Derivation of Wetland Difference Products by Comparing the NWI Geospatial Dataset with C-CAP (10-m) and NLCD (2019) Data

A Technical Report

Zhenhua Zou [1], Chengquan Huang [1], Megan Lang [2], and Rusty Griffin [3]

[1] Department of Geographical Sciences, University of Maryland, College Park, MD
[2] National Wetlands Inventory, U.S. Fish and Wildlife Service, Falls Church, VA
[3] National Wetlands Inventory, U.S. Fish and Wildlife Service, Madison, WI

March 2022

Table of Contents

Derivation of Wetland Difference Products by Comparing the NWI Geospatial Dataset with C-CAP (10-m) and NLCD (2019) Data

Summary

The primary goal of this study was to improve the efficiency of NWI Geospatial Dataset workflows and better meet associated Congressional mandates by identifying areas of difference between the Geospatial Dataset and more contemporary land cover maps. This was achieved by examining the differences between the NWI Geospatial Dataset and two recent land cover datasets – the National Oceanic and Atmospheric Administration (NOAA) Coastal Change Analysis Program (C-CAP) 10 m land cover dataset and the 2019 Multi-Resolution Land Characteristics Consortium (MRLC) National Land Cover Database (NLCD) dataset. The resultant difference products highlighted wetland change hotspots for three change types: 1) wetland to impervious (W2I), 2) upland to open water (U2W), and 3) drier-end vegetated wetland to open water (V2W). Although the Difference Products were intended to map all differences, and not just those due to land cover change, the vast majority of differences reflect change. A robust examination of the C-CAP-based Difference Products relative to fine spatial resolution (~1 m) imagery showed that 85% (W2I), 83% (U2W), and 81% (V2W) of difference was due to confirmed land cover change. The overall accuracies of the C-CAP based products were 97% (W2I), 98% (U2W), and 96% (V2W). For the NLCD-based product, 85% (W2I), 80% (U2W), and 75% (V2W) of difference was confirmed as change. The overall accuracies of the NLCD-based products were 90% (W2I), 84% (U2W), and 87% (V2W). The impervious area estimates for individual NWI wetland polygons provided by the C-CAP-based polygon level product were highly correlated with estimates derived through visual interpretation of Google Earth (GE) imagery. The R^2 values were 0.79 and 0.97 for polygons having impervious areas < 5000 m^2 and >= 5000 m^2, respectively. For the NLCD-based polygon level product, the R^2 was 0.92.

Derivation of these difference products required innovative algorithms designed to minimize spurious differences caused by misclassification errors in the input land cover products and geospatial mismatches between those products and the NWI data. Implementation of these algorithms resulted in >15,000 lines of Python and JavaScript code. Throughout this study, more than 100TB of data were processed using >200,000 CPU hours, which would not be possible without using the massive parallel processing capabilities provided by University of Maryland's (UMD) High Performance Computer systems and the Google Earth Engine (GEE). Given the limitations of the input data, including both NWI and the two land cover datasets used, the difference products derived through this study were not intended for deriving change estimates for individual polygons or pixels. However, the change hotspots and other issues revealed by these products should be highly valuable for many applications related to the use, maintenance, and update of the NWI dataset.

1. Introduction and Objectives

The NWI geospatial dataset provides critical information on the Nation's wetlands but may be out of date in areas where substantial wetland changes have occurred after data production. Differences between this dataset and more recent land cover products may be a good indicator of those changes. Two land cover products are especially valuable for identifying change hotspots across the Nation. One is a 10-m land cover dataset produced by the NOAA Coastal Change Analysis Program (C-CAP). This dataset was derived based on high resolution imagery acquired from 2013 to 2017 for areas along the coast at a 10 m resolution, and hence may be useful for identifying wetland changes in coastal regions where NWI data were created before 2013-2017. The other dataset is the 2019 MRLC National Land Cover Dataset (NLCD). The NLCD dataset covers the entire conterminous United States (CONUS), and hence can support CONUS-wide analysis.

In order to support the maintenance and update of the NWI geospatial dataset, a suite of difference products have been developed by carefully comparing this dataset with the C-CAP and NLCD data. The comparison of NWI data with each of the two more recent land cover datasets resulted in two types of difference products:

- Pixel level difference product: This is a raster dataset that shows specific pixel locations where wetland changes might have occurred since NWI data production or there are other dataset differences that may need to be addressed.
- Polygon level difference product: This is a vector dataset that shows the amount of urban impervious surfaces in individual NWI wetland polygons.

An example of these two product types is shown in Figure 1. Together, the two product types can be used to identify wetland change hotspots for three change types:

1. wetland loss to urban development (mostly impervious surfaces),
2. change from upland to open water, and
3. change from drier-end vegetated wetland to open water (permanent water).

A description of these products is provided in a readme file (Appendix A). The following sections provide more details on the input datasets as well as the derivation and assessment of the final difference products.

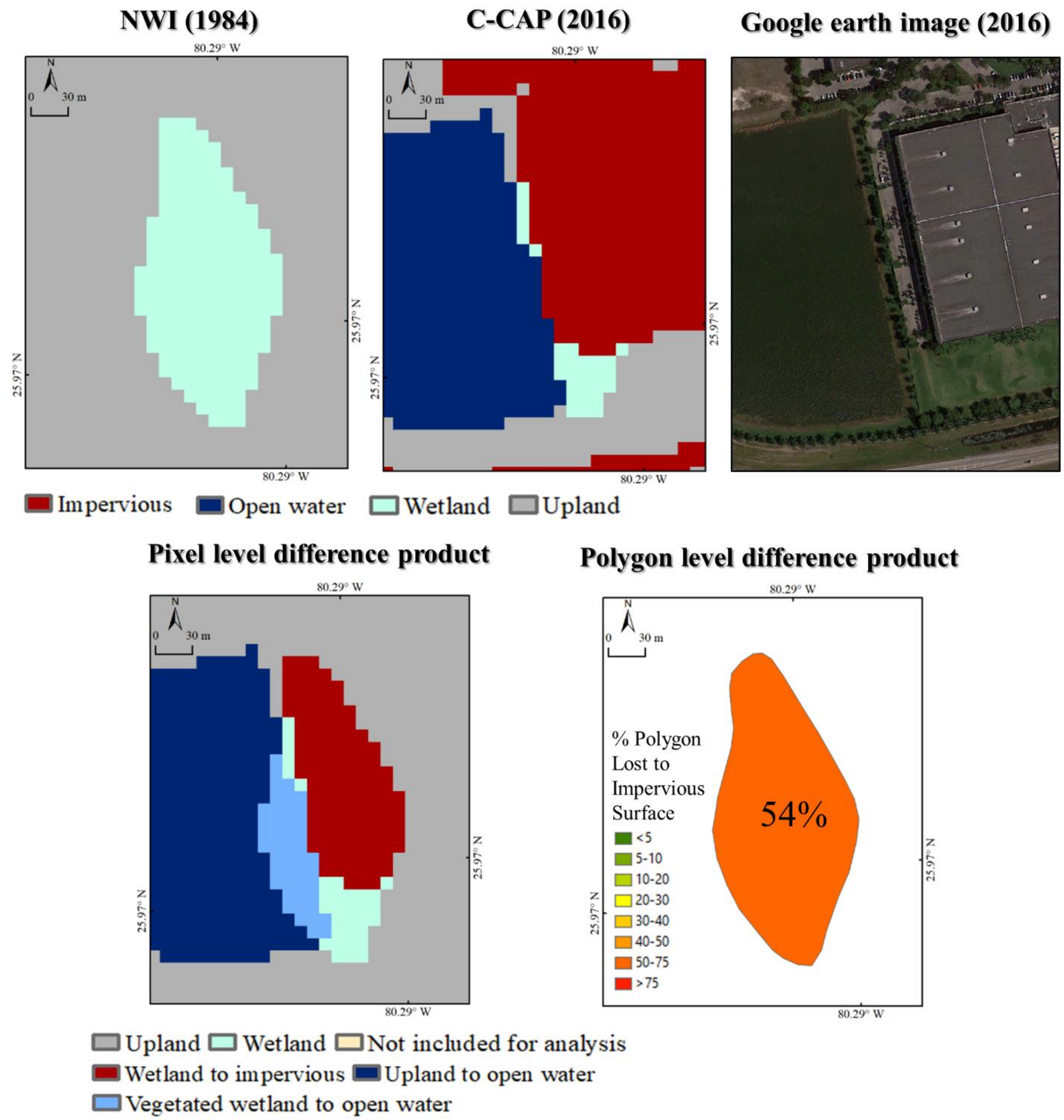

Fig. 1. An example of pixel (bottom left) and polygon (bottom right) level difference products derived by comparing the NWI and C-CAP 10-m data. The top row shows a simplified version of the C-CAP land cover map (middle), a raster version of the NWI data at the 10 m resolution (left), and a 2016 high resolution image (right).

2. Input Data

Three datasets were used to derive the difference products, including the NWI geospatial dataset, the C-CAP 10 m land cover data, and the 2019 NLCD data. In addition, we calculated temporal greenness and inundation metrics using Landsat images. These metrics were used to separate ephemeral flooding from more permanently inundated wetlands and to reduce errors in the NLCD data arising from misclassification of many small water bodies as impervious surface.

NWI geospatial dataset – The dataset used in this study was released in May 2021. It included updates made after 2010 in many areas. However, most of the country was mapped using imagery acquired before 2010 (Fig. 2).

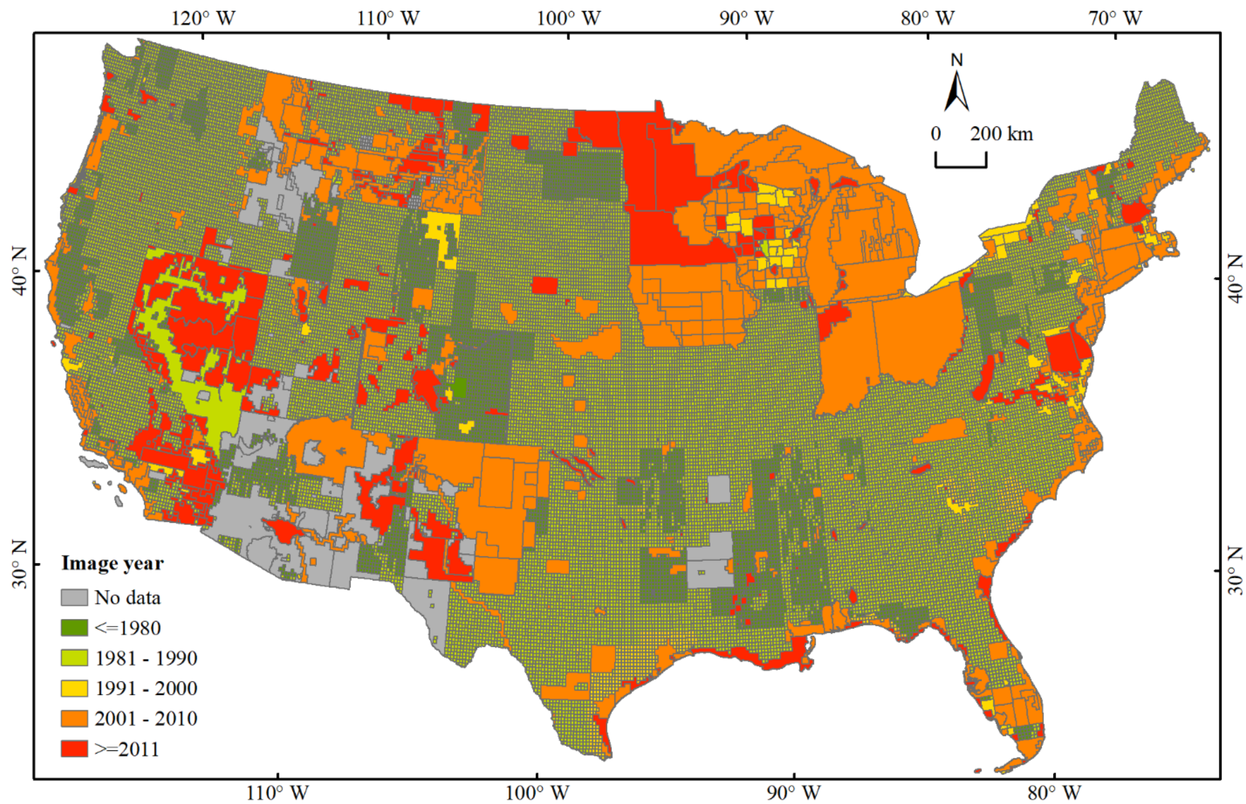

Fig. 2. The acquisition year of images used to create the NWI dataset used in this study.

Two groups of wetlands were excluded from this study. The first group included areas that were unlikely to have wetland loss to impervious over a few decades, including:

M1: Marine/Subtidal; M2: Marine/Intertidal

E1: Estuarine/Subtidal

L1: Lacustrine/Limnetic

The other group included all R wetlands (riverine), including many features derived through automated incorporation of the U.S. Geological Survey (USGS) National Hydrography Dataset. A comprehensive assessment revealed that these wetland polygons demonstrated too much geospatial inaccuracy for a meaningful comparison with the land cover data.

C-CAP data – This dataset consisted of 10-m land cover maps developed by NOAA's Coastal Change Analysis Program (C-CAP) for 25 states located along the coasts of the Atlantic and Pacific Oceans, the Gulf of Mexico, and the Great Lakes. These maps were derived by classifying high resolution imagery acquired between 2013 and 2017. The classifications were then aggregated to the 10 m resolution to produce the final land cover maps. It should be noted that for areas where the NWI dataset was updated recently, the C-CAP data could be as old as or even older than the NWI data. These areas were excluded from the C-CAP-based difference products, because differences between the two datasets within these geographies are less likely to be related to land cover change. Fig. 3 shows the location of the C-CAP data used in this study and the land cover types mapped by the C-CAP program.

