy on total estimated shorebird habitat.

Shorebird Model Approach
Based on a brief literature review, I believe that the modeling approach chosen for the SJNM Shorebird Habitat Model is a novel method and an extremely adaptable technique for estimating shorebird habitat using readily available DEMs, land cover classifications, and aerial photography. As with all spatial habitat models, a critical issue is: What is the finest resolution that the habitat can be mapped with existing data? Therefore, the challenge for applying this new modeling technique is to determine the finest spatial resolution that existing DEMs can support.

The strategy used by this approach is to incrementally raise the water level across the study area and record the newly inundated area at each step (while applying habitat quality factors), adding it to the total (as in integral calculus). Since one of the primary purposes of the model is to generate a habitat mitigation estimate, this technique is particularly appropriate in that only a single total area of shorebird habitat is required. Rather than a spatially explicit habitat suitability map, only a single area estimate is required.

One interesting possibility is that this high degree of aggregation could make the proposed total inundated area metric extremely robust to DEM inaccuracy. Even if the DEM is relatively inaccurate, the total inundated area metric may not be sufficiently sensitive to this inaccuracy to cause marked fluctuations in the final metric. Although Twedt did not express the advantages of the metric in exactly this way, he clearly describes the approach's advantages (Twedt, 2010, p 4., my emphasis added):

> Assumptions and Rationale –
>
> 1. The contour lines developed by USACE, Memphis District that are associated with 1-foot increments in Mississippi River stage, as recorded at the New Madrid gauge, *provide a reasonably accurate representation of the floodwater extent associated with each of these river stages.*

2. Use Geographic Information System (GIS) to derive interpolated elevations between 1-foot contour lines at <1-foot intervals (e.g., at 2 inch [5 cm], 4 inch [10 cm], or 0.1 foot [3 cm] intervals) so as to depict the theoretical distribution of floodwater extent associated with Mississippi River stages between the 1 foot river stages.
   a. Where possible, interpolation will be aided by LIDAR and DTM data. Elsewhere, interpolations will be based only on distance between contour lines.
   b. *Although distance interpolation may be imprecise, the assumption is that variation in flood area is averaged, thereby providing a reasonable approximation of the flooded area. Thus, this representation may not depict the exact geographic distribution of flooding but the total area inundated is presumed accurate.*

The essence of the approach is to incrementally raise the water across a DEM and record the inundated area at each step. As in integral calculus, the step-increment is chosen arbitrarily small to obtain an accurate overall estimate of the total volume, not because the step-increment is relevant to the application or the accuracy of the data. Too small of a step-increment and computational time is unnecessarily wasted, while too coarse of a step and estimate resolution is lost. Therefore, the question for this type of model is not whether a DEM will support 0.1 foot contours, but whether the calculation of a metric of this type requires a 0.1 foot step-increment to obtain a meaningful estimate of shorebird habitat.

As an ornithologist, and not a GIS analyst, Twedt's explanation of the GIS portion of the methodology is somewhat ambiguous and therefore potentially misleading. I believe a lack of documentation of the GIS methods has created much of the misunderstanding surrounding the review of this model. Because I was unable to find any detailed or specific description of the actually data processing steps used, I developed an approach (as there are always many) to operationalize and test the approach he describes.

Available DEMs

The question of the best available elevation data is always complicated by three main factors: 1. data collection methods, 2. availability date, and 3. areal extent. Usually there is not a perfect coincidence of these three factors and compromise is always necessary. The table below describes the DEMs available for this study area.

DEM	Elevation Collection Method	Source Agency	Areal Extent	Ground Condition	DEM Available
Lidar	LIDAR, breaklines, hydrographic cross-sections	USACE, PhotoScience	Mississippi River, Memphis District	2004	2004 (?)
DEM10	digitized USGS 7.5 Quandrangle hypsography, spot elevations, hydrography	USGS, Univ. of Missouri-Columbia, CARES	Missouri	1969	2003
NED10	Photogrammetric mass points, breaklines	USGS, NED	New Madrid Floodway	1995	2008
NED30	Photogrammetric mass points, breaklines	USGS, NED	New Madrid Floodway	1995	2008

Table 1. Characteristics of Available DEMs.

