Mapping Inundation Uncertainty

National Oceanic and Atmospheric Administration (NOAA)
Coastal Services Center

August 2010

NOAA Coastal Services Center
LINKING PEOPLE, INFORMATION, AND TECHNOLOGY

2234 South Hobson Avenue
Charleston, South Carolina 29405-2413
(843) 740-1200
www.csc.noaa.gov

Regional Offices:
NOAA Pacific Services Center, NOAA Gulf Coast Services Center, and
Offices in the Great Lakes, Mid-Atlantic, Northeast,
and West Coast

Table of Contents

Depicting Elevation Uncertainty in Inundation Maps

All forms of mapping have errors; some are known and others are not. Elevation data and the information that is derived from them are no different. Since mapping inundation using elevation data is becoming a very commonly performed geospatial analysis, this paper will attempt to answer the following questions about these errors, outline a new uncertainty mapping procedure that quantifies mapping errors, and communicate the results in easy-to-understand terms:

- How are the mapping results affected by the various known data errors?
- Are all areas of the derived results equal in confidence?
- Does the uncertainty vary with the inundation depth?
- What other non-elevation factors are involved and how are they incorporated into the uncertainty, if at all?

Overview

Mapping inundation using derived elevation data and tidal surfaces as the primary variables is a common practice. This type of mapping—often called a "single-value surface model" or "bathtub model"—has only two variables: the inundation level and the ground elevation. More complex hydraulic and geomorphic models, which include additional variables, can also be used to depict inundation. These models have their own error budgets, which can be complex, depending on model assumptions. This overview and description of error maps will focus primarily on the elevation surface variables and how their associated errors affect the resulting inundation maps.

A single-value surface modeling uses 1) a water elevation surface, which may come from a hydraulic model, overlaid on 2) the topography (elevation). Elevation is typically referenced to the North American Vertical Datum of 1988 (NAVD88), which is the vertical datum established for vertical control surveying in the U.S. NAVD88, however, it is not a tidal datum, so a value of "0" does not equate to any particular local tidal value. This can be an issue when dealing with coastal inundation that is inherently tied to local tidal variability (i.e., the ocean). To correct for this the elevation data can be converted from the NAVD88 datum to a tidal datum.

Use of tidal datum surfaces (i.e., a tidal correction surface) helps adjust elevations to a uniform tidal stage (e.g., mean high higher water) for a defined study area, but converting datums does create some uncertainty – the first source of uncertainty. Tidal elevations are not constant, however, often there are only a limited number of gauge stations to use when interpolating the tidal values across a large area. For example, mean higher high water (MHHW) at Station X is 3 feet (NAVD88) while at Station Y, 15 miles away, it is 3.5 Feet (NAVD88). Between these two stations there can be additional differences, depending on variables such as basin geometry, water depths, and location. VDatum is a software package designed to address these needs by providing higher accuracy tidal elevations for the US.

VDatum is a free tool developed by the National Oceanic and Atmospheric Administration (NOAA) and the National Geodetic Survey (NGS) that can be used to provide the tidal values for an area, and it is available for many locations (see http://vdatum.noaa.gov/). VDatum converts elevation values between NAVD88 and tidal datums using tidal stations and hydraulic modeling but has a level of error (see http://vdatum.noaa.gov/docs/est_uncertainties.html) on the order of 5 to 20 centimeters depending on location in the U.S.

The second source of uncertainty—a potentially higher source, with more spatial variability—in a single-value surface model is the elevation data itself. Digital elevation models (DEMs) derived from lidar data are used in many inundation mapping applications. Lidar is among the most accurate of the remote sensing elevation techniques, but the data from lidar can have limitations in certain land covers. In addition, not all lidar data are collected to the same accuracy standards; the vertical accuracy of the lidar can vary from 5 centimeters (root mean square error, or RMSE) to more than 30 centimeters (RMSE) and may vary within any one collection area. To quantify this variability, accuracy assessments are often performed with lidar collections. The results of the accuracy assessments help to document the errors and the statistical properties of the errors. The collection and processing factors that contribute to the errors, the process of determining the error values, and the tests that are performed are not covered in this text—but additional information on accuracy assessments can be found at www.csc.noaa.gov/digitalcoast/data/coastallidar/_pdf/What_is_Lidar.pdf. For inundation mapping, the most important result of the accuracy assessment is the vertical RMSE that depicts the accuracy of the data (NDEP, 2004). This value, the tidal datum error, and water level, is used to compute the uncertainty for each DEM cell.

An elevation-related error that is not included in the model is the inaccuracy of extending the tidal corrections landward. Generally, the tidal to NAVD88 conversion is only provided in areas where there are tidal waters. VDatum corrections typically do not extend far beyond the present shoreline, so some method of interpolation is required with the assumption that tidal conditions will not change dramatically with different water levels.