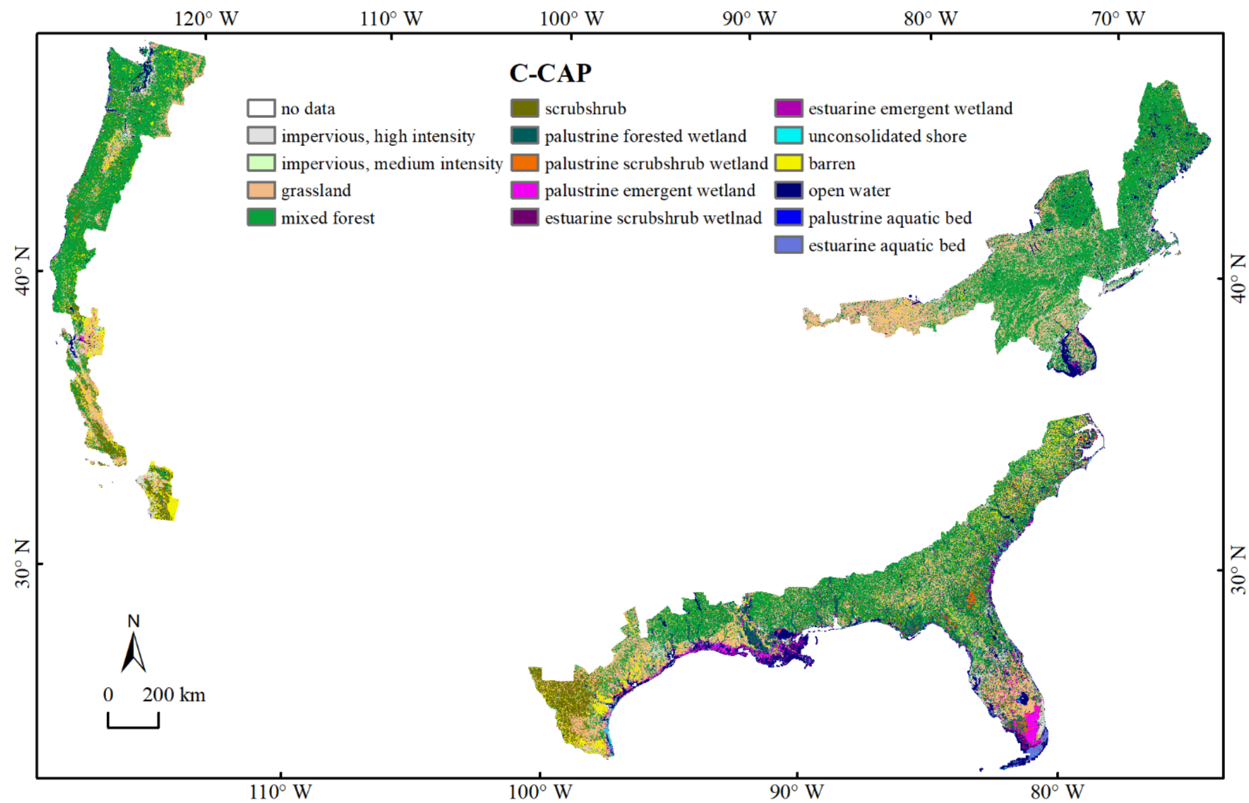

Fig. 3. A browse image of the C-CAP dataset showing the classification scheme and geographic coverage of the dataset

NLCD data – Led by the U.S. Geological Survey (USGS), the Multi-Resolution Land Characteristics (MRLC) Consortium produces National Land Cover Databases (NLCD) regularly. Most NLCD datasets include a land cover classification and subpixel fractional estimates of tree canopy cover and impervious surface mapped at the 30 m resolution for the conterminous United States (CONUS). The land cover map (Fig. 4) and impervious fraction layer (Fig. 5) of the 2019 NLCD dataset, which was the latest release available to this study, were used.

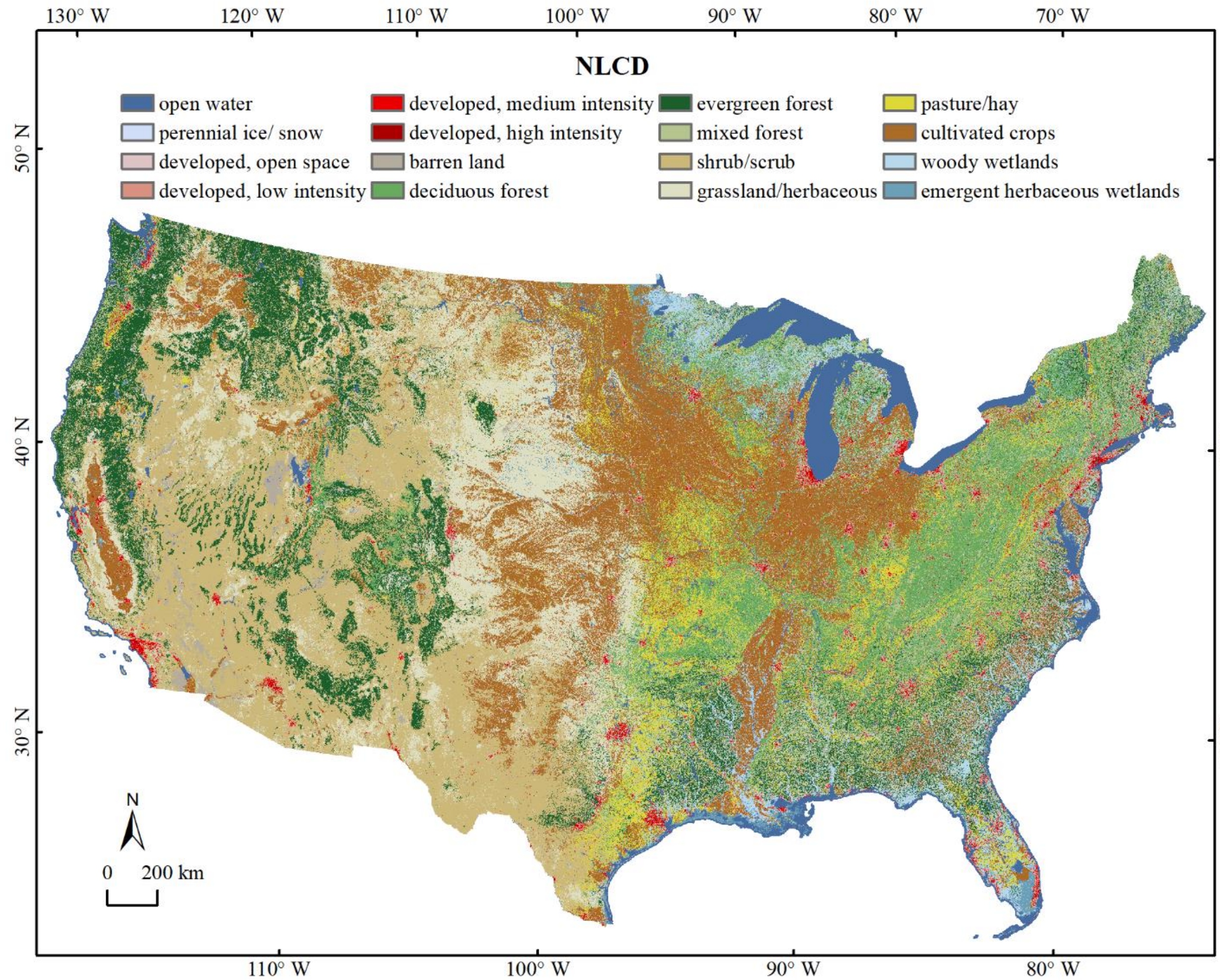

Fig. 4. Land cover classification map of the 2019 NLCD dataset.

Multi-temporal inundation and greenness metrics – These metrics were derived for CONUS by analyzing more than 15,000 Landsat images acquired during the growing season of 2018-2020. Inundation was mapped using the Dynamic Surface Water Extent (DSWE) algorithm (Jones 2015, Jones 2019). As an indicator of vegetation greenness, the Normalized Difference Vegetation Index (NDVI) was calculated using the red and near infrared bands. For each pixel location, temporal metrics such as water frequencies and median NDVI values were calculated based on all available clear view observations. Derivation of these metrics took several weeks using an unknown large number of computer nodes provided by the Google Earth Engine (GEE) cloud computing platform. Fig. 6 and 7 show two of the resultant metrics: inundation frequency and median NDVI.

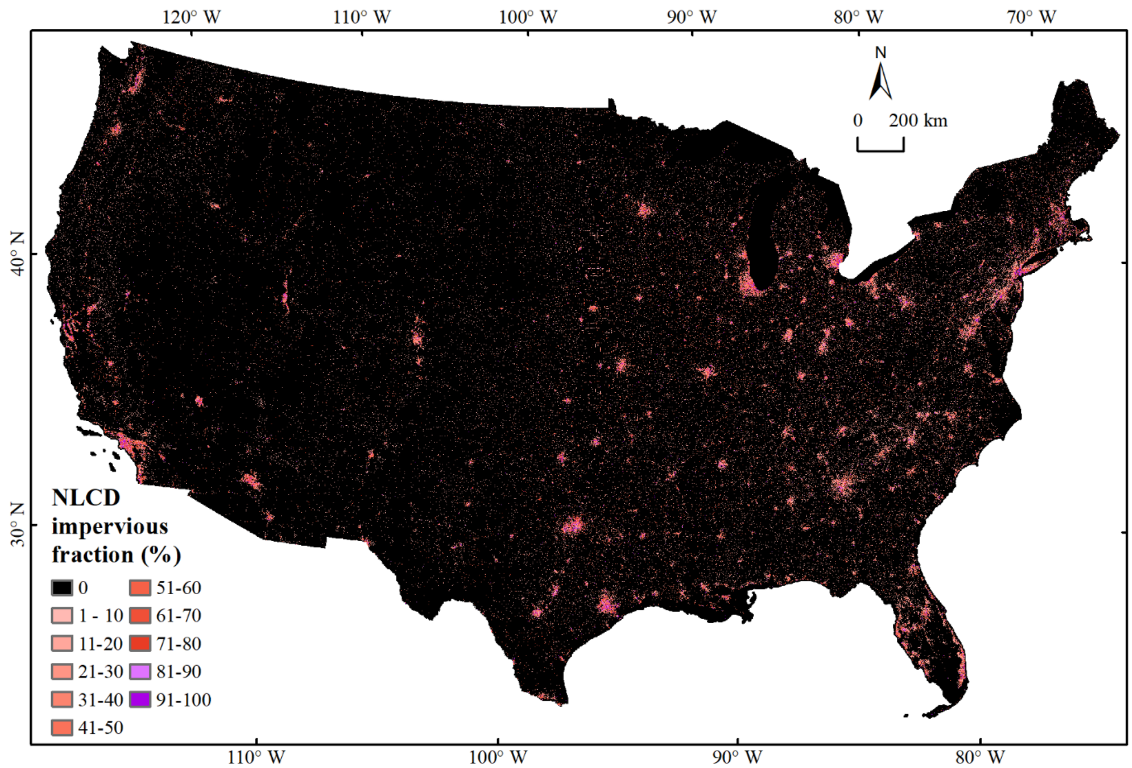

Fig. 5. Impervious fraction map of the 2019 NLCD dataset.

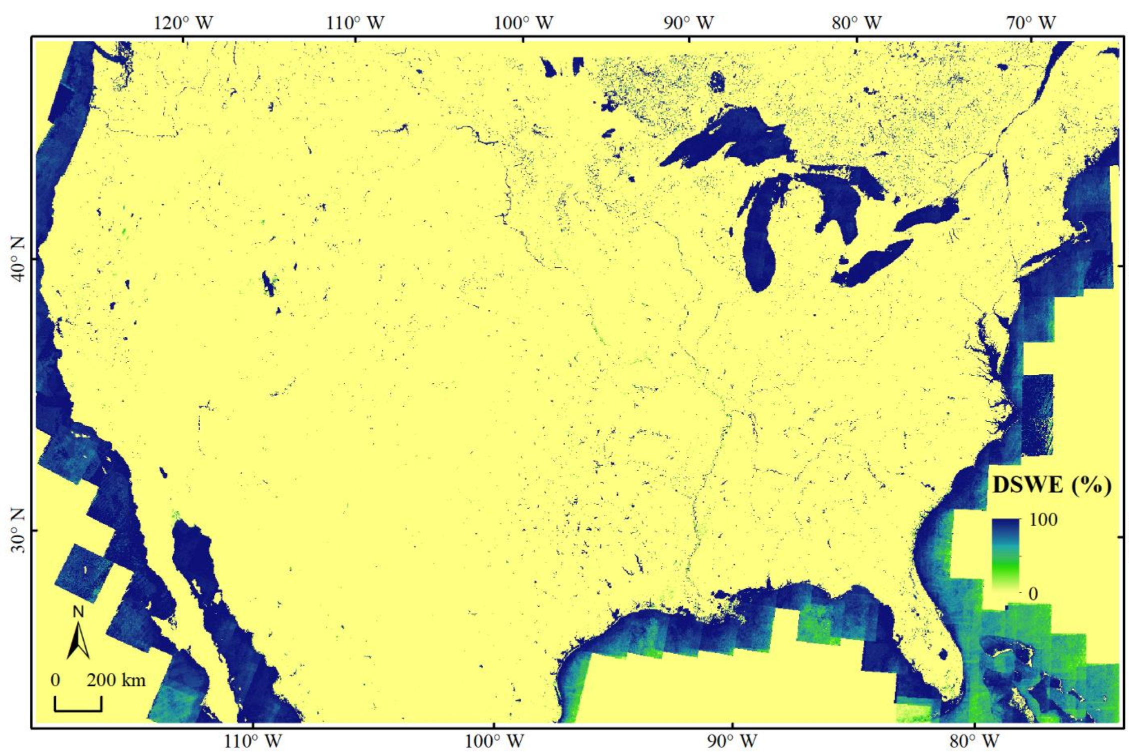

Fig. 6. The 3-year (2018-2020) inundation (DSWE) frequency (%) map for CONUS

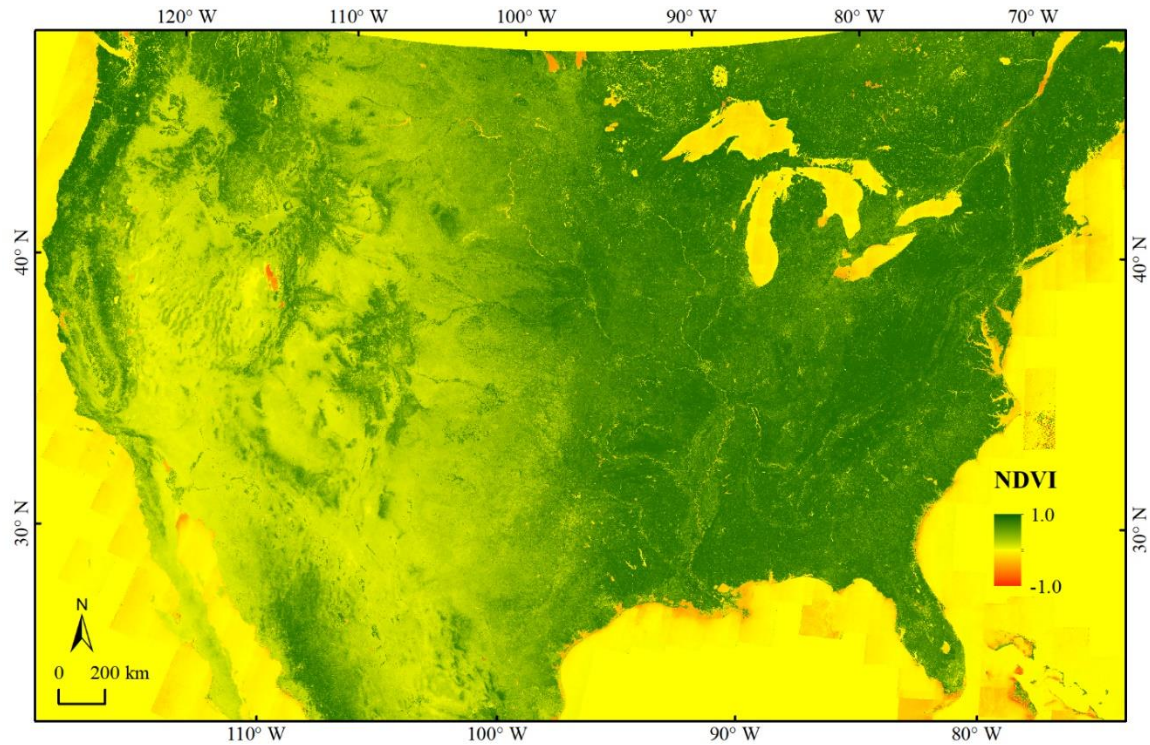

Fig. 7. The 3-year (2018-2020) median greenness (NDVI) map for CONUS

3. Development and validation of C-CAP-based products

3.1 Methodology

The C-CAP- and NLCD-based difference products were derived following a workflow consisting of several major steps. The first step was to reproject non-NWI data to have the same NAD83 CONUS Albers projection used by the NWI dataset. The second was to create subsets of all datasets based on the USGS Hydrological Unit polygons. We used HUC12 polygons in this study to create subsets that could be handled by available computing resources and to facilitate efficient multi-thread processing. There were 83,334 HUC12 watersheds across the CONUS, 17,054 of which overlapped with the C-CAP study area. A set of difference products were produced for each HUC polygon area in the third step. Once this was complete for all HUCs, the results were stitched together in the last step to create mosaics for the entire CONUS or the C-CAP study region.