Although LIDAR typically has the highest quality (spatial resolution and elevation accuracy), it is often not available for the entire study area. In this case, LIDAR (flown for USACE in 2004) is available for the New Madrid portion of the study area, but not for the St. John's portion. For the St. John's portion of the study area, the best available elevation data is a DEM based on digitized USGS 7.5 Quadrangle hypsography (contour lines), spot elevations, and hydrography. This DEM is abbreviated as DEM10 in this analysis. This is one of the most common techniques used by USGS to generate DEMs throughout the U.S. when no better source of elevation data (photogrammetric mass points and breaklines, or LIDAR) is available. In 1995, USACE acquired elevation data (photogrammetric mass point and breaklines) for the New Madrid portion of the study area (but not for the St. John's) and this has now (since 2008) been incorporated into the USGS National Elevation Dataset (NED). Since the NED now uses best available data to build its NED products, the 10m (1/3rd arc-second), and 30m (one arc-second) NED products currently (April, 2011 version) use this 1995 elevation data as its source for the New Madrid portion of the study area. Apparently, the 2004 LIDAR dataset has not been incorporated by the USGS into the generation of the 10m or 30m NED at the time of this writing.

Therefore, based on this availability of DEMs, the Shorebird Model was developed using LIDAR source data for New Madrid and digitized USGS contours, spot elevations, and hydrography for the St. John's portion. These two data sources potentially represent opposite ends of the spectrum of DEM source data quality and create a problem of differing elevation data quality across the single SJNM study area. The question this analysis will attempt to answer is to determine the effect of this difference in DEM quality on the calculation of the shorebird habitat metric. That there is a difference in elevation accuracy between the DEMs is not as important to the review of the SJNM Shorebird Model as is the question of whether that difference in elevation accuracy substantially affects the shorebird habitat metric.

Key Study Question

Can DEMs derived from digitized USGS 7.5 Quadrangle hypsography, spot elevations, and hydrography (best available for SJ) produce similar results for the shorebird habitat metric as LIDAR derived DEMs (best available for NM)?

Methods

The analysis strategy employed in this study used the highest quality elevation source data available (Lidar) as the most accurate estimate of the true elevation. Other DEMs were then compared against this definition of the "true" elevation. Since LIDAR is available only for the NM portion of the study area, this study will focus analysis on the NM portion of the SJNM study area. This study examines the accuracy of the available DEMs for the NM study area, including the DEM derived from digitized USGS 7.5 Quadrangle hypsography, spot elevations, and hydrography (DEM10) which is the best available DEM for the SJ study area. By comparing these different types of DEMs in the NM study area, conclusions will be drawn about the relative suitability of various DEMs for calculating the shorebird habitat metric.

Analysis was performed using ESRI ArcGIS Desktop 10, Spatial Ecology's Geospatial Modeling Environment, and the R Language and Environment for Statistical Computing. All datasets used in this analysis and detailed process step documentation are available upon request from the author.

All DEMs were resampled to the same grid size (3 meters) using nearest neighbor resampling to preserve the original elevation values from the source DEMs. This was done to eliminate spatial resolution differences during the sensitivity analysis.

DEM accuracy assessment was performed using several methods. First, an effort was made to measure the accuracy of DEMs used in this study using National Geodetic Survey (NGS) high accuracy ground surveys (Order I). The Root Mean Squared Error (RMSE) statistic and boostrapped 95% confidence intervals (elevations were not normally distributed) were calculated to measure the difference between the 61 NGS control points available for the NM study area and the DEM values for each elevation grid (Table 2). Second, difference grids were calculated to measure the difference of each DEM from the Lidar grid and allow visual examination of the spatial distribution of error (Figure 1). Descriptive statistics, the RMSE, and boostrapped 95% confidence intervals (elevation data was not normally distributed) are reported in Table 3.