Mapping or Depicting Data Uncertainty

Previous Uncertainty Techniques

Clearly, the varying levels of elevation-surface and water-level-surface accuracy will affect the accuracy of inundation mapping. There are several techniques that can be used to depict this uncertainty. An example of a linear error (L.E.) technique was detailed by Gesch (2009); it describes the L.E. of the "inundation extent" (i.e., the line depicting the extent of inundation) based on the 95% confidence level of the elevation data. In its simplest form, the 95% accuracy value (1.96 x RMSE) is added to the mapped inundation extent to depict additional areas above the mapped area that may be flooded (Figure 1). Chapter 2 in the US Climate Change Science Program CCSP 4.1 – Coastal Sensitivity to Sea-Level Rise (CCSP, 2009) suggests that mapping inundation extents is akin to contouring and, as such, provides guidance on the minimum sea-level rise scenarios that can be mapped based on the underlying data accuracy. These values are directly related to the NSSDA (2003?) guidance on contour intervals; however, this technique does not include guidance on incorporating water surface uncertainties.

The tenet that mapping inundation scenarios is essentially a contour interval exercise is a distinct aspect of the techniques and guidance discussed above. The assertion that it is may not, however, truly capture the intent of this form of mapping. For example, mapping a single shoreline based on elevation data does not inherently require that specific data accuracy is achieved; i.e., there is no standard that stipulates a certain RMSE to map shorelines and comparison of various shorelines is performed regardless of accuracy. It is, however, common practice to detail the error estimates based on the data used and to incorporate that into any analysis performed.

The differences in technique or analogy to a commonly used process to standardize SLR mapping is still ongoing. The uncertainty mapping described in the following sections falls more in line with the shoreline analogy forwarded above in contrast to the contour interval analogy cited by CCSP and USGS.

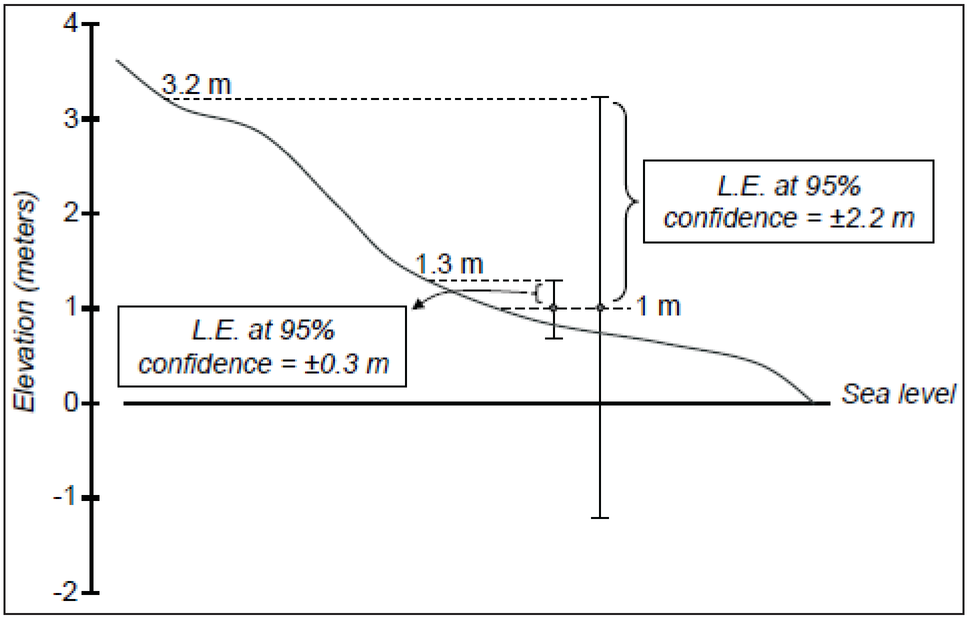

Figure 1. Diagram of linear error (L.E.) mapping for different elevation data sources (From Gesch, 2009). A 1 meter inundation level is shown with two example data sets: one with a 95% accuracy confidence of 2.2 meters, and one with a 95 % accuracy confidence of 30 centimeters.

NOAA Coastal Services Center's Elevation Uncertainty Technique

The techniques used to generate "areas of higher uncertainty" in the NOAA Coastal Services Center's sea level rise (SLR) and inundation mapping visualization tool are similar in principle to those cited by the USGS. They both use the reported RMSE of the data in most cases and its relationship to a normal error distribution. Some major differences are the use of an 80% confidence level instead of a 95% confidence level, the use of a cumulative percentage in determining the uncertainty (one tail vs two tail), and mapping this interval both above and below the inundation extent. Water level surface inaccuracies are also included; for many parts of the U.S. where VDatum coverage exists (http://vdatum.noaa.gov/about/availability.html), the standard deviation of the water level errors have been documented. In areas without VDatum, the errors may be greater and may have to be estimated.