3.1.1 Derivation of the pixel level product

To produce the pixel level C-CAP-based difference product, the NWI wetland polygons were rasterized to create a 10-m raster dataset whose spatial extent matched that of the C-CAP data exactly. The original NWI wetland types were regrouped into the following types:

- Upland
- Drier-end vegetated wetland
- Wetland and deep water not included for analysis (including M1, M2, E1, L1, and R, see section 2)
- Other wetland

Here the drier-end vegetated wetland class was created with the intention to map changes from this class to open water. This class was defined in consultation with NWI as follows:

- System:
  - E2 (Estuarine, Intertidal)
  - P (Palustrine)
- Vegetation types:
  - EM (Emergent)
  - SS (Scrub-Shrub)
  - FO (Forested)
  - Different combinations of the above (EM1/SS, EM2/FO1, …)
- Water regimes:
  - A (Temporarily Flooded)
  - B (Seasonally Saturated)
  - C (Seasonally Flooded)
  - D (Continuously saturated)
  - J (Intermittently Flooded)
  - N (Regularly Flooded)
  - P (Irregularly Flooded)
  - R (Seasonally Flooded-Fresh Tidal)
  - S (Temporarily Flooded-Fresh Tidal)

Throughout this report, the term vegetated wetland refers to these drier-end vegetated wetland types. The original C-CAP classes shown in Fig. 3 were regrouped to create a raster map consisting of four classes: impervious, other upland, open water, and other wetland. Comparison of the two raster datasets resulted in a pixel level difference product that included the follow classes:

- Change classes:
  - Wetland to impervious surface
  - Upland to open water
  - Drier-end vegetated wetland to open water
- Static classes
  - Upland (upland remaining upland)
  - Wetland (wetland remaining wetland, including deepwater)
  - Wetland and deep water not included in this analysis

The initial difference product developed above had many spurious differences resulting from misalignment issues between NWI and C-CAP as well as rounding errors arising from data reprojection and vector-raster data conversion. These spurious differences were typically small objects such as individual pixels, small groups of pixels, or slivers located around the edges of wetland regions (Fig. 8). To minimize these spurious differences, a sophisticated filtering procedure was developed. Difference objects removed by this procedure included those that:

- were small ($\leqslant$ 4 pixels) and touched the edges of wetlands or uplands,
- were located at the edges of wetland areas that were slim or had linear shapes, and
- were slim or had linear shapes and were close to riverine wetland polygons.

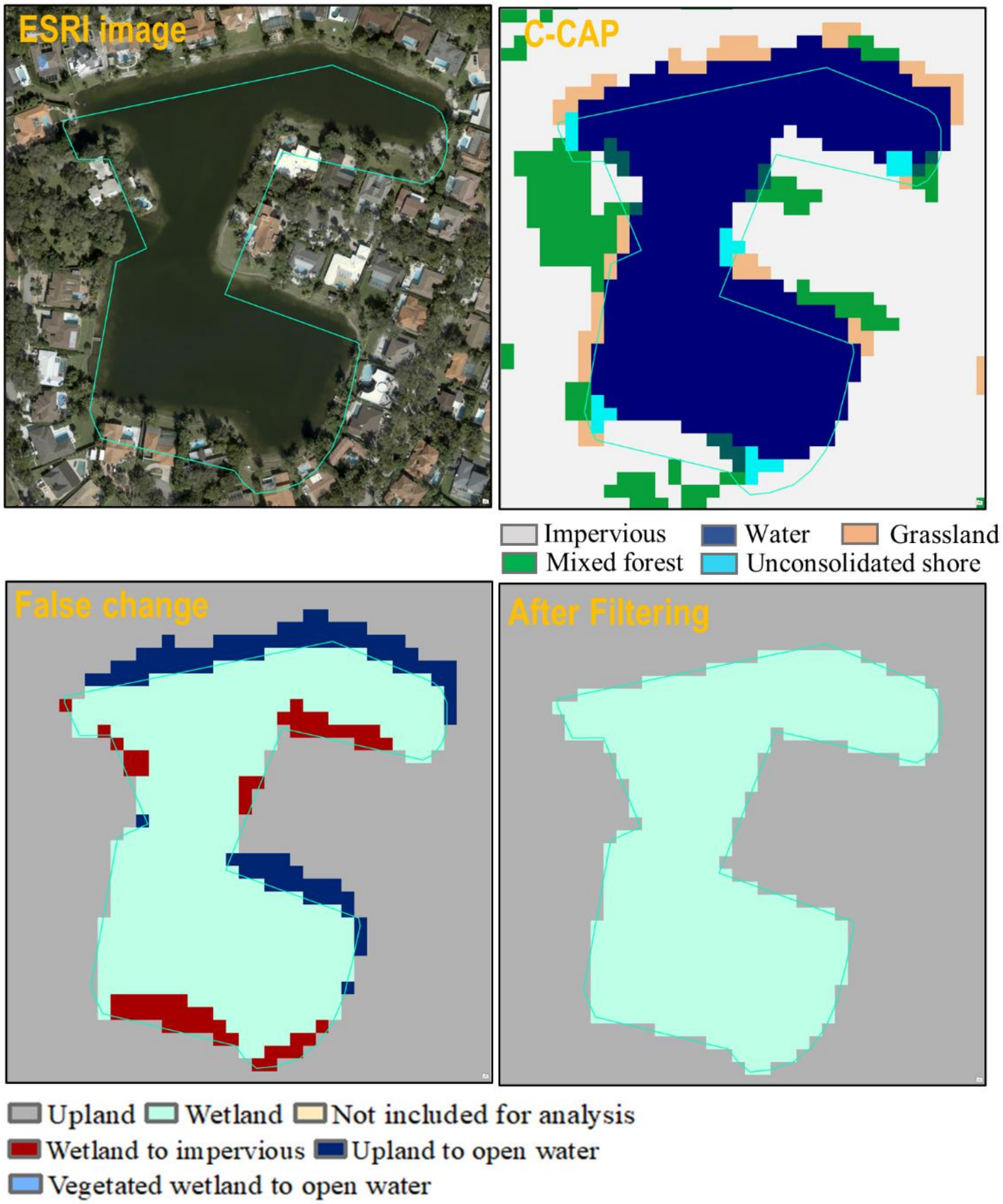

Fig. 8. Misalignment of NWI polygons with high resolution images (upper left) and C-CAP data (upper right) resulted in many spurious differences (lower left), most of which were removed by using an innovative filtering procedure developed during this study.

Fig. 8 shows an example demonstrating the effectiveness of how spurious differences were removed by this innovative filtering process. The final C-CAP-based difference product was produced by applying this filtering procedure to the initial results derived by directly comparing the raster version of the NWI dataset with the C-CAP data. Fig. 9 shows a browse map of the final C-CAP-based pixel level difference product along with two zoom-in examples.

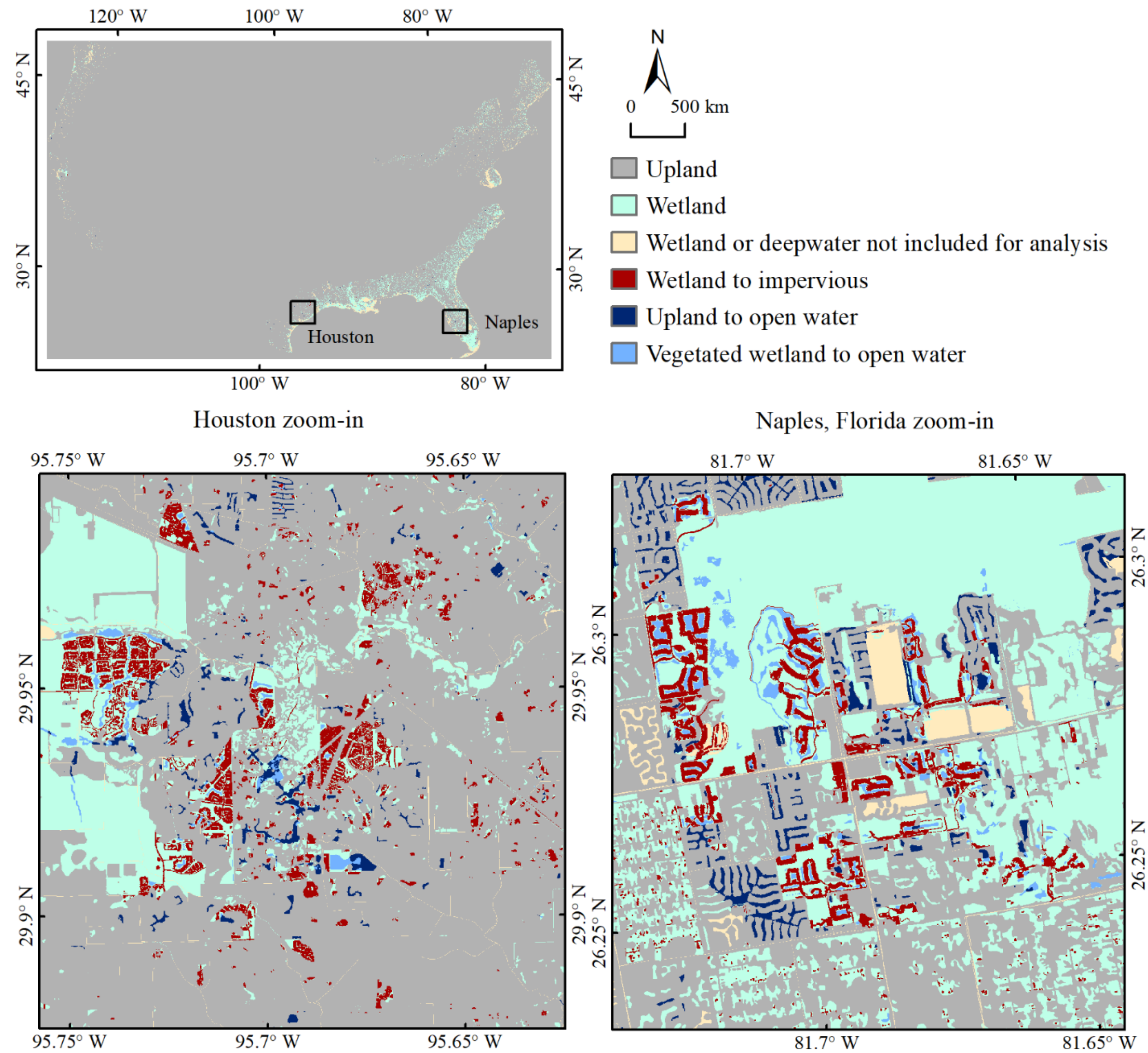

Fig. 9. An overview of the C-CAP-based pixel level difference product and two zoom-in examples.

From the pixel level difference product, we also calculated the amount of change within each HUC12 watershed and census tract polygons. These watershed and census tract level products could be used by NWI or other wetland management agencies to select watersheds that need to be updated the most according to available resources, agency priorities, and/or other practical considerations.

3.1.2 Derivation of the polygon level product

The polygon level product was derived by overlaying the C-CAP land cover dataset on top of the NWI polygons and calculating the amount of impervious areas within each NWI polygon. Because edge pixels are often mixed pixels, impervious pixels located at the edge of impervious patches were considered to have only 50% impervious areas. The rest of C-CAP impervious pixels were considered to have 100% impervious cover. Methods for calculating polygon statistics based on a raster dataset typically require that the polygon dataset be converted to a raster dataset that has the same pixel size and spatial extent as the input raster dataset. Given that NWI polygons were mostly delineated based on imagery that had meter or sub-meter resolutions, many of the fine spatial details in this dataset could be lost if the polygons were rasterized to have the same 10 m pixel size as the original C-CAP data. To minimize errors that may arise from information loss in this conversion process, we resampled the C-CAP land cover data to 1 m resolution and rasterized the NWI dataset to create a dataset having the same resolution and spatial extent. By overlaying the resultant 1 m raster datasets on top of each other, the amount of impervious area within each NWI polygon was calculated at the 1 m resolution as follows:

$$Impervious\ area = \sum_{i=1}^{n} P_i F_i \quad (1)$$

Where Pi is the area of pixel i, Fi is the impervious fraction value of pixel i, and n the total number of 1 m pixels within an NWI polygon.

Fig. 10 shows an overview of the C-CAP-based polygon level product along with two zoom-in examples.

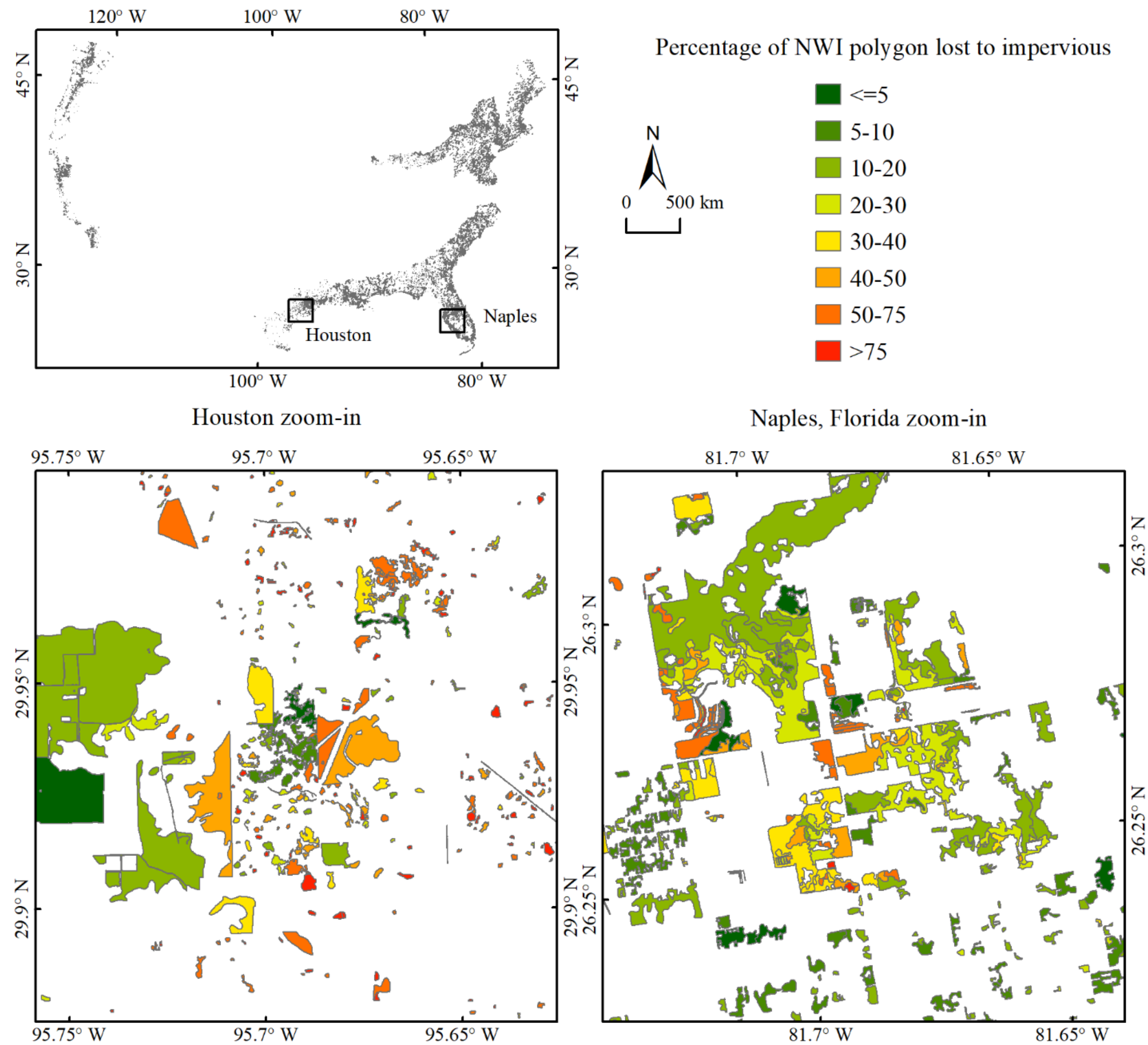

Fig. 10. An overview of the C-CAP-based polygon level difference product (top) along with two zoom-in examples (bottom).