The shorebird habitat metric sensitivity analysis was achieved by calculating the amount of area inundated at each 0.1 foot increments and recording the marginal and cumulative area inundated at each step. An Python script was written using the ESRI ArcGIS Desktop 10 ArcPy API to model inundation. This script was run for each DEM in this study and results were imported into R for statistical analysis and graphing. Graphs of marginal (Figure 2) and cumulative inundated area (Figure 3) were created. Error statistics for marginal (Table 4) and cumulative inundated area (Table 5) were calculated with RMSE and boostrapped 95% confidence intervals (inundated areas were not normally distributed) being reported.

DEM Accuracy Assessment
NGS survey control points were used to measure the accuracy of DEMs used in this study (Table 2). These vertical accuracy calculations seem to indicate that the Lidar DEM is relatively inaccurate (2.11 ft RMSE). However, only a limited number of survey control points (61 1st Order) were available for the NM study area and the majority were not located in ideally flat terrain. Since many survey control points came from top-of-levee surveys, small horizontal displacements may cause the error statistic to be artificially inflated.

DEM	RMSE (ft)	95% C.I. Lower Bound (ft)	95% C.I. Upper Bound
Lidar	2.11	1.61	2.61
DEM10	2.23	1.69	2.83
NED30	2.77	2.31	3.22
SRTM	4.23	3.80	4.68

Table 2. Error statistics of various DEMs compared to 61 1st Order NGS survey control points. Root Mean Square Error (RMSE). 95% confidence interval calculated using bootstrapping functions found in the boot package for R. SRTM refers to Shuttle Radar Topography Mission, 90m pixel DEM.

Calculating the difference between the LIDAR derived DEM and the 10m DEMs provides an important visual and statistical picture of the accuracy of the DEMs relative to the LIDAR derived DEM. These maps highlight the spatial variability of elevation accuracy. Values close to zero indicate low difference (i.e., high accuracy, symbolized by yellows and greens on the map), higher values indicate high difference (i.e, low accuracy, positive values symbolized by red where the Lidar elevations are higher, and negative values symbolized by blue where the Lidar elevations are lower).