In the simplest case, without any water level inaccuracies included, the inundation from a single-value water surface model is dependent only on the elevation uncertainty. Errors in elevation data are typically reported as either the RMSE or the Accuracy_z and inherently incorporate horizontal errors (FEMA, 2003). The Accuracy_z is computed from the RMSE by multiplying the RMSE x 1.96 when the data are normally distributed; Accuracy_z is the same as the 95% confidence level in these cases. In addition, RMSE is equivalent to the standard deviation (SD) when the data are not biased (DE SMITH et al., 2007). This is an important point for the NOAA Coastal Services Center method; this paper assumes that the RMSE is analogous to the SD (i.e., the data are not biased), which allows for the generation of a type of z-score or 'standard-score' from the data.

Mapping of the uncertainty takes on the same basic equation used to compute a standard-score (1) but substitutes RMSE for SD. The standard-score is simply the number of standard deviations a particular value falls from the mean (Figure 2).

$$\text{Standard-Score}_{(\text{value})} = (\text{Value} - \text{Mean}_{(\text{population})}) / \text{SD}_{(\text{population})} \qquad (1)$$

This equation is rewritten to substitute the inundation level (water surface) and the specific elevation at an X,Y location and RMSE of the elevation data set (2).

$$\text{Standard-Score}_{(X,Y)} = (\text{Inundation}_{(\text{water surface})} - \text{Elevation}_{(X,Y)}) / \text{RMSE}_{(\text{Elevation Data})} \quad (2)$$

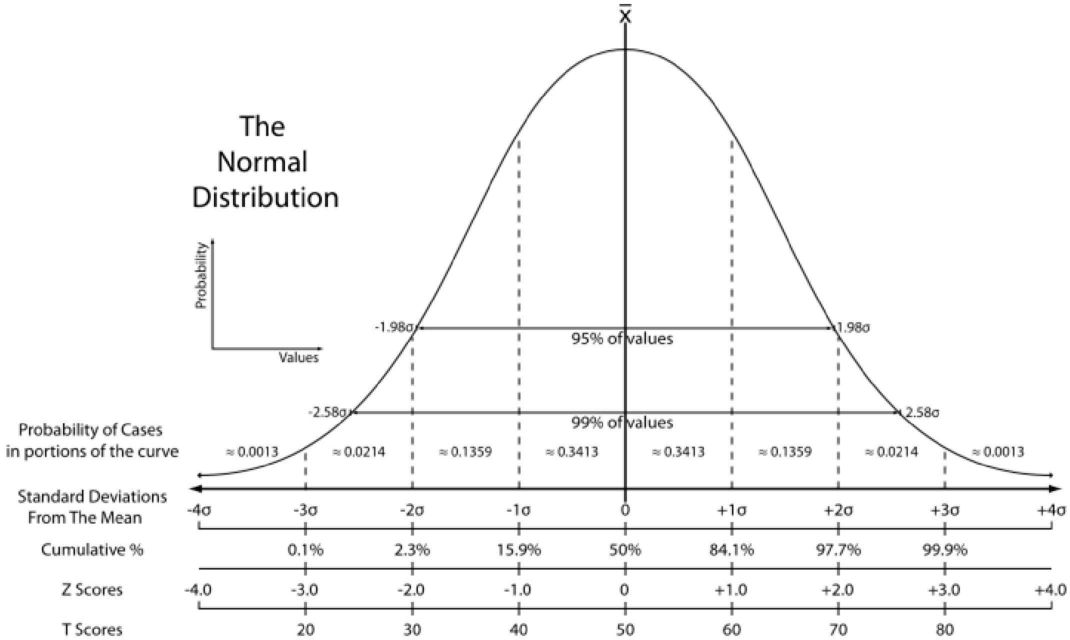

Figure 2. Z-scores/Standard-Scores and the normal distribution.

Using this equation (2), the standard-score can be defined for any elevation in relationship to a certain water level. For example, the standard-score at a given location with an elevation of 1.2 meters, which is based on a lidar data set that has an RMSE of 20 centimeters (0.2 meters) and an inundation level of 0.75 meters would be:

$$(0.75 - 1.2) / 0.2 = -2.25$$

This standard-score (-2.25) can then be used to find the percentile rank (Cumulative % in Figure 2). In this case the percentile rank, using a look-up table, would be 1%. So one could say, given the quality of this elevation data set, that this location has a 1% rank (chance) of being inundated by a water level of 0.75 meters.

The differences between this process and the L.E. process are the continuous nature of the standard-score computations (i.e., each elevation value has its own standard-score) and the use of the 1.96 x RMSE value to attain a 95% value. A standard-score of 1.96 above the level of inundation (what would be called -1.96) would equate to a 2.5% rank of flooding in this equation, not a 5% rank, since a cumulative or single tail approach is used (i.e., there is an assumption that values at the lower extreme are going to be inundated). The cumulative approach uses the entire area under the graph (Figure 3), not the discreet area between the standard deviations from the mean (i.e., the area between 1.96 and -1.96).