3.2 Product assessment

3.2.1 Assessment of the pixel level product

For the pixel level product, we conducted an accuracy assessment for each of the three change categories, including wetland loss to impervious, upland to open water, and vegetated wetland to open water. Random samples (10 m pixel) were selected for each change category and a no-change category relative to that change category. Table 1 provides a list of the change and no-change categories considered and the number of samples selected for each category. Fig. 11 shows the distribution of the selected samples.

Table 1. Number of random samples (10 m pixels) selected to evaluate the C-CAP based pixel level difference product

Categories	Sample size (number of pixels)
Wetland loss to impervious	200
Wetland remaining wetland	800
Upland to open water	200
Upland remaining upland	1500
Vegetated wetland to open water	200
Vegetated wetland remaining vegetated wetland	800

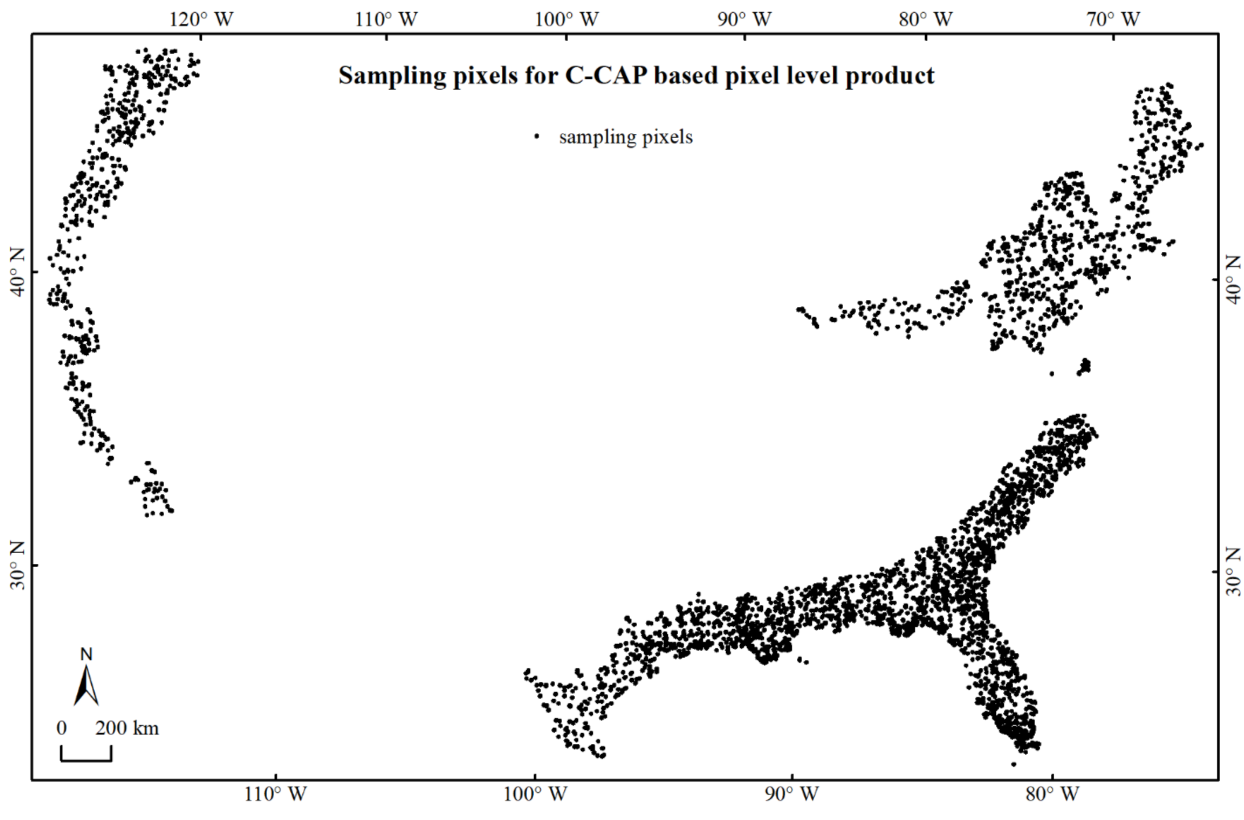

Fig. 11. Distribution of pixels selected for evaluating the C-CAP based pixel level difference product

For each selected 10 m pixel, 25 points evenly distributed within the pixel boundary were generated and overlaid on top of available Google Earth (GE) images (Fig. 12). The high-resolution GE image acquired in the year closest to the year of the C-CAP data over a given sample location was selected as the reference for determining the land cover type over that location. The land cover type at each of the 25 points was determined based on the high resolution reference image. The land cover types considered at the interpretation point level included impervious surface, vegetation, water, urban, bareland, etc. Pixels with at least 5 (or 20%) of the 25 points labeled as impervious surface were classified as impervious, while those having at least 13 (>50%) of the 25 points labeled as water were classified as water.

Fig. 12. The land cover type at a 10 m C-CAP pixel (red square) was determined based on Google Earth high resolution imagery in two steps. First, the percentages of impervious surface, water, and other surface covers within the 10 m pixel were estimated with the assistance of the 25 points evenly distributed within the 10 m pixel. Those percentages were then used to determine the land cover type for the 10 m pixel.

The confusion matrices derived based on the validation points for the three change categories are shown in Tables 2-4. The overall accuracy of wetland loss to impervious, upland to open water, and vegetated wetland to open water were 0.969, 0.980, and 0.960, respectively.

Table 2. Confusion matrix for the wetland loss to impervious change class

		Reference from Google Earth high-resolution image			
		Wetland loss to impervious (fraction >=20%)	Wetland remaining wetland fraction <20%)	Total	User accuracy
Class types from NWI-CCAP difference map	Wetland loss to impervious	169	31	200	0.845
	Wetland remaining wetland	0	800	800	1.000
	Total	169	831	1000	Overall Accuracy
	Producer accuracy	1.000	0.963		0.969

Table 3. Confusion matrix for the upland to open water change class

		Reference from Google Earth high-resolution image			
		Upland to open water (fraction >=50%)	Upland remaining upland (fraction <50%)	Total	User accuracy
Class types from NWI-CCAP difference map	Upland to Open water	166	34	200	0.830
	Upland remaining upland	0	1500	1500	1.000
	Total	180	1520	1700	Overall Accuracy
	Producer accuracy	1.000	0.978		0.980

Table 4. Confusion matrix for the vegetated wetland to open water change class

		Reference from Google Earth high-resolution image			
		Vegetated wetland to open water (fraction >=50%)	Vegetated wetland remaining vegetated wetland (fraction <50%)	Total	User accuracy
Class types from NWI-CCAP difference map	Vegetated wetland to open water	161	39	200	0.805
	Vegetated wetland remaining vegetated wetland	1	799	800	0.999
	Total	162	838	1000	Overall accuracy
	Producer accuracy	0.994	0.953		0.960

3.2.2 Assessment of the polygon level product

In order to assess the accuracy of the impervious estimates within the NWI polygons, we divided the percentages of NWI polygon lost to impervious surface into 20 intervals with a step of 5%. Fifty polygons were randomly selected at each interval (Table 5). Fig. 13 shows the distribution of the selected polygons.

Table 5. Number of samples (NWI polygons) selected to evaluate the C-CAP based polygon level difference product

Percentage of impervious area	Number of NWI polygons selected
0%-5%	50
5%-10%	50
10%-15%	50
15%-20%	50
20%-25%	50
25%-30%	50
30%-35%	50
35%-40%	50
40%-45%	50
45%-50%	50
50%-55%	50
55%-60%	50
60%-65%	50
65%-70%	50
70%-75%	50
75%-80%	50
80%-85%	50
85%-90%	50
90%-95%	50
95%-100%	50

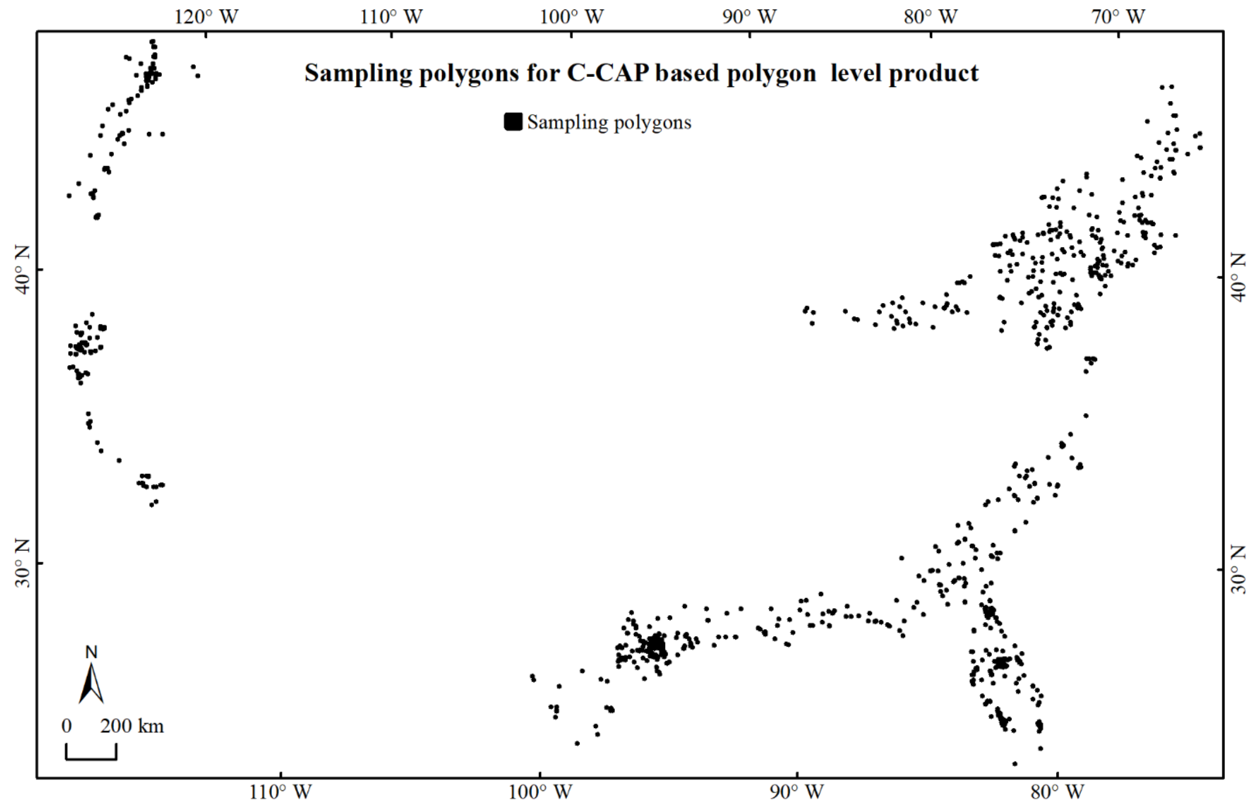

Fig. 13. Distribution of NWI polygons selected for evaluating the C-CAP based polygon level difference product

To determine the impervious area within each selected NWI polygon, we overlaid that polygon on top of the high-resolution GE image acquired in the year closest to the year of the C-CAP data at the polygon location, digitized the area covered by impervious surfaces manually according to the GE image, and calculated the impervious area within that NWI polygon (Fig. 14).

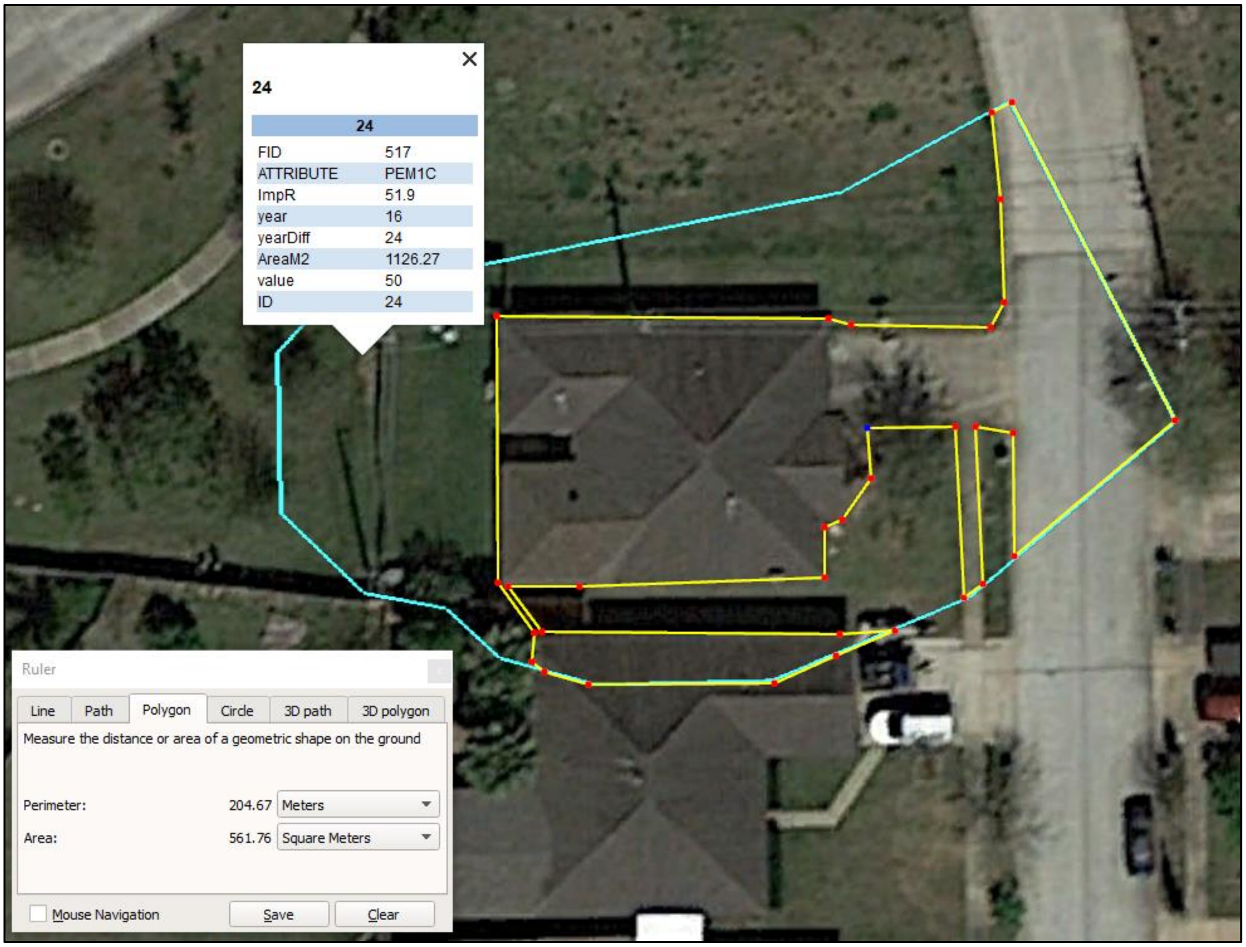

Fig. 14. The impervious area within an NWI polygon (cyan lines) was digitized manually (yellow lines) based on high resolution Google Earth imagery.