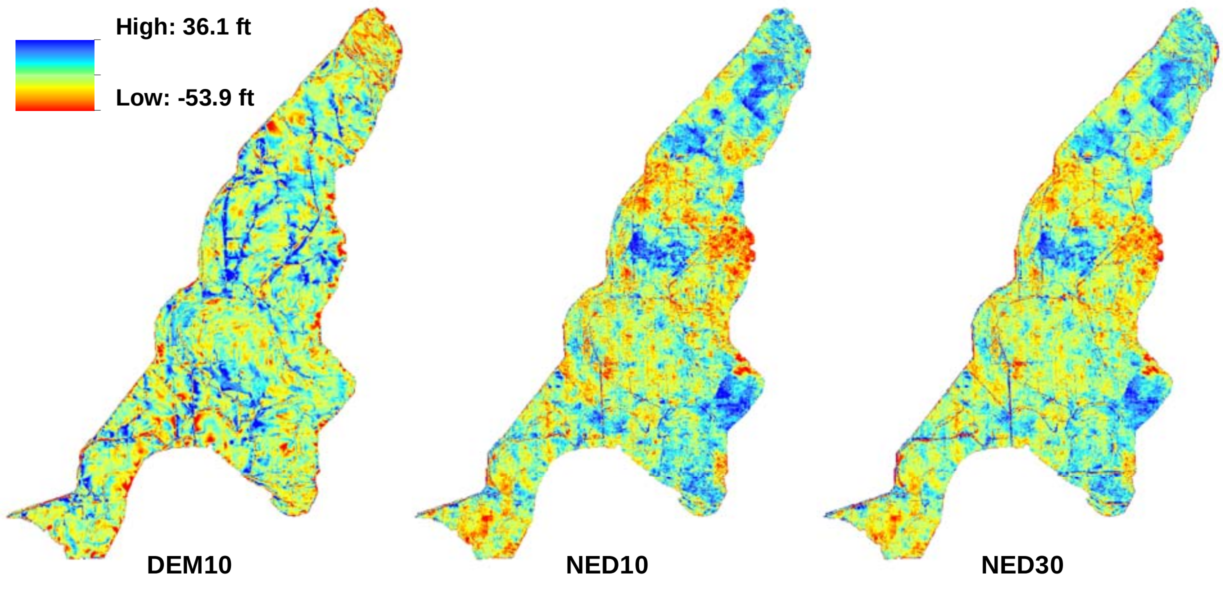

Figure 1. Differences between LIDAR derived DEM and other DEMS

Similarities in the spatial distribution of error between NED10 and NED30 are due to the fact that they are both derived from the same source data (1995 Photogrammetric mass points and breaklines). The differences between the NED DEMs and the DEM10 are due to the fact that it was derived from an entirely different set of source data (1969 digitized USGS 7.5 Quadrangle hypsography, spot elevations, and hydrography).

DEM	Min (ft)	Max (ft)	Mean (ft)	Standard Deviation (ft)	RMSE (ft)	95% C.I. Lower Bound (ft)	95% C.I. Upper Bound
DEM10	-52.53	36.14	-0.27	2.35	2.36	2.24	2.47
NED10	-51.11	25.38	-0.86	1.34	1.57	1.50	1.65
NED30	-53.86	30.36	-0.86	1.55	1.78	1.65	1.91

Table 3. Elevation differences of three DEMs from LIDAR derived DEM. Root Mean Square Error (RMSE). 95% confidence interval calculated using bootstrapping functions found in the boot package for R.

The newer 10m and 30m NED DEMs have means of -0.86 ft and a standard deviation of 1.34 ft and 1.55 ft respectively, while the older 10m DEM has a mean of -0.27 ft and a standard deviation of 2.35 ft. The three DEMs have means relatively close to zero, but the older DEM (DEM10) has greater spread around zero (indicated by the higher standard deviation) than the newer NED DEMs. The standard deviation indicates that for the newer NED10 and NED30 DEMs, ~66% of elevation values are less than 1.34 ft and 1.55 ft different than the LIDAR derived elevation. For the older DEM10, the standard deviation is 2.35 ft. This higher standard deviation indicates a higher degree of inaccuracy in the older 10m DEM when compared with the newer NED DEMs. The RMSE statistics in Table 3 also indicate that the DEM10 DEM is less accurate (RMSE 2.36 ft) than either the NED10 (RMSE 1.57 ft) or NED30 (RMSE 1.78 ft) elevation models.

Shorebird Habitat Metric Sensitivity Analysis

Since the DEM is being used to derive the shorebird habitat metric, a study of the appropriateness of a DEM must focus on the effects of DEM accuracy on the metric in question. The sensitivity analysis portion of this study developed a script to calculate a marginal and cumulative inundated area metric similar to the shorebird habitat metric. However, due to time limitations, this study did not carry the analysis beyond the area calculation step to apply the habitat quality weight factors. It was deemed sufficient for the purpose of assessing DEM accuracy requirements to stop at the evaluation of area, although testing effects on habitat quality would be fruitful if the model will be used more broadly.