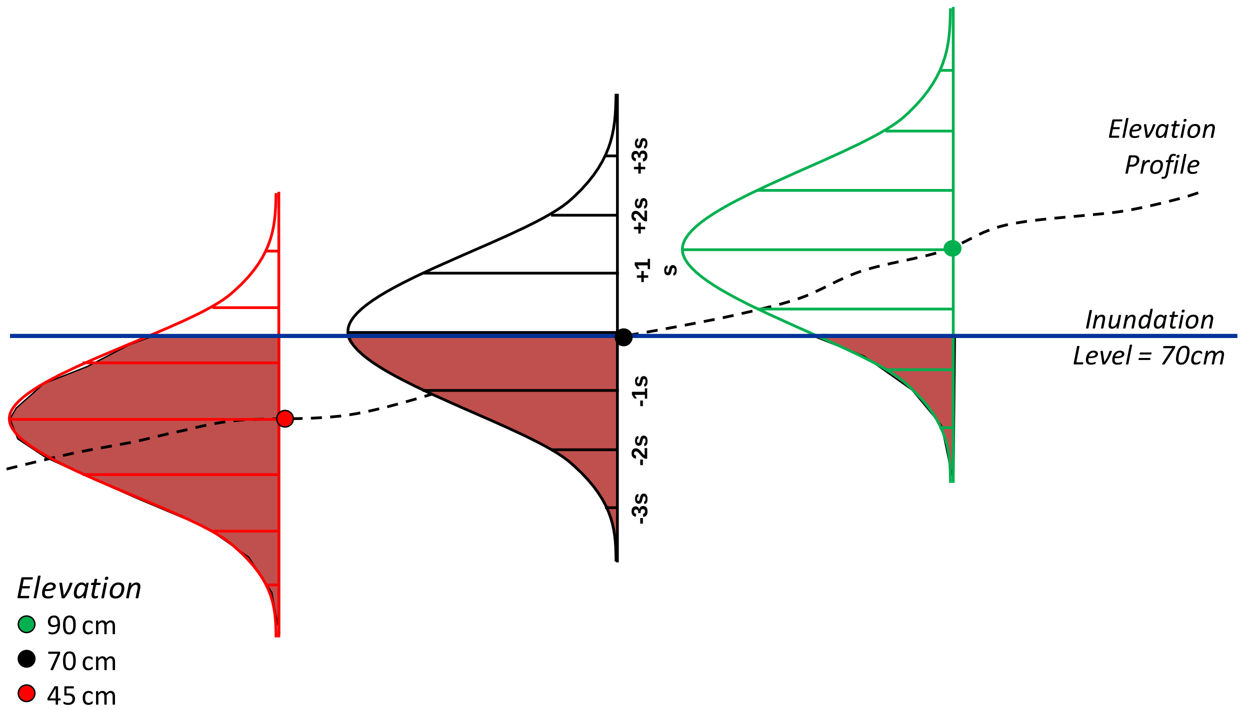

Figure 3. Example of NOAA Coastal Services Center method for several locations below, at, and above an inundation level.

These slight differences aside, the important aspects are the inundation level, elevation values, and their accuracy. For example, using the same 1.2 meter elevation and 0.75 meter inundation, but assuming the elevation value came from a data set with an RMSE of 100 centimeters (instead of 20) the rank score would be different.

$$(0.75 – 1.2) / 1.0 = -0.45 \text{ or a } 33\% \text{ rank}$$

Based on this data set, which is not as accurate as the first example, there is a 33% chance this location's elevation is enough in error that it would actually be inundated by a water level of 0.75 meters. With a rank of 33% this location would be mapped as a low confidence area (see below for mapping section), since there is a higher uncertainty in determining whether the location would or would not be inundated.

Mapping the Standard-Scores

Mapping the standard-scores is very similar to mapping elevation. The following examples illustrate a mapped inundation extent (Figure 4) and a confidence map based on a 50 centimeter water level rise (Figure 5). To begin, a single-surface water level (i.e., 50 centimeters above MHW) is mapped on the elevation data to provide water depths and the extent of inundation (Figure 4). Then, the corresponding values (water depth – initial elevation) are converted to a standard-score based on the accuracy of the initial elevation data set (Figure 5), which, in this case, has an RMSE of 30 centimeters. In Figure 5, areas with dark colors, blue or green, have high confidence in being inundated or not inundated respectively. Areas in grey or with lighter colors have a lower confidence of being inundated (i.e., below red inundation line) or not inundated (i.e., above red inundation line).