Of the 1000 selected NWI polygons, 826 had impervious areas less than 5000 m^2. For these polygons, the coefficient of determination (R^2) of the relationship between impervious area provided by the polygon level difference product and GE-based estimates was 0.79 (Fig. 15). For polygons with impervious area larger than 5000 m^2, the R^2 was 0.97 (Fig. 16), suggesting that the difference product can provide more accurate impervious estimates for polygons with larger impervious values than those with smaller values.

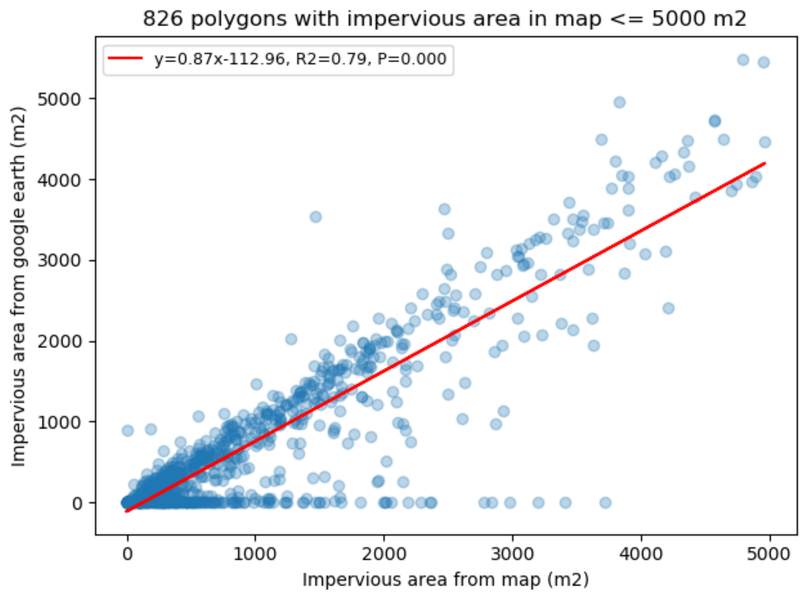

Fig. 15. Relationships between impervious areas provided by the C-CAP-based polygon level difference product and GE-based estimates for NWI polygons having 5000 m^2 or less impervious surfaces.

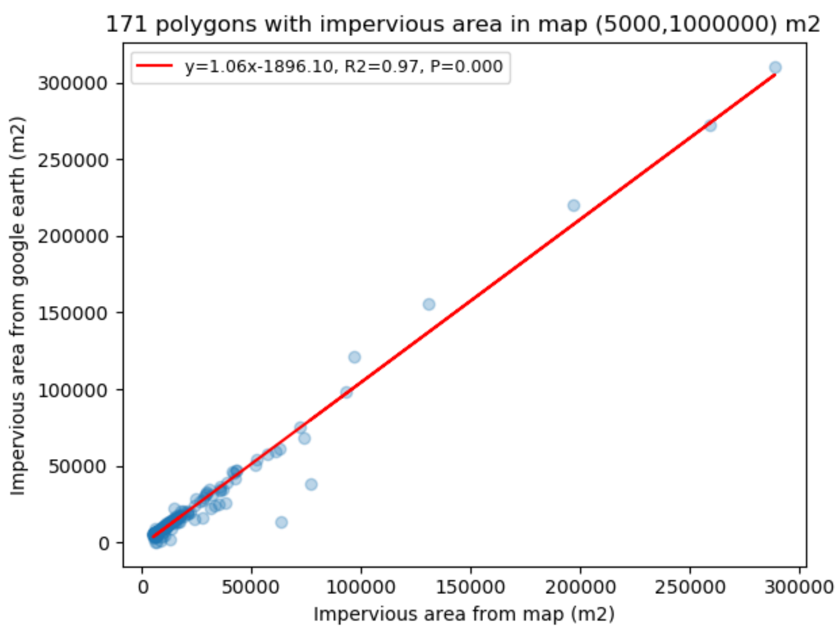

Fig. 16. Relationships between impervious areas provided by the C-CAP-based polygon level difference product and GE-based estimates for NWI polygons having >5000 m^2 impervious surfaces.

4. Development and validation of NLCD-based products

4.1 Methodology

4.1.1 Derivation of the pixel level product

The NLCD-based difference products were derived following the same general approach used to derive the C-CAP-based difference products. However, there were a few differences that required substantial adjustments to the specific methods used. For example, the NLCD land cover map had four urban classes, including open space and low, medium, and high intensity developed. These classes were defined based on impervious values provided in a separate impervious data layer. The impervious values were kept in the final NLCD-based pixel level difference product, and were used to calculate the impervious area within each NWI polygon. We did not include the developed open space in this study because it included non-impervious areas like parks, golf courses, and urban vegetation.

An extensive examination of the NLCD dataset revealed that it had two error types that could result in substantial errors in the derived difference product. One was that because in any given area the NLCD maps were created based on Landsat images acquired on a few selected dates, some of the water pixels mapped by NLCD could be the result of ephemeral flooding or other short-term surface inundation. Since we were primarily concerned with changes to permanent or semi-permanent water, including ephemeral water mapped by NLCD would result in errors in the differences related to changes from upland or vegetated wetland to open water. The other error type in the NLCD data was substantial overestimation of impervious cover over some water areas, especially small water bodies (Fig. 17).

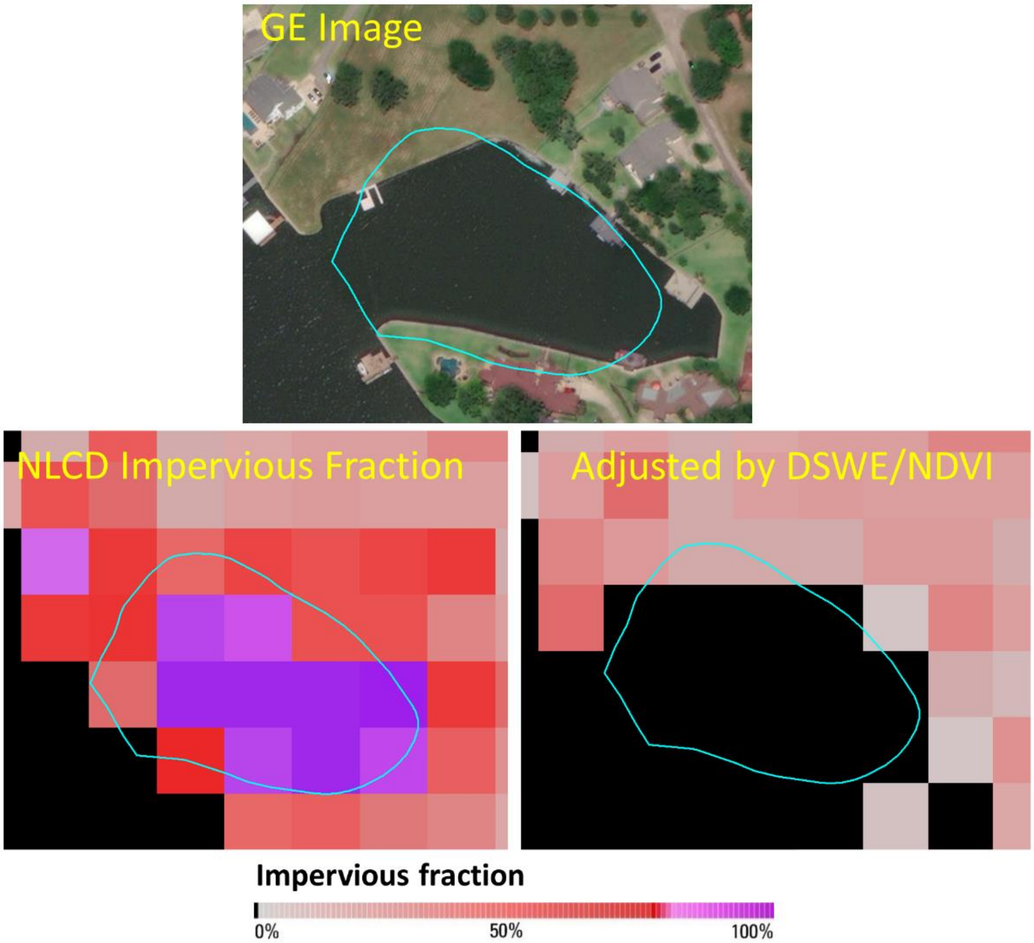

Fig. 17. A Google Earth (GE) image (upper) shows the overestimation of impervious fraction by NLCD over a waterbody (lower left), which was corrected using 3-year inundation (DSWE) and greenness (NDVI) metrics described in section 2 (lower right).

To reduce the impact of these two error types in the NLCD data, we used Google Earth Engine to analyze all Landsat images (>15000) acquired during the growing season over three years (2018-2020). As discussed in section 2, this resulted a suite of temporal metrics describing surface inundation frequency and vegetation phenology for every 30 m pixel across CONUS (Fig. 6 and 7). A surface inundation frequency threshold of 50% was used to remove short term inundation caused by ephemeral flooding in the NLCD data. Many impervious overestimation errors over water and vegetated areas were removed or reduced using both inundation frequency and greenness metrics (Fig. 17).

A major difference between the NLCD-based pixel level product and the C-CAP-based product was that the land cover condition in 2019 (NLCD mapping year) was characterized using a continuous variable, not the binary land cover category used in the C-CAP-based product. Specifically, for the wetland to impervious change type, a continuous number was used to describe the fraction of impervious cover within a 30 m pixel ranging from 20% (lower end of intensity developed) to 100%. For the upland or vegetated wetland to open water change types, the inundation frequency at that pixel location ranging from 50% (lower end of semi-permanently inundated) to 100% was recorded. To represent these continuous variables properly, one pixel level difference product was produced for each of the three change types, including wetland to impervious, upland to open water, and vegetated wetland to open water. An integrated visualization of the three pixel level difference product is provided in Fig.18, including a browse map along with two zoom-in examples.

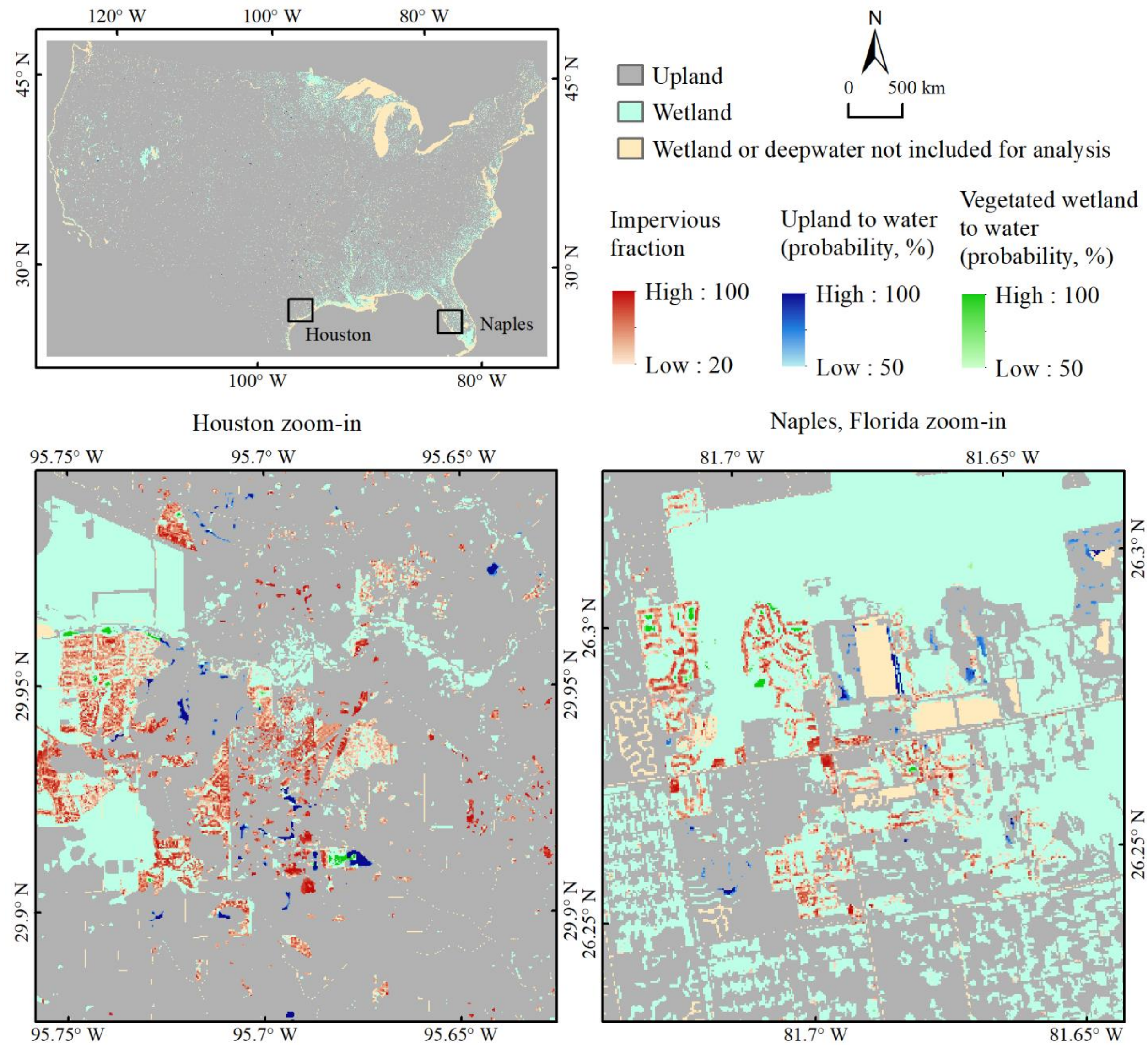

Fig. 18. An integrated visualization of the NLCD-based pixel level difference products – overview (top) and zoomed-in examples (bottom)

4.1.2 Derivation of the polygon level product

Derivation of the NLCD-based polygon level difference product from the adjusted NLCD impervious fraction layer followed essentially the same procedure developed in section 3.1.2 for creating the C-CAP-based polygon level difference product. The only difference was that for the NLCD-based products, the actual impervious value after being adjusted using inundation frequency and greenness metrics for pixels located within an NWI polygon were used to calculate the impervious area within that polygon. Fig. 19 shows an overview of the NLCD-based polygon level product along with two zoom-in examples.