The first step of the sensitivity analysis was to calculate the marginal inundated area at each 0.1 foot water level increments. Figure 2 displays the results of this analysis. This simulation raised the water level in 0.1 foot increments across each DEM (y-axis) and the marginal amount of land inundated at each step was recorded (x-axis). The Lidar DEM represented the highest spatial accuracy dataset and other DEMs were compared to this standard. In Figure 2, notice the distinctive error signature present in the DEM10 dataset. The clue to the source of this artifact is that the spikes are spaced at round 5 foot elevation intervals (285 ft, 290 ft, 295 ft, 300 ft, etc.). Upon further investigation, it was determined that these elevation intervals coincide with the contour intervals on the USGS Quadrangle sheet this DEM was derived from. The pattern in this graph results from DEMs derived from digitized USGS 7.5 Quandrangle hypsography (contour lines), spot elevations, and hydrography (streams/lakes/rivers). Identified as a common problem in NED DEMs, this effect is often referred to as terracing or ringing. This error structure is of particular because it is present in the best available DEM for the SJ portion of the study area.

The problem with this type of DEM error structure is that the inundated area estimate accuracy varies with elevation. A water elevation that happens to fall on one of the elevation contours (285, 290, 295, 300 ft, etc.) will overestimate the inundated area, while water levels that fall midway between the contour lines (287.5, 292.5, 297.5, 302.5 ft., etc.) will underestimate the inundated area. The error in inundated area estimated at any water elevation can be measured as the x-axis (inundated area) distance between the DEM10 line and the Lidar line. When the Lidar and DEM10 lines cross the inundated area estimates are the same.

The inundated area estimates for NED10 and NED30 track very closely in Figure 2 due to the fact that they are derived from the same underlying source data (photogrammetric mass points and breaklines). This is an interesting finding since it demonstrates that a coarser spatial resolution DEM (NED30, 30 m pixels) produces a similar inundated area estimate as a finer spatial resolution DEM (NED10, 10 m pixels), thus saving storage and computation time. Notice that the NED10 or NED30 DEMs derived from photogrammetric mass points and breaklines contain no apparent error structures in Figure 2. However, the NED10 and NED30 DEMs do possess a spatial error structure evidenced as a grid shape in Figure 1 (presumably the result of photogrammetric data processing methods used). This photogrammetric error structure is not obviously apparent in Figure 2.