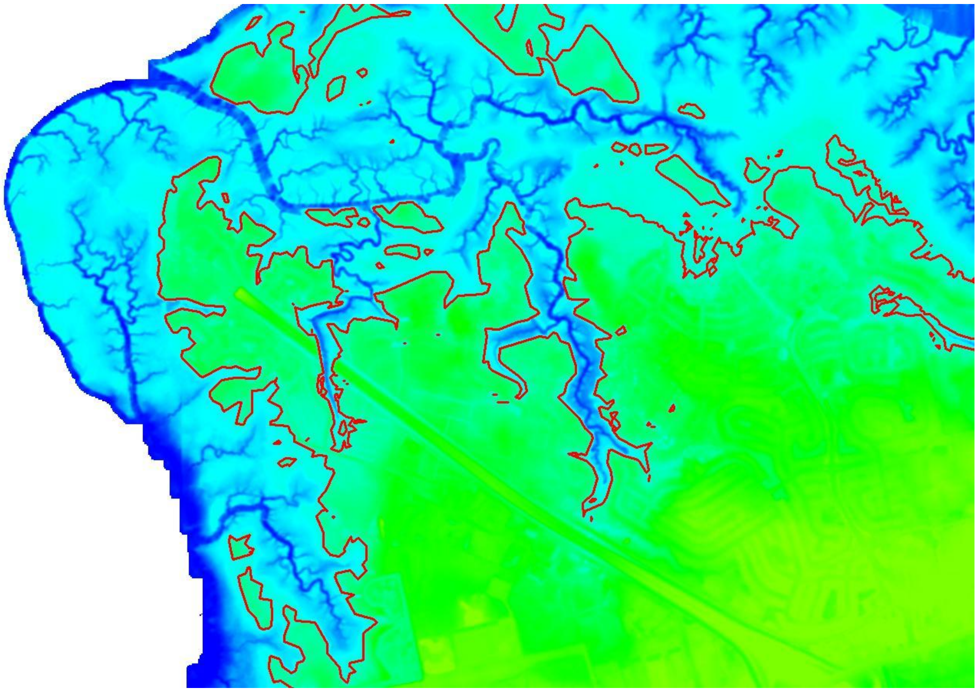

Figure 4. Inundation map of an example area showing the depths of water (darker blue to lighter blue), dry areas (light green to yellow), and the "mapped" inundation extent (red line).

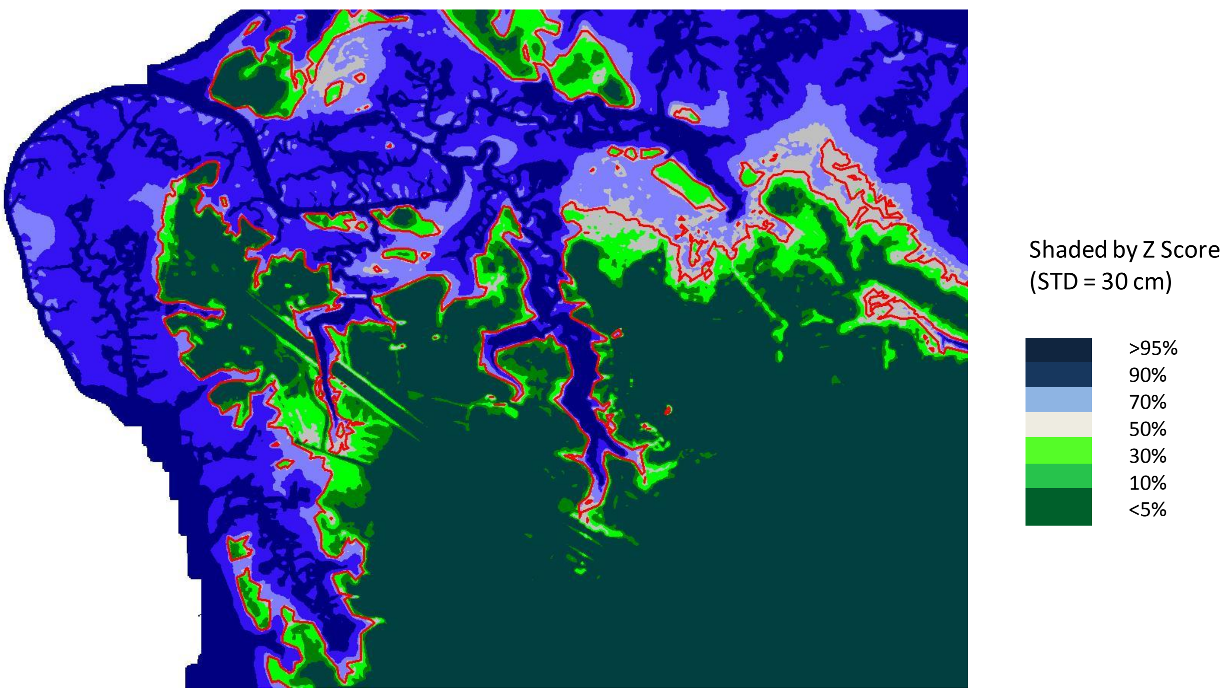

Figure 5. Z-scores of inundation data, where dark blues indicate a high chance of being inundated and dark greens indicate a very low chance of being inundated. Gray and lighter shades of blue and green indicate areas of lower confidence in the mapped result.