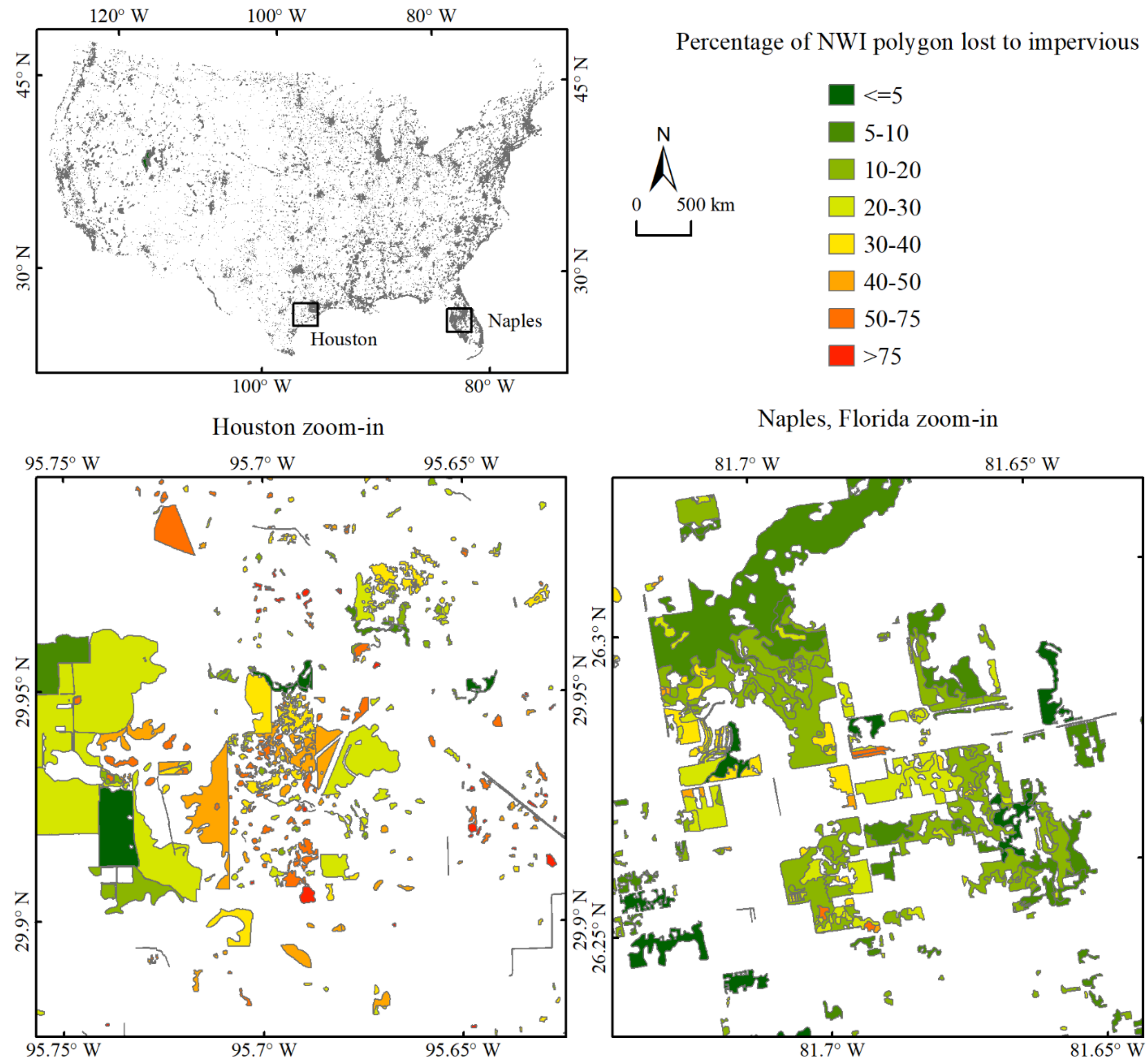

Fig. 19. An overview of the NLCD-based polygon level difference product (top) along with two zoom-in examples (bottom).

4.2 Product assessment

4.2.1 Assessment of the pixel level product

As with the assessment of the C-CAP-based pixel level difference product, we conducted accuracy assessment of the NLCD-based pixel level product for each of the three change categories, including wetland loss to impervious, upland to open water, and vegetated wetland to open water. Random samples (30 m pixels) were selected for each change category and a no-change category relative to that change category. Table 6 provides a list of the change and no-change categories considered and the number of samples selected for each category. Fig. 20 shows the distribution of the selected samples. The same visual analysis procedure used to determine the land cover type of C-CAP pixels as shown in Fig. 12 was used to determine the land cover type of the selected 30 m samples.

Table 6. Number of random samples (30 m pixels) selected to evaluate the NLCD-based pixel level difference product

Categories	Sample size (number of pixels)
Wetland loss to impervious	250
Wetland remaining wetland	250
Upland to open water	250
Upland remaining upland	250
Vegetated wetland to open water	250
Vegetated wetland remaining vegetated wetland	250

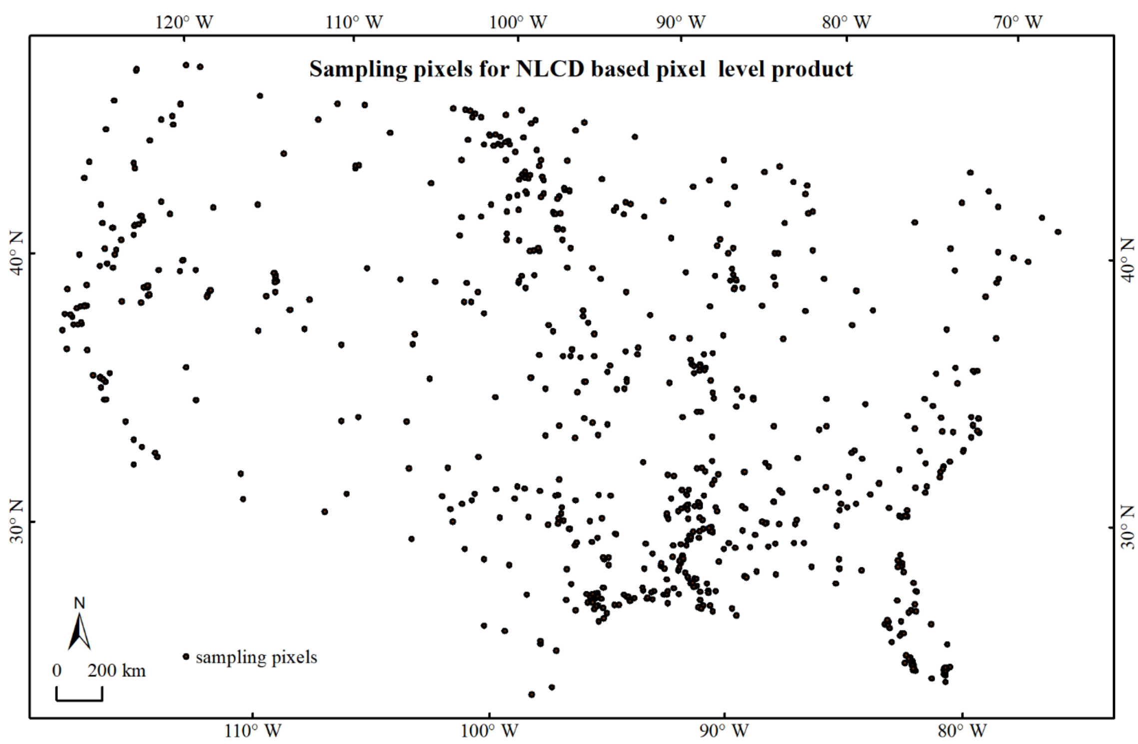

Fig. 20. Distribution of pixels selected for evaluating the NLCD-based pixel level difference product

The confusion matrices derived based on the validation points for the three change categories are shown in Tables 7-9. The overall accuracy of wetland loss to impervious, upland to open water, and vegetated wetland to open water were 0.902, 0.838, and 0.866, respectively.

Table 7. Confusion matrix of wetland loss to impervious

		Class types from google earth high-resolution image			
		Wetland loss to impervious (fraction >=20%)	Wetland without change (fraction <20%)	Total	User accuracy
Class types from NWI-NLCD difference map	Wetland loss to impervious	169	31	200	0.845
	Wetland not changed to impervious	18	282	300	0.940
	Total	187	313	500	Overall Accuracy
	Producer accuracy	0.904	0.901		0.902

Table 8. Confusion matrix of upland to open water

		Class types from google earth high-resolution image			
		Upland to open water (fraction >=50%)	Upland without change (fraction <50%)	Total	User accuracy
Class types from NWI-NLCD difference map	Upland to Open water	100	25	125	0.800
	Upland without change	56	319	375	0.851
	Total	156	344	500	Overall Accuracy
	Producer accuracy	0.641	0.927		0.838

Table 9. Confusion matrix of vegetated wetland to open water

		Class types from google earth high-resolution image			
		Vegetated wetland to open water	Vegetated wetland without change	Total	User accuracy
		(fraction >=50%)	(fraction <50%)		
Class types from NWI-NLCD difference map	Vegetated wetland to open water	94	31	125	0.752
	Vegetated wetland without change	36	339	375	0.904
	Total	130	370	500	Overall accuracy
	Producer accuracy	0.723	0.916		0.866

4.2.2 Assessment of the polygon level product

Similar to the assessment of the C-CAP-based polygon level product described in section 3.2.2, we divided the percentages of impervious area within NWI polygons in the NLCD-based polygon level product into 20 intervals using a step of 5%. Fifty polygons were randomly selected at each interval (Table 10). Since the NLCD dataset was developed with a minimum mapping unit of 4 pixels (~4000m^2), we excluded NWI polygons smaller than 4000m^2 from the assessment. Out of the 1000 polygons selected, 350 were larger than 4000m^2. Fig. 21 shows the distribution of the 350 polygons used to evaluate the NLCD-based polygon level product.

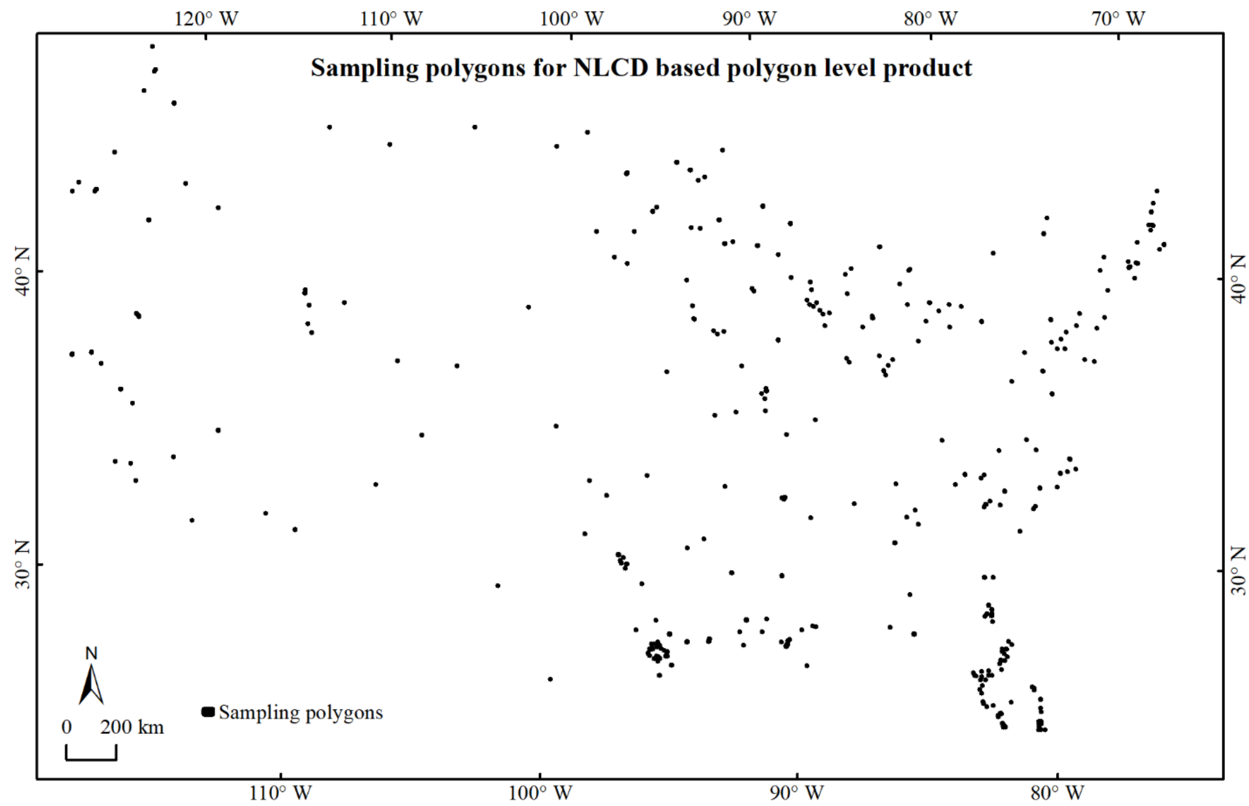

Fig. 21. Distribution of NWI polygons selected for evaluating the NLCD-based polygon level difference product

For each selected polygon, we manually digitized the impervious areas within that polygon as shown in Fig. 14 based on the available GE image with an acquisition year closest to 2019. The impervious areas within the 350 NWI polygons larger than 4000m^2 as estimated by the difference product were correlated with the GE-based estimates with an R^2 value of 0.92 (Fig. 22).

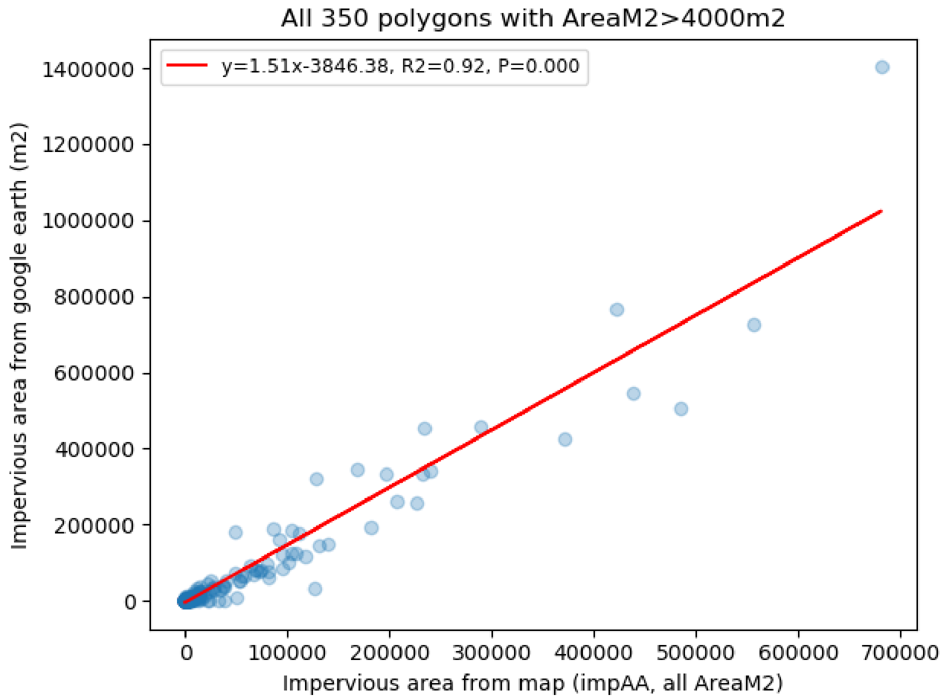

Fig. 22. Relationships between impervious areas provided by the NLCD-based polygon level difference product and GE-based estimates for NWI polygons larger than 4000 m^2.

5. Example Use Cases and Caveats

The pixel level products could be used to identify wetland difference hotspots at specific pixel locations where wetland-related land cover changes might have occurred or there are other dataset differences that could be addressed. The NLCD-based product provides additional information on the amount of impervious area or inundation frequency within each pixel.

The polygon level products could be used to find NWI polygons where wetland loss to urban development is likely to have occurred. The impervious values (area/cover) of a polygon indicate the relative amount of wetland loss within that polygon. Use of both pixel and polygon level products together could help identify the polygons that are likely to have had wetland loss to impervious surfaces and where within the polygon changes have occurred.