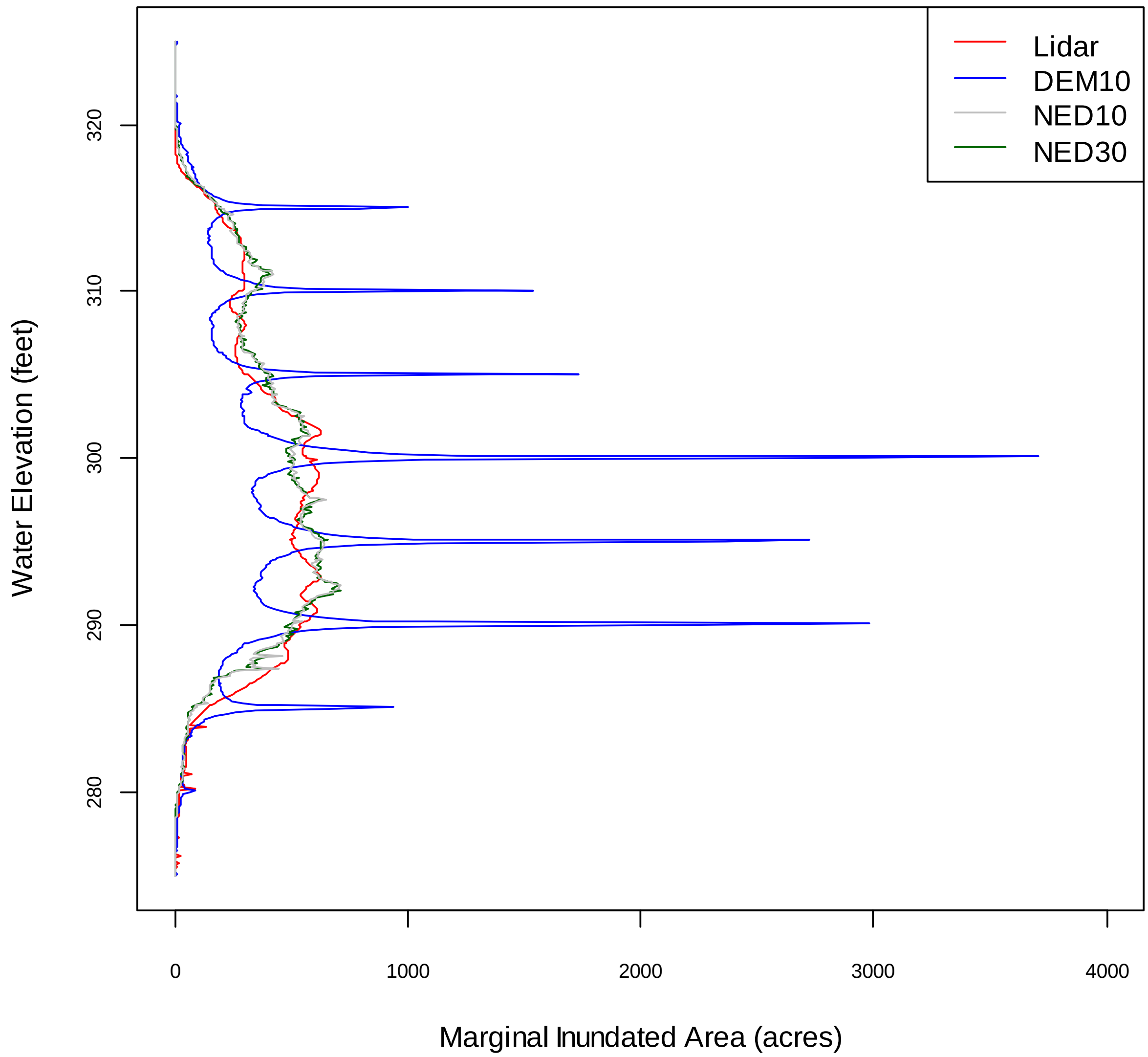

Figure 2. Marginal Inundated Area for several DEMs.

The next step in the sensitivity analysis was to calculate the cumulative inundated area at each 0.1 foot water level increments. Cumulative inundated area is the critical metric (not marginal inundated area) since the shorebird habitat metric accumulates the area inundated each day through the shorebird migration season. Figure 3 displays the results for this simulation as water levels are raised in 0.1 foot increments across each DEM (y-axis) and the cumulative amount of land inundated up to that water level was recorded (x-axis).

The significant observation from Figure 3 is that the estimates of cumulative inundated area for all DEMs roughly track the Lidar DEM, without major deviations. Despite the error structure present in the graph for the DEM10 marginal inundated area, that error structure does not affect the cumulative inundated area estimate. The DEM10 estimate follows the Lidar line as it periodically under- and then over-estimates cumulative inundated area, as discussed above. Again, the NED10 and NED30 datasets closely track each other as discussed above, but consistently overestimate the cumulative inundated area.