Mapped inundation extents do not, themselves, portray any sensitivity to data accuracy. An inundation boundary (red line in Figures 4 and 5) can be generated from data of any quality; the boundary (i.e., red line) simply corresponds to a standard-score of “0” (see Figures 2 and 3). The data quality is, however, highlighted in the standard-scores. For example, in Figure 6 the inundation level (50 centimeters) and resulting extent (red line) are the same as the scenario depicted in Figure 5; however, differences in the map trends (i.e., standard-score rank) are in response to a higher RMSE scenario (e.g., in this case 100 vs. 30 centimeters) of the data. In Figure 6, lighter shades of blue and green and gray dominate around the line of inundation, indicating lower confidence in the determination of “wet” (blue) vs. “dry” (green) areas. Another way to visualize this is to assume that the gray color is the width of the “line” of inundation. A wide line (i.e., a large grey area) will less accurately portray the extent of inundation than a narrow line (Figure 4).

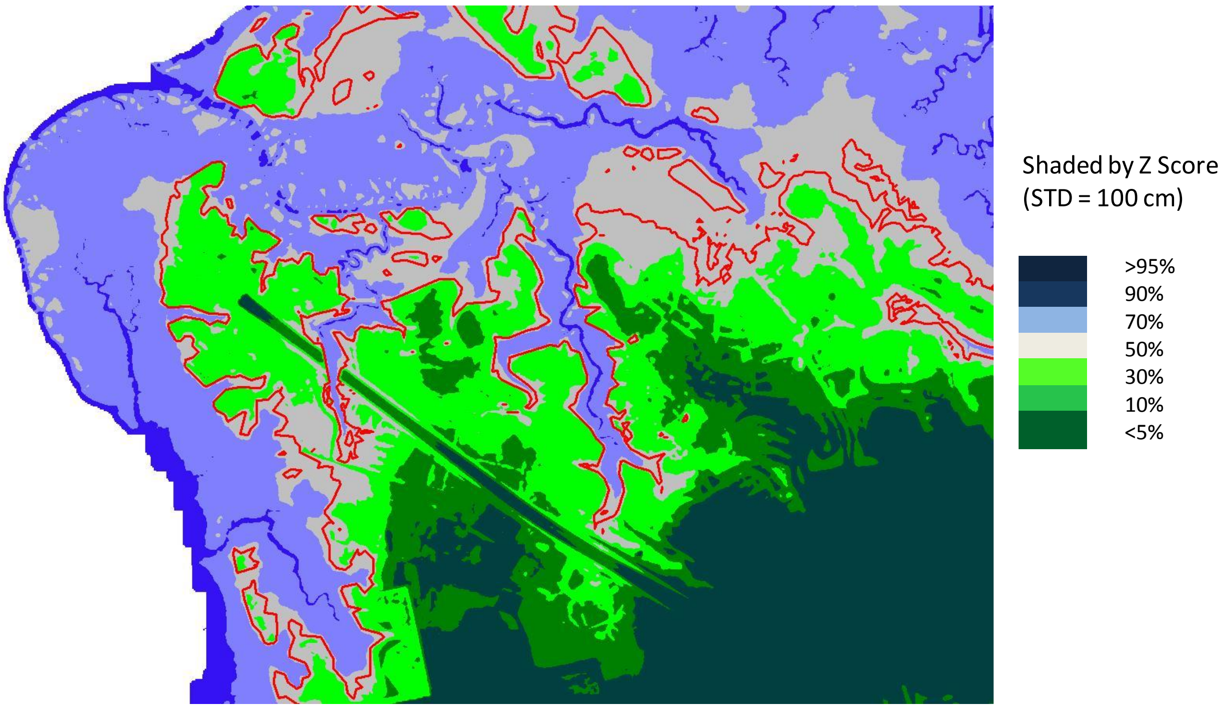

Figure 6. Z-score ranks from the same inundation level (50 centimeters) as in Figure 4 but with an elevation data set that has a 100 centimeter RMSE.

The presentation of the uncertainty in map form can be accomplished in many different ways. The simplest way is to show the areas of high confidence or low confidence. The NOAA Coastal Services Center has tentatively adopted an 80% rank (either as inundated or not inundated) as the zone of relative confidence. The use of 80% has no special significance but is a commonly used ‘rule of thumb’ measure to describe economic systems (Epstein and Axtell, 1996). Areas with standard-scores outside the zone of uncertainty (in this case 80%) are either shaded blue or left transparent (Figure 7); areas with a lower uncertainty are shown in orange. The choice of uncertainty levels (i.e., rank scores) that define the “confidence” is dependent on a particular user’s needs (Figure 7).

