Monitoring & Assessment of Climate Change Impact on Geomorphology in the Coastal Areas of Bangladesh

Deliverable 3.1, 4.1: Step-by-step guide with methodological guidelines

DHI • Agern Allé 5 • • DK-2970 Hørsholm • Denmark
Telephone: +45 4516 9200 • Telefax: +45 4516 9292 • dhi@dhigroup.com • www.dhigroup.com

Monitoring & Assessment of Climate Change Impact on Geomorphology in the Coastal Areas of Bangladesh

Deliverable 3.1, 4.1: Step-by-step guide with methodological guidelines

Prepared for CTCN

Represented by Ho-Sik Chon

Satellite image of the Sundarbans

Project number	11818880
Date	2020-07-13
Revision	0.1 draft
Classification	Restricted

CONTENTS

1 Introduction

The coastline of Bangladesh is at the forefront of climate change and coastal communities are becoming increasingly more vulnerable to climate impacts. Satellite-derived data and information can support, and significantly enhance, the ability of authorities to mitigate climate induced impacts. This manual introduces basic guidelines and concepts of remote sensing and step-by-step guidelines for using satellite data to assess coastal changes and river morphological change.

1.1 Background

Climate change impact in Bangladesh

The Intergovernmental Panel on Climate Change (IPCC) projected in 2007 that the world's oceans would rise from 20 to 60 cm in the coming century in response to global warming. This estimate, however, was modified in a 2019 IPCC special report[1] where IPCC estimated that if current emission level continue global sea level rise could rise up to 110 cm by the end of the 21st century.

Bangladesh is, according to the United Nations, one of the few countries to experience potentially catastrophic consequences of climatic change. Bangladesh is already severely affected by tropical cyclones and storm surges. Catastrophic events in the past have caused damage up to 100 km inland. Imaging how devastating these catastrophes would be with accelerated sea-level rise. Thus decision-makers and practitioners will need more than ever to incorporate sea level rise and other climate change considerations into policies and coastal management plans. Still, appropriate and timely measures can only be taken on the basis on valid scientific information and for the time being the net impact of sea level rise on the land area of Bangladesh is still surrounded by much controversy.

The most obvious outcome of sea level rise due to storms, tsunami or global sea level change is the permanent inundation of Bangladeshi coastal areas which can have serious impacts upon the natural environment and socio-economic conditions in the coastal zone. Over time, inundation changes the position of the coastline and drowns natural habitat and coastal structures. Inundation can also exacerbate coastal erosion by transporting submerged sediments offshore and extending the effects of coastal flooding by allowing storm waves to act further.[2] There is however a number of confounding factors which need to be accounted for before it can be

[1] IPCC, 2019: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. In press.

[2] Dwarakish, G. & Vinay, S. & Natesan, Usha & Asano, Toshiyuki & Kakinuma, Taro & Venkataramana, Katta & Pai, Jagadeesha & Babita, M.. (2009). Coastal vulnerability assessment of the future sea level rise in Udupi coastal zone of Karnataka state, west coast of India. Ocean & Coastal Management - OCEAN COAST MANAGE. 52. 467-478. 10.1016/j.ocecoaman.2009.07.007.

stated that a country like Bangladesh is and actually will loose ground as a result of forecast climatic change and associated rise in sea level. For one thing, vast amounts of sediments from the Himalayas are transported to and deposited in the delta by the rivers and thus creating “new” land. In addition, the relative vertical movement of land due to land subsidence or uplift can either exacerbate or suppress the rise in sea level. Finally, many estimates of the land area that would be lost due to sea level rise are misleading since they do not factor in the embankments that protect much of the coast.[3]

The purpose of this training manual

Satellite remote sensing has long been considered an ideal technology for terrestrial and marine monitoring due to its ability to provide synoptic, repetitive and consistent information of the Earth’s surface. The objective of this manual is to introduce the basic concepts of satellite-based remote sensing with a special emphasize on coastal change mapping and assessment of river morphological dynamics. It will illustrate how different data acquired by various satellite sensors can be used to get scale variable (temporal and spatial) information on coastal dynamics and river morphological changes in Bangladesh.

The manual provides basic introductions to key software/cloud processing systems used to pre-process and process satellite imagery and visualise results. Furthermore, it presents in a non-prescriptive way, step-by step guidelines for using satellite data to; Extract coastlines and river networks; Assess short- and long-term morphological changes; Quantify changes and visualise results; Explorative approaches to assess soil salinity.

[3] Inman, Mason. (2009). Where warming hits hard. Nature Reports Climate Change. 3. 18-21. 10.1038/climate.2009.3.

2 Remote sensing: a brief introductory

In this section the basic concepts and theory of remote sensing is introduced and with a specific emphasis on remote sensing for surface water and vegetation mapping and monitoring

2.1 Basics of remote sensing

Types of remote sensing imagery

Remote sensing refers to the detection of electromagnetic energy reflected from or emitted by the Earth's surface, by air- or space-borne sensors. The electromagnetic radiation used in data acquisition by remote sensing sensors is normally divided into three distinct regions: an optical region (~ 0.4 µm to 3 µm); a thermal infrared region (~ 8 µm to 14 µm) and a microwave region (~ 1 mm to 1 m)[4]. The most common sensors are 'passive' optical sensors i.e. sensors that record electromagnetic radiation emitted from the sun and reflected by the ground surface or emitted by the ground surface itself. A sensor uses photo detectors sensitive to the direct contact of photons on its surface and generates a recordable electric signal proportional to the energy of the photon. This recording forms the remote sensing image. Since the energy decreases as the wavelength increases, thermal sensors need to collect energy from a larger surface (i.e., collecting more photons to get enough energy beyond the detection threshold) than the sensors of shorter wavelengths, so thermal sensors have larger spatial resolutions. Typical applications of thermal sensors are, e.g., surface fire mapping and heat loss monitoring.

A different type of sensor is referred to as 'active'. In contrast to 'passive' sensors that record naturally available radiation, an active sensor emits electromagnetic radiation itself and later measures the energy returned or bounced back from the target to the sensor. LIDAR (LIght Detection And Ranging) is an active sensor system that operates in the optical region, while RADAR (RAdio Detection And Ranging) operates in the microwave region[5]. An important advantage of microwave remote sensing over optical remote sensing is its ability to penetrate clouds, thus enabling data collection at almost any time. The most common active microwave sensors are the Synthetic Aperture Radar (SAR) imaging systems. They can be classified according to the wavelength (frequency) bands and polarization modes used in the data acquisition e.g. single frequency (e.g. X-band ~ 3 cm; C-band ~ 6 cm; L-band ~ 24 cm) or multiple frequency (i.e. a combination of two or more frequency bands) and single polarization (VV[6], HH, HV or VH) or multiple polarization (combination of two or more polarization modes). A special field of SAR application is termed radar

[4] In the ranges of 3 to 8 µm and 14 µm to 1 mm the atmosphere is practically opaque, it does not transmit the electromagnetic energy.

[5] Note that it is also possible to record passive microwave energy that is naturally emitted from the Earth's surface

[6] VV = Emitted with Vertical polarization and received with Vertical polarization; HH = emitted with Horizontal polarization and received with Horizontal polarization, …

interferometry where multiple SAR images of the same location are used to extract three-dimensional information based on the analysis of image interferograms. Common use of radar interferometry includes topographic mapping and assessments of canopy structures.

Remote sensing data is mostly distributed as digital images, consisting of a two-dimensional array of individual picture elements called pixels arranged in columns and rows. Each pixel represents an area on the Earth's surface. The location of a pixel is denoted by its row and column coordinates in the two-dimensional image. There is a one-to-one correspondence between the column-row location of a pixel and the geographical coordinates (e.g. longitude, latitude). In order to be useful, the exact geographical location of each pixel on the ground must be derivable from its row and column indices, given the imaging geometry and the satellite orbit parameters. A pixel represents the intensity level or the radiation that was captured in a given wavelength interval. This value is normally the average intensity for the whole ground area covered by the pixel. The intensity of a pixel is digitized and recorded by the sensor as a digital number.

Radiometric, spectral, spatial and temporal resolution

A digital number is stored with a finite number of bits. The number of bits determines the radiometric resolution of the image. For example, an 8-bit digital number ranges from 0 to 255, while an 11-bit digital number ranges from 0 to 2047, i.e., that many levels of intensities can be recorded for a pixel. The detected intensity value needs to be scaled and quantized to fit within the given range. In a radiometrically calibrated image, the actual intensity value can be derived from the digital number, by the inversion of the above scaling.

The spectral resolution refers to the number and dimensions of specific wavelength intervals (bands) in the electromagnetic spectrum to which a remote sensing detector is sensitive. Objects on the Earth interact with the reflected or emitted electromagnetic energy according to their material and surface characteristics. Furthermore, this interaction is also wavelength dependent, therefore, the spectral bands contain data sufficient for identifying the different objects. Besides using individual spectral bands, spectral indices are also very useful. Spectral indices use combinations of the reflectance bands to accentuate properties of the earth’s surface, which can be used for improved mapping. The standard typology for spectral resolution is to refer to the sensor systems as panchromatic (single band), multi-spectral (several spectral bands) or hyper-spectral (hundreds of spectral bands).

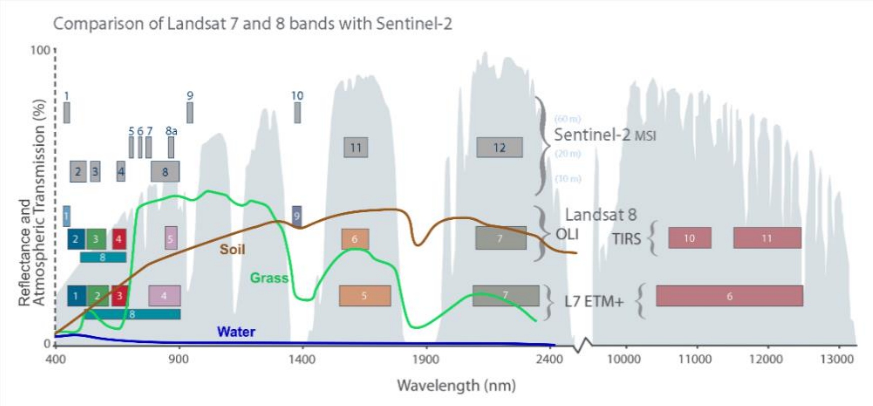

Figure 1: Comparison of Sentinel-2 and Landsat-7 & 8 spectral bands (coloured boxes) and the reflectance curves of soil, grass and water (solid lines). Light grey bars are proportional to the transmission of electromagnetic energy through the atmosphere. Modified from: [*https://landsat.gsfc.nasa.gov/wp-content/uploads/2015/06/Landsat.v.Sentinel-2.png*](https://landsat.gsfc.nasa.gov/wp-content/uploads/2015/06/Landsat.v.Sentinel-2.png)

Spatial resolution refers to the size of the smallest object that can be resolved on the ground. With some simplification, this refers to the size of the area represented by a pixel in the original image. Spatial resolution is often referred to as very high (~ 5 m or less); high (~ 5 m to 30 m); medium/moderate (~ 100 m to 500 m); low/coarse (~ 1 km or more). It is important to note that pixel size and spatial resolution are not interchangeable as it is possible to display an image with a pixel size different from its spatial resolution.

In addition to the radiometric, spectral and spatial resolutions, the concept of temporal resolution is also important to consider in remote sensing. Temporal resolution refers to how frequently a given location on the earth surface can be imaged by the sensor system. Temporal resolution is becoming important when persistent clouds offer limited clear views of the Earth's surface (often in the tropics) or when remote sensing is used to monitor dynamic processes causing considerable changes in a short time, like floods or forest fires.

Sentinel Satellites Initiate New Era in Earth Observation

In recent years we have experienced a paradigm shift with the advent of the Sentinel missions of the European Copernicus initiative. The Sentinel missions will, amongst others, provide long-term access to enhanced high and medium spatial resolution radar and optical observations opening a new era for the systematic mapping, assessment and monitoring of coastal and river morphological dynamics worldwide:

- The C-band radar of the Sentinel-1 mission provides all-weather and day-and-night imagery that