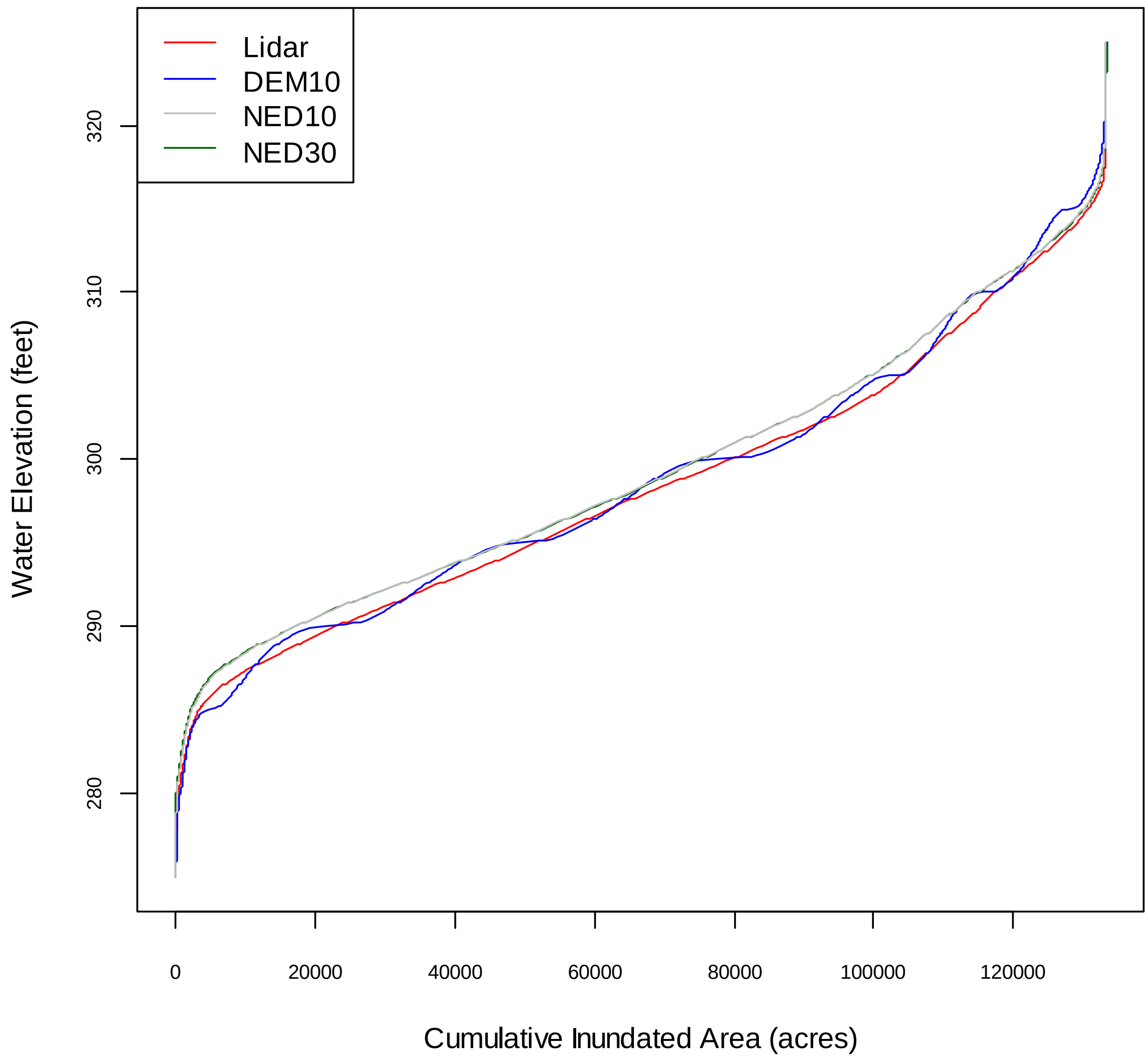

Figure 3. Cumulative Inundated Area for several DEMs.

DEM	RMSE	95% confidence interval lower bound	95% confidence interval upper bound
DEM10	312.9	222	408
NED10	56.5	52	61
NED30	57.4	53	62

Table 4. Error statistics of marginal inundated area (acres) of various DEMs vs. the Lidar DEM. Root Mean Square Error (RMSE). 95% confidence interval calculated using bootstrapping functions found in the boot package for R.

DEM	RMSE	95% confidence interval lower bound	95% confidence interval upper bound
DEM10	1,929	1,799	2,072
NED10	3,115	2,965	3,271
NED30	3,108	2,936	3,270

Table 5. Error statistics of cumulative inundated area (acres) of various DEMs vs. the Lidar Root Mean Square Error (RMSE). DEM. 95% confidence interval calculated using bootstrapping functions found in the boot package for R.