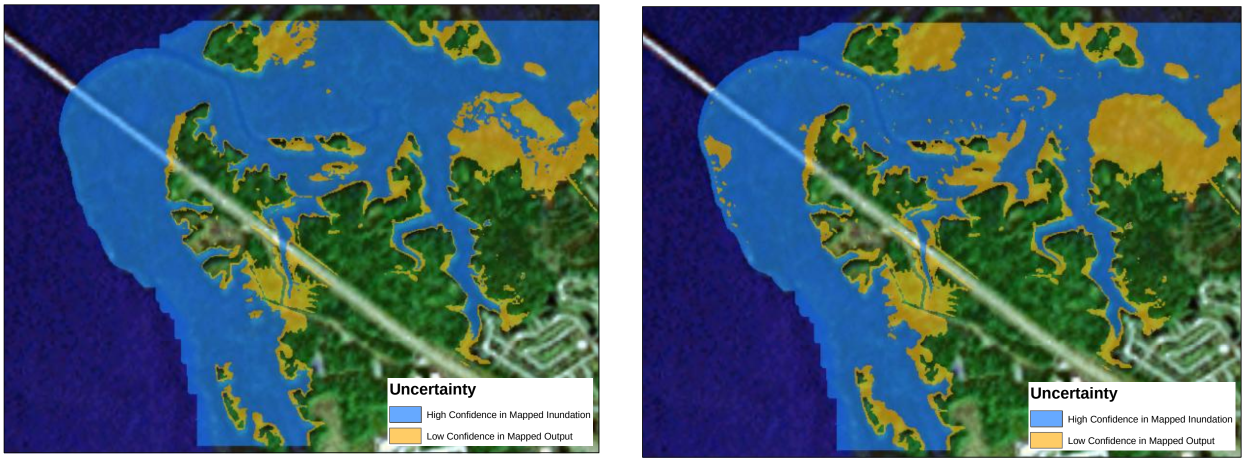

Figure 7. Simplified uncertainty presentation using an 80% confidence rank (left) and 95% confidence rank (right). Note the slightly larger area of orange color, especially in the marsh areas, when using a 95% confidence level.

Standard-Scores and Basic Uncertainty Concepts

The results of the standard-score outputs also provide some basic concepts about the uncertainty of inundation mapping. Most importantly, they highlight that the uncertainty is not uniform everywhere. Some areas have low uncertainty and others have a higher uncertainty, and this is often independent of data accuracy. This aspect has a great deal to do with slope; areas with low slopes will have higher uncertainty, since a small vertical error will generate a large horizontal error. Steeper areas, on the other hand, can have large vertical errors but only a limited horizontal expression of that error; a coastal bluff or scarp is a good example of where large vertical errors would result in small horizontal errors.

A less obvious aspect, but following from the slope discussion, is that the level of uncertainty can depend on the inundation level and where that level falls in regard to the local morphology (i.e., topography). A particular area can have low uncertainty at one inundation level and high uncertainty at a second inundation level simply because the second level falls at an elevation with lower slopes. One simple example would be an area with a low seawall or revetment fronting a parking lot. Low inundation extents would generate low uncertainty since the seawall or revetment is nearly vertical (i.e., has a very steep slope). Once the level of modeled inundation tops the seawall, however, there will be much higher uncertainty, since the parking lot is flat and any vertical uncertainty will cause large horizontal uncertainty. The result is that higher inundation levels do not necessarily have less uncertainty.

Incorporation of Water Level Uncertainty

To this point, the paper has focused only on the error associated with the elevation (topography) data. However, there is also error in determining the tidal elevations (i.e., the water surface). NAVD88 elevation values must be transformed to a local tidal vertical datum, and there is error associated with this transformation. The error associated with the water surface is not tied to elevation error, so a sum of squares error assessment can be made. This would be written as (3):

$$\text{Total SD} = (\text{Error (SD)}_1^{\,2} + \text{Error (SD)}_2^{\,2} + \ldots.\ \text{Error (SD)}_N^{\,2})^{0.5} \qquad (3)$$

For the purposes of a single-value water surface inundation model, the total SD error would be summed as below (4) based on the assumption that $RMSE_{(Elevation\ Data)}$ is equal to $SD_{Elevation}$:

$$SD_{Inundation} = (SD_{Elevation}^{2} + SD_{water\ surface}^{2})^{0.5} \quad (4)$$

Substituting the $SD_{Inundation}$ for $RMSE_{(Elevation\ Data)}$ in equation (2) would yield (5):

$$Standard\text{-}Score_{(X,Y)} = (Inundation_{(water\ surface)} - Elevation_{(X,Y)}) / SD_{Inundation} \quad (5)$$

This equation is essentially the same as that previously used, except that the RMSE term now incorporates the error associated with transforming the water level datum (i.e., going from NAVD88 to a tidal datum). In most areas the combined term of $SD_{Inundation}$ will not be significantly higher than the larger of the two errors. For example, given an RMSE of 20 centimeters for the elevation data and an SD of 10 centimeters for the water level error, the total error, $SD_{Inundation}$, is 22.3 centimeters.