Compared with the NLCD data, the C-CAP data appear to be more accurate and had a much finer spatial resolution. Therefore, the C-CAP based products should be prioritized where both products are available.

Due to some mapping errors in the NLCD/C-CAP datasets and misalignment issues between the NWI data and the two land cover datasets, neither the pixel level nor the polygon level products were intended for providing accurate estimates of change areas within individual pixels or NWI polygons. (This partially explains why the products are referred to as difference, and not change products.) Given the relatively large pixel sizes and minimum mapping units of the NLCD and C-CAP datasets as compared to NWI data, differences over small areas are not as reliable as those over larger areas. In particular, difference values in polygons smaller than 4000 m^2 in the

NLCD-based polygon level product or less than 1000 m^2 in the C-CAP-based polygon product are less likely to be a result of wetland change than those in larger polygons. Misalignment issues between the NWI data and the two land cover datasets were caused by multiple factors, including small geolocation errors with the C-CAP data in some regions, geometric inaccuracies with certain NWI polygons, and rounding errors from converting the NWI polygon data to raster data for comparison with the NLCD and C-CAP land cover data, among others. These misalignment issues often resulted in spurious differences along the edges of NWI polygons or impervious areas. The procedure we developed for minimizing these spurious differences could lead to the removal of differences that could indicate real wetland change.

Note that since differences were assessed relative to the date of imagery used to create NWI data, results will not be consistent across NWI project boundaries.

6. Acknowledgement

Funding for this project was provided by the U.S. Fish and Wildlife Service (USFWS). Derivation of the difference products would not have been possible without the guidance and insights from the USFWS, NOAA and USGS. We especially wish to thank Nate Herold and George Xian for their guidance on the use of the C-CAP and NLCD data. The multi-year water (DSWE) and greenness (NDVI) metrics used in this study were generated using the Google Earth Engine (GEE).

Appendix A. Readme file for the NWI Wetland Difference Products

A1. Products description

Two types of difference products were derived by comparing NWI with national land cover products – one is a raster dataset and the other a vector dataset. The raster product was intended to provide pixel level change indicators, while the vector dataset shows the amount of impervious area, and presumably wetland loss, within each NWI polygon.

A1.1 C-CAP-based pixel level difference product

This and the C-CAP-based polygon level product described in section 2.2 were derived by comparing NWI with the C-CAP dataset over the coastal regions of the conterminous United States (CONUS). A small fraction of those regions was excluded from this analysis because the NWI data were more recent than the C-CAP data in those areas. This dataset has a 10 m spatial resolution, and each pixel has an integer value indicating the land cover type/change status at that pixel location. An explanation of the pixel values is provided in Table A1. Fig. A1 shows an overview of this product along with two zoomed-in examples.

Table A1: Class values and names for the C-CAP-based pixel level difference product

Value	Class name/description
0	Upland
1	Wetter Wetland (does not include less frequently inundated vegetated wetlands [see class 6]) without change
2	Marine (not included for change analysis)
3	Estuarine Subtidal (not included for change analysis)
4	Lacustrine Limnetic (not included for change analysis)
5	Riverine (not included for change analysis)
6	Drier Vegetated Wetland (does not include more frequently inundated wetlands [see class 1]) without change
52	Wetland (both class 1 and class 6) to C-CAP impervious
121	Upland to C-CAP open water
221	Wetland (only class 6) to C-CAP open water
255	Background

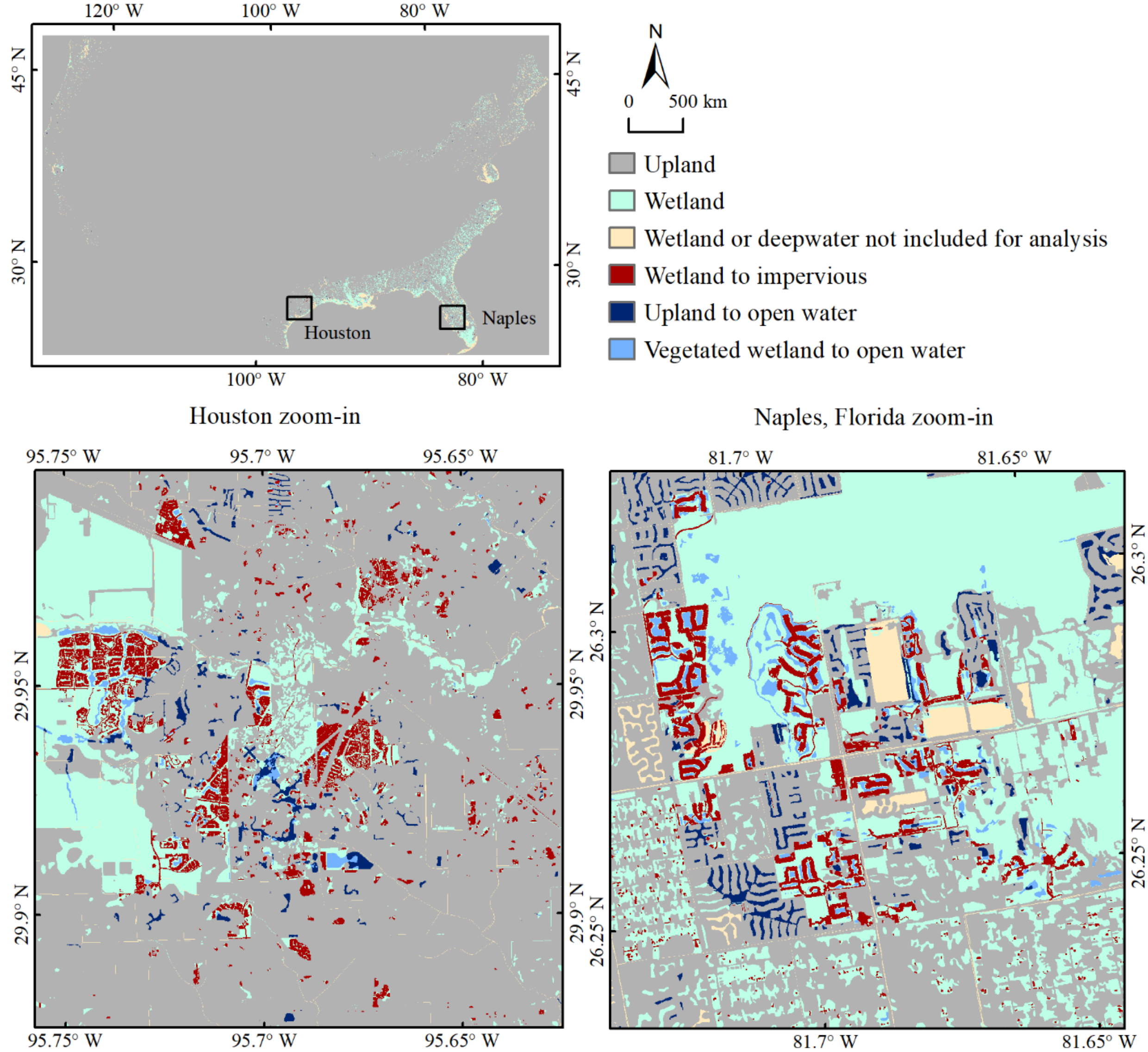

Fig. A1. An overview of the C-CAP-based pixel level difference product (top) along with two zoom-in examples (bottom).

A1.2 C-CAP-based polygon level difference product

This product provides information on the amount of wetland lost to impervious for individual NWI wetland polygons as calculated based on the C-CAP data.

The attributes of each polygon are listed in Table A2. Fig. A2 shows an overview of this product along with two zoomed-in examples.

Table A2: Attribute table of the C-CAP-based polygon level difference product

Attribute name	Description
ATTRIBUTE	Cowardin code of the wetland polygon
WETLAND_TY	NWI wetland description
ACRES	Polygon area in acres
SHAPE_Leng	Polygon perimeter in meters
SHAPE_Area	Polygon area in m^2
ImpAcre	Impervious area (acres) in a wetland polygon
ImpP	Percentage of impervious cover in the polygon (100 x ImpAcre/ACRES)

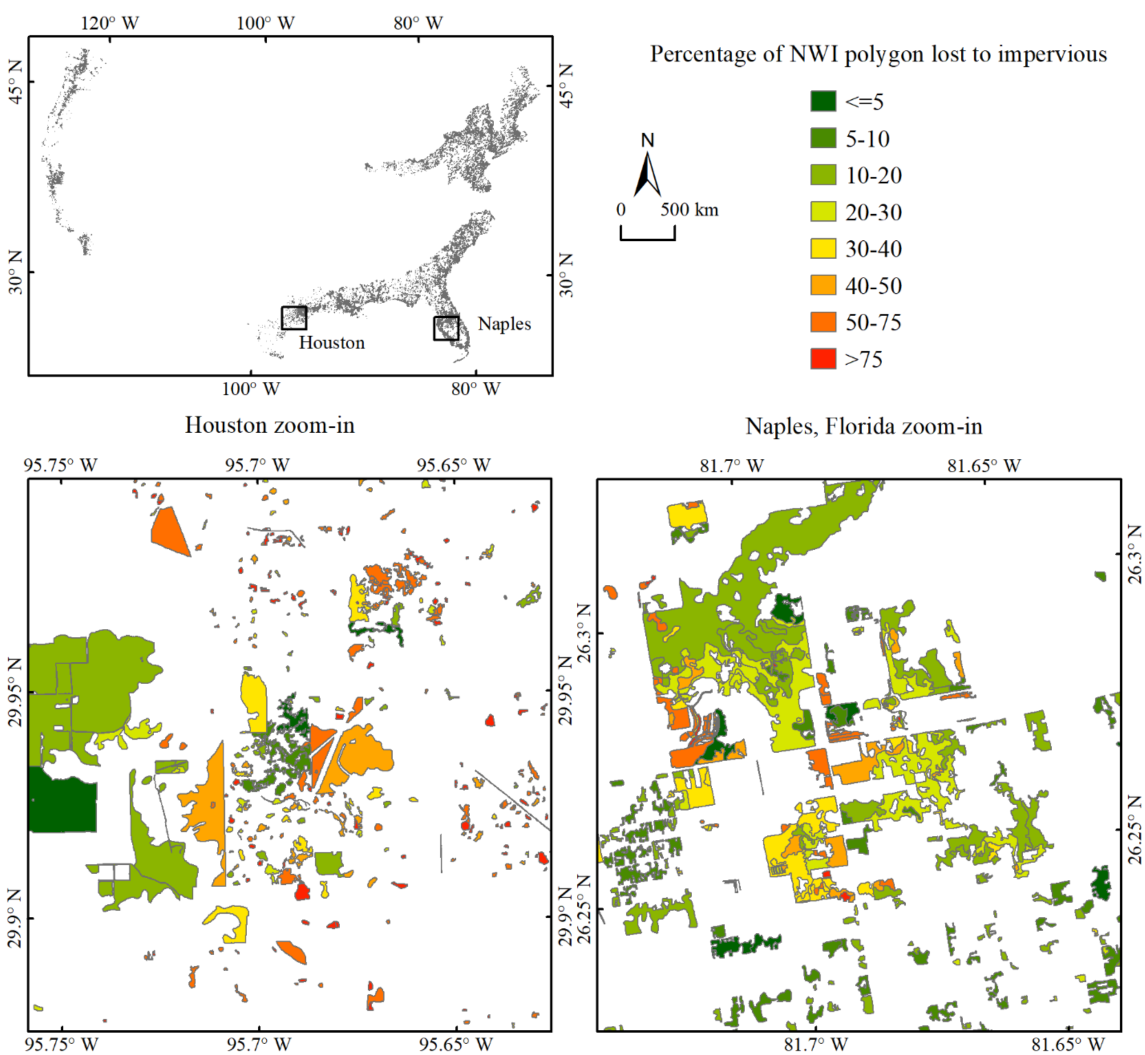

Fig. A2. An overview of the C-CAP-based polygon level difference product (top) along with two zoom-in examples (bottom). Only polygons that had a wetland loss of > 10 acres or > 10% of the total area of each polygon are shown here.

A1.3 NLCD-based pixel level difference product

This and the NLCD-based polygon level product described in section 2.4 were derived by comparing NWI with the NLCD 2019 dataset over CONUS. The NLCD dataset provides information on fractional impervious cover within each 30 m pixel. Further, the frequency of water presence at each pixel location was calculated based on an analysis of all Landsat images acquired during the growing season (May – September) of 2018-2020. In order to preserve these thematic details, one raster file was produced for each of the three change types listed in section 1. Descriptions of the pixel values in these files are provided in Tables A3-A5. To support visualization of these products in the context of existing wetlands, a raster version of the original NWI dataset was also produced. A description of the pixel values in this file is provided in Table A6. An integrated visualization of these products is shown in Fig. A3 along with two zoomed-in examples.

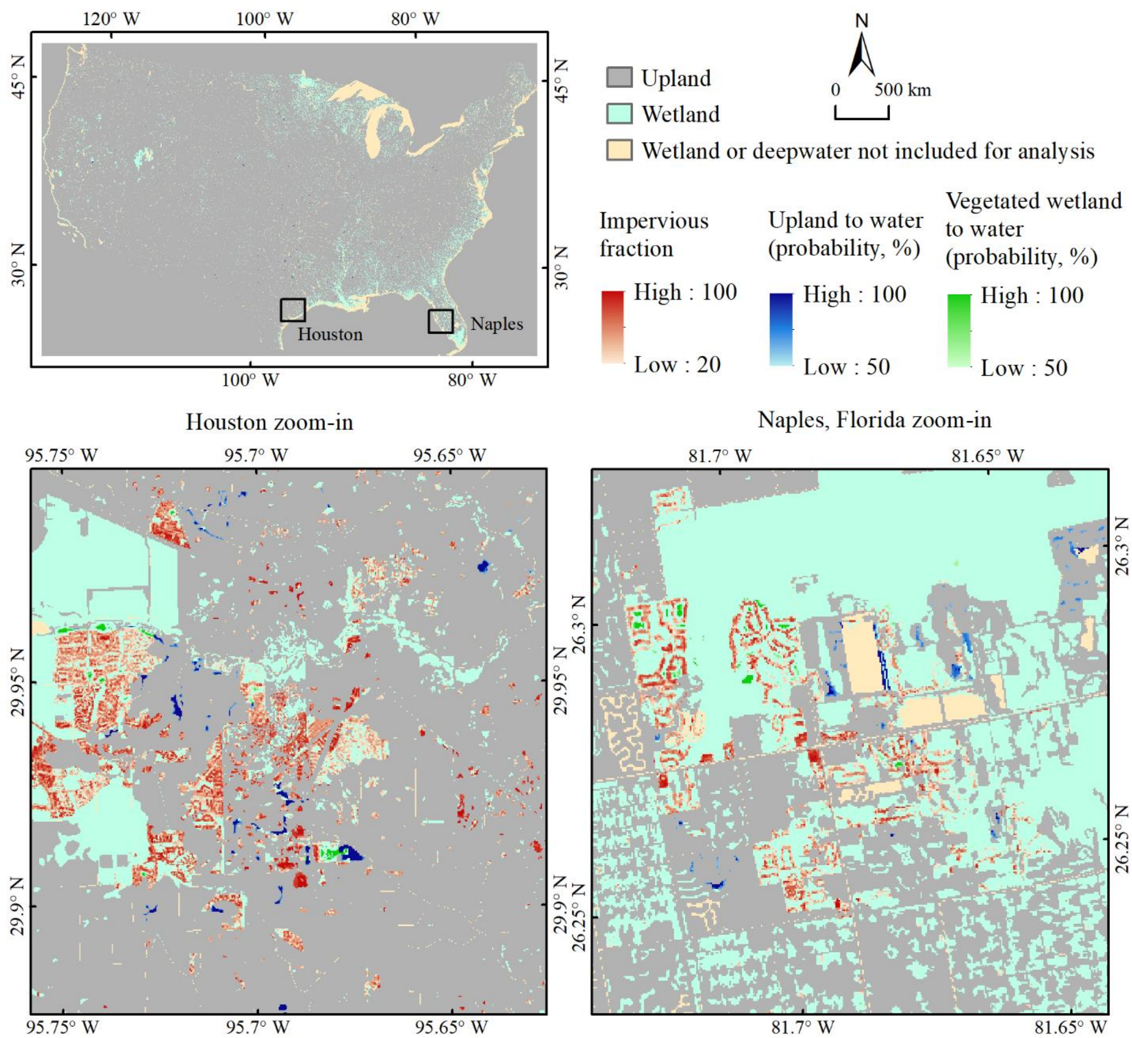

Fig. A3. An integrated visualization of the NLCD-based pixel level difference products – overview (top) and zoomed-in examples (bottom)

Table A3: Values in the NLCD-based pixel level difference product (wetland to impervious)

Value	Description
20-100	Wetland loss to impervious surfaces - impervious fraction in those pixels
255	All other pixels

Table A4: Values in the NLCD-based pixel level difference product (upland to open water)

Value	Description
50-100	Upland to open water – inundation frequency in those pixels
255	All other pixels

Table A5: Values in the NLCD-based pixel level difference product (vegetated wetland to open water)

Value	Description
50-100	Vegetated wetland to open water – inundation frequency in those pixels
255	All other pixels

Table A6: Values in the raster version of the NWI dataset

Value	Meaning
0	Upland
1	Wetter Wetland (does not include less frequently inundated vegetated wetlands [see class 6]) without change
2	Marine (not included for change analysis)
3	Estuarine Subtidal (not included for change analysis)
4	Lacustrine Limnetic (not included for change analysis)
5	Riverine (not included for change analysis)
6	Drier Vegetated Wetland (does not include more frequently inundated wetlands [see class 1]) without change
10	Upland or Ocean Water in coastal regions
255	Background

A1.4 NLCD-based polygon level difference product

This product provides information on the amount of impervious areas in individual NWI wetland polygons. The attributes of each polygon are listed in Table A7. Fig. A4 shows an overview of this product along with two zoomed-in examples.

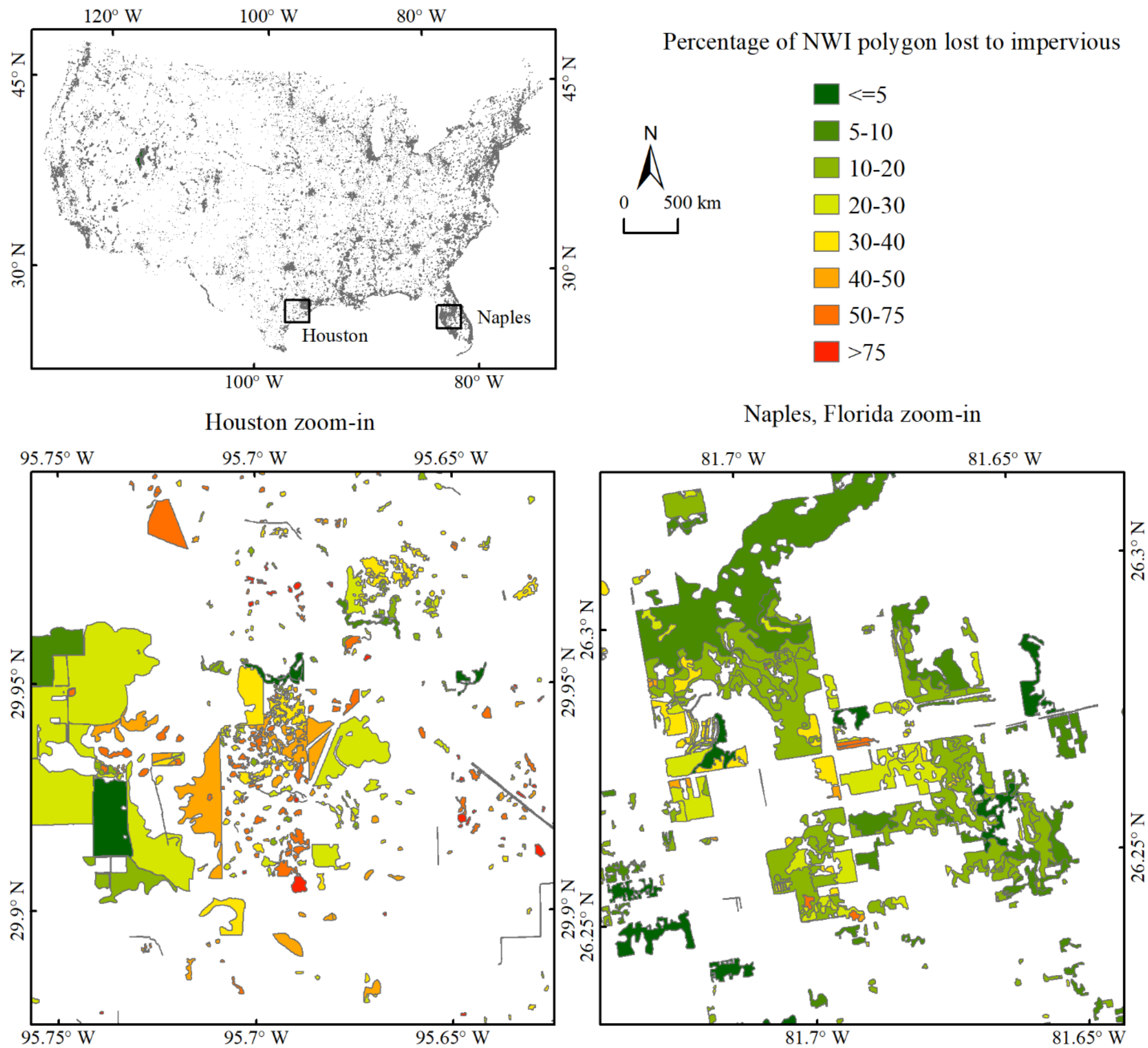

Fig. A4. An overview of the NLCD-based polygon level difference product (top) along with two zoom-in examples (bottom). Only polygons that had a wetland loss of > 10 acres or > 10% of the total area of each polygon are shown here.

Table A7: Attribute table of the NLCD-based polygon level difference product

Attribute name	Description
ATTRIBUTE	Cowardin code of the wetland polygon
WETLAND_TY	NWI wetland description
ACRES	Polygon area in acres
SHAPE_Leng	Polygon perimeter in meters
SHAPE_Area	Polygon area in m^2
ImpAcre	Impervious area (acres) in a wetland polygon
ImpP	Percentage of impervious cover in the polygon (100 x ImpAcre/ACRES)

A2. Data format

- Pixel level products: Erdas Imagine format (*.img)
- Polygon level products <=4 GB: Shapefile format (*.shp)
- Polygon level products >4 GB: GeoPackage format (*.gpkg)

A3. Geospatial reference information

All products have the same projection:

NAD_1983_Contiguous_USA_Albers
WKID: 5070 Authority: EPSG
Projection: Albers
False_Easting: 0.0
False_Northing: 0.0
Central_Meridian: -96.0
Standard_Parallel_1: 29.5
Standard_Parallel_2: 45.5
Latitude_Of_Origin: 23.0
Linear Unit: Meter (1.0)
Geographic Coordinate System: GCS_North_American_1983
Angular Unit: Degree (0.0174532925199433)
Prime Meridian: Greenwich (0.0)
Datum: D_North_American_1983
Spheroid: GRS_1980
Semimajor Axis: 6378137.0
Semiminor Axis: 6356752.314140356
Inverse Flattening: 298.257222101q

A4. Data access

The datasets are organized by state and are provided in the following box folder:

https://umd.box.com/s/fblc1vagvr18iyw2jq3xga5uvkwmafq9

List of files for each state (using AL as an example):

AL (state of Alabama):

NWI_C-CAP:

NWI_C-CAP_10m_pixel_state_AL.img

NWI_C-CAP_10m_pixel_state_AL.img.aux.xml

(NWI vs C-CAP pixel level difference product)

NWI_C-CAP_polygon_state_AL.shp

NWI_C-CAP_polygon_state_AL.cpg

NWI_C-CAP_polygon_state_AL.dbf

NWI_C-CAP_polygon_state_AL.shx

NWI_C-CAP_polygon_state_AL.prj

(NWI vs C-CAP polygon level difference product, wetland to impervious)

NWI_NLCD differences:

NWI_NLCD_30m_pixel_nwi_to_imp_state_AL.img

NWI_NLCD_30m_pixel_nwi_to_imp_state_AL.img.aux.xml

(NWI vs NLCD pixel level difference product, wetland to impervious)

NWI_NLCD_30m_pixel_upland_to_water_state_AL.img

NWI_NLCD_30m_pixel_upland_to_water_state_AL.img.aux.xml

(NWI vs NLCD pixel level difference product, upland to water)

NWI_NLCD_30m_pixel_vegWetland_to_water_state_AL.img

NWI_NLCD_30m_pixel_vegWetland_to_water_state_AL.img.aux.xml

(NWI vs NLCD pixel level difference product, vegetated wetland to water)

NWI_30m_pixel_state_AL.img

NWI_30m_pixel_state_AL.img.aux.xml

(NWI data in 30-m raster format)

NWI_NLCD_polygon_state_AL.cpg

NWI_NLCD_polygon_state_AL.dbf

NWI_NLCD_polygon_state_AL.prj

NWI_NLCD_polygon_state_AL.shp

NWI_NLCD_polygon_state_AL.shx

(NWI vs NLCD polygon level difference product, wetland to impervious)

For users who need to get the products for the entire CONUS, CONUS-wide mosaics are available under the “conus” subfolder. Due to constraints of file size limit/display speed of vector data, only polygons with significant loss to impervious were included in the mosaics of the polygon level products.

List of CONUS-wide Datasets:

NWI_C-CAP_10m_pixel_conus.zip

(NWI-C-CAP pixel level difference product)

NWI_C-CAP_only_nwi_polygons_with_significant_lost_to_impervious_selected.zip

(NWI vs C-CAP polygon level difference product)

NWI_NLCD_30m_pixel_nwi_to_imp_conus.zip

(NWI vs NLCD pixel level difference product, nwi to impervious)

NWI_NLCD_30m_pixel_upland_to_water_conus.zip

(NWI vs NLCD pixel level difference product, upland to water)

NWI_NLCD_30m_pixel_vegWetland_to_water_conus.zip

(NWI vs NLCD pixel level difference product, vegetated wetland to water)

NWI_raster_30m_pixel_conus.zip

(NWI data in 30-m raster format)

NWI_NLCD_only_nwi_polygons_with_significant_loss_to_impervious_selected.zip

(NWI vs NLCD polygon level difference product)

A5. Contact information

Zhenhua Zou, zhzou@umd.edu
Chengquan Huang, cqhuang@umd.edu
Megan_Lang@fws.gov

Appendix B. Land Cover Classes, Code and Definitions

Table B1. Land cover classes of the 2019 NLCD land cover dataset

Value	Meaning
11	Open Water- areas of open water, generally with less than 25% cover of vegetation or soil.

12	Perennial Ice/Snow- areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover.
21	Developed, Open Space- areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20% of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.
22	Developed, Low Intensity- areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20% to 49% percent of total cover. These areas most commonly include single-family housing units.
23	Developed, Medium Intensity -areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50% to 79% of the total cover. These areas most commonly include single-family housing units.
24	Developed High Intensity-highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80% to 100% of the total cover.
31	Barren Land (Rock/Sand/Clay) - areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of total cover.
41	Deciduous Forest- areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75% of the tree species shed foliage simultaneously in response to seasonal change.
42	Evergreen Forest- areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75% of the tree species maintain their leaves all year. Canopy is never without green foliage.
43	Mixed Forest- areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. Neither deciduous nor evergreen species are greater than 75% of total tree cover.
51	Dwarf Scrub- Alaska only areas dominated by shrubs less than 20 centimeters tall with shrub canopy typically greater than 20% of total vegetation. This type is often co-associated with grasses, sedges, herbs, and non-vascular vegetation.
52	Shrub/Scrub- areas dominated by shrubs; less than 5 meters tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions.
71	Grassland/Herbaceous- areas dominated by gramanoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing.
72	Sedge/Herbaceous- Alaska only areas dominated by sedges and forbs, generally greater than 80% of total vegetation. This type can occur with significant other grasses or other grass like plants, and includes sedge tundra, and sedge tussock tundra.
73	Lichens- Alaska only areas dominated by fruticose or foliose lichens generally greater than 80% of total vegetation.

74	Moss- Alaska only areas dominated by mosses, generally greater than 80% of total vegetation.
81	Pasture/Hay-areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20% of total vegetation.
82	Cultivated Crops -areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20% of total vegetation. This class also includes all land being actively tilled.
90	Woody Wetlands- areas where forest or shrubland vegetation accounts for greater than 20% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.
95	Emergent Herbaceous Wetlands- Areas where perennial herbaceous vegetation accounts for greater than 80% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.

Table B2. Land cover classes of the C-CAP 10m dataset

Value	**Meaning**
0	Background
1	Unclassified
2	Impervious Developed
5	Open Space Developed
8	Grassland
11	Upland Trees
12	Scrub/Shrub
13	Palustrine Forested Wetland
14	Palustrine Scrub/Shrub Wetland
15	Palustrine Emergent Wetland
16	Estuarine Forested Wetland
17	Estuarine Scrub/Shrub Wetland
18	Estuarine Emergent Wetland
19	Unconsolidated Shore
20	Barren Land
21	Open Water
22	Palustrine Aquatic Bed
23	Estuarine Aquatic Bed

Table B3. Water regime codes and their full names (Cowardin, Carter et al. 1979)

Nontidal		**Saltwater Tidal**		**Freshwater Tidal**	
Code	Name	Code	Name	Code	Name
A	Temporarily Flooded	L	Subtidal	Q	Regularly Flooded-Fresh Tidal
B	Seasonally Saturated	M	Irregularly Exposed	R	Seasonally Flooded-Fresh Tidal
C	Seasonally Flooded	N	Regularly Flooded	S	Temporarily Flooded-Fresh Tidal
D	Continuously Saturated	P	Irregularly Flooded	T	Semipermanently Flooded-Fresh Tidal
E	Seasonally Flooded/Saturated			V	Permanently Flooded-Fresh Tidal
F	Semipermanently Flooded				
G	Intermittently Exposed				
H	Permanently Flooded				
J	Intermittently Flooded				
K	Artificially Flooded				

Reference

Cowardin, L. M., V. Carter, F. C. Golet and E. T. LaRoe (1979). Classification of wetlands and deepwater habitats of the United States. Washington, DC, USA, US Department of the Interior, US Fish and Wildlife Service.

Jones, J. (2015). "Efficient Wetland Surface Water Detection and Monitoring via Landsat: Comparison with in situ Data from the Everglades Depth Estimation Network." Remote Sensing **7**(9): 12503.

Jones, J. W. (2019). "Improved Automated Detection of Subpixel-Scale Inundation—Revised Dynamic Surface Water Extent (DSWE) Partial Surface Water Tests." Remote Sensing **11**(4): 374.