Analysis of the root mean squared error statistics for marginal and cumulative inundated areas reveals a similar pattern as the graph interpretation above. Although DEM10 has a higher marginal inundated area RMSE statistic (less accurate) than the NED DEMs (312 acres vs. ~57 acres), the DEM10 has a lower cumulative inundated area RMSE statistic (more accurate) than the NED DEMs (1,900 acres vs. 3,100 acres).

Discussion

The question of whether the DEM available for the SJ portion of the study area (similar to DEM10) can produce comparable estimates of shorebird habitat as a LIDAR derived DEM I believe is quantified by the results presented in Table 5. Despite DEM10 being a less accurate DEM than other DEMs (Table 3), and despite DEM10 having a peculiar error structure evidenced in marginal area calculations (Table 4), the aggregation of marginal inundation area estimates into a single cumulative inundated area metric largely erased these lower level errors. Constructing a metric from many incremental estimates is the basis of integral calculus and a common strategy in simulation. Unfortunately people are often only aware of the case of error propagation, but this is a classic case of the data analysis strategy of using aggregation to escape the low data resolution problem. This strategy seeks to escape the low resolution problem by moving up a scale level. Often this strategy provides a new way forward. By aggregating a large number of low spatial resolution estimates of inundated area into a single aspatial metric, based on the finds of this analysis, it appears that a single shorebird habitat estimate can be calculated.

Table 5 also indicates the uncertainty associated with these estimates and can be used to adjust habitat mitigation quantities. For example, rather than bear the added expense and delay to fly LIDAR, the shorebird habitat estimated for the SJ portion of the study area using a DEM similar to DEM10 could simply adjust its mitigation area estimate upward by the 95% confidence interval value. Model development using available data is the norm since the ideal data is seldom available.

Given the availability of LIDAR derived DEMs and their stunning quality relative to the USGS contour line derived DEMs of the recent past, I believe there was sufficient basis for concern about the use of these older DEMs is such flat terrain and being used for inundation mapping. The author’s interest in this question stemmed from the numerous flood-fighting inundation mapping accuracy questions that have been raised recently. Complex metrics such as the shorebird habitat metric involve so many calculations with varying factors that compelling theories could be invoked to advocate for either error propagation or error reduction outcomes. Ultimately, only a test calculation can determine how all of the competing theories and effects will work out in the final metric.

Recommendations

1. Far from being within the purview of only a single discipline, effective habitat modeling requires the knowledge and expertise of several disciplines that no one individual can possibly hope in a lifetime to possess. Effective habitat modeling requires an interdisciplinary team of individuals that command the varied skills of biology, ecology, geospatial methods, statistics, and simulation and modeling.

2. Involving experts with the above full range of skills from the beginning of the model design phase onward is the most cost effective method of model development. Trying to fix issues that are the result of a flawed design after the fact or could have been resolved early is frustrating and inefficient.

3. Effective documentation of the detailed data processing steps used in this analysis and made available during the peer review phase would likely have eliminated much of the confusion and misinformation surrounding the subject of DEM accuracy.

References

Battelle Memorial Institute (2011) *Draft Planning Model Quality Assurance Review Report for the Model Review of the Assessment of Shorebird Habitat within the St. Johns-New Madrid Basins, Missouri.* U.S. Army Corps of Engineers Ecosystem Restoration Planning Center of Expertise, Memphis District, Contract No. W912HQ-10-D-002, Task Order: 0008

Gesch, D.B. (2007) Chapter 4 – The National Elevation Dataset, in Maune, D., ed., *Digital Elevation Model Technologies and Applications: The DEM Users Manual, 2nd Edition*: Bethesda, Maryland, American Society for Photogrammetry and Remote Sensing, p. 99-118.

Twedt, DJ. 2010, Assessment of Shorebird Habitat within the St. Johns-New Madrid Basins, Missouri. Study Plan 2302-9S270. USGS Patuxent Wildlife Research Center, Vicksburg, MS. 26 pp.