Assessment of the error (i.e., SD) of the water surface/tidal correction is difficult in areas without VDatum, but in areas with VDatum coverage the error has been documented (http://vdatum.noaa.gov/docs/est_uncertainties.html). In areas without VDatum, this paper recognizes that the water surface/tidal correction has error but lacks a way to quantify it. Therefore the assumption is that the water surface as defined by tidal stations without the use of hydraulic modeling (as in VDatum) is in error of twice the amount of nearby or similar tidal areas that have VDatum error estimates. Moreover, the error in interpolating the tidal elevations inland is poorly constrained; it is nowhere documented. Given this situation, this study assumes that the interpolation of tidal elevations inland carries no additional error.

The mapped output of the uncertainty zone is portrayed in the same manner as in Figure 7, but it will have a slightly larger area of uncertainty. Areas depicted as having a higher uncertainty should be assumed to have nearly equal chances of being inundated or not being inundated given the specific water level setting.

Factors Not Included in Uncertainty Assessment

The single-value water level model is like taking a snapshot of the present physical conditions and holding them constant in the model run. In some cases and locations this is an acceptable assumption, but in many locations and under most inundation processes this is not the case. The coastal zone is highly dynamic, and various processes acting simultaneously will affect the extent of inundation; time and changes in energy act together to sculpt the coastal environment. Some of these processes include coastal barrier island migration and evolution, changes in sedimentation rates and deposition patterns, and changes in tidal elevations (i.e., coastal hydraulics) as a result of changes in land forms and their geometry. These factors are not included in simple single-value surface models or modified bathtub approaches that are based on interpolated water surface from tidal models.

Aside from the natural dynamics, the largest unknown and unincorporated factor in most single water surface models or, for that matter, geomorphic models is probably the dynamics of anthropogenic activities, which will be significant in some coastal areas (Titus and others, 2009; Gesch and others, 2009). It is assumed that most developed areas will be protected (Titus and others, 2009), but this is neither displayed nor factored into adjacent inundation extents in the Coastal Services Center’s single value surface inundation model.

The various unknown factors will affect the levels of uncertainty differently, depending on location and landform. It is probably sufficient to say that uncertainty in dynamic areas (i.e., barrier islands, spits) is actually higher than shown. That said, the mapped uncertainty in many undeveloped, low-lying, inland areas is probably fairly accurate when there is little additional “natural and/or anthropogenic uncertainty”.

Summary

Data errors that are directly related to elevation and water height data can be used to begin defining areas with mapped inundation that do not have the same level of confidence as other areas. The errors associated with most lidar and tidal data are not trivial and can approach 2 feet, which is on par with sea level rise estimates for 2100. The technique described only maps uncertainty in the elevation and tidal surface, not in the various geomorphic and anthropogenic variables and processes.

The technique and definition of “uncertainty” levels described herein is slightly different than other common methods (See Gesch, 2009), but the use of standard deviation to derive the assessment are similar. This study has tentatively chosen the 80% rank as the threshold between areas that have high uncertainty and those that seem well represented by the mapped output; however, any choice of uncertainty or rank can be used with this technique.

References

Epstein, Joshua and R. Axtell. 1996. Growing Artificial Societies: Social Science from the Bottom-Up, MIT Press, pp. 208.

De Smith, M. J., Goodchild,M. F. and Longley, P. A. (2007). Geospatial analysis: a comprehensive guide to principles, techniques and software tools. Leicester, Matodor.

Gesch, D.B., 2009. “Analysis of Lidar Elevation Data for Improved Identification and Delineation of Lands Vulnerable to Sea Level Rise.” *Journal of Coastal Research*. Special issue 53. Pages 49 to 58.

Gesch, D.B.; Guitierrez, B.T., and Gill, S. K., 2009. *Coastal Elevations*. In: Titus, James G., Eric K. Anderson, Donald R. Cahoon, Stephen Gill, Robert E. Thieler, and S. Jeffress Williams (eds.), 2009. *CCSP, 2009: Coastal Sensitivity to Sea-Level Rise: A Focus on the Mid-Atlantic Region*. Washington D.C.: U.S. Environmental Protection Agency, USA. 320 pages.

National Digital Elevation Program (NDEP). 2004. Guidelines for Digital Elevation Data: Version 1.0. National Digital Elevation Program. Accessed on June 22, 2010 at www.ndep.gov/NDEP_Elevation_Guidelines_Ver1_10May2004.pdf.

Titus, J.G., D.E. Hudgens, D.L. Trescott, M. Craghan, W.H. Nuckols, C.H. Hershner, J.M. Kassakian, C.J. Linn, P.G. Merritt, T.M. McCue, J.F. O’Connell, J. Tanski, and J. Wang, 2009. State and “Local Governments Plan for Development of Most Land Vulnerable to Rising Sea Level along the U.S. Atlantic Coast.” *Environmental Research Letters*. Volume 4, Number 4. Pages 44008 to 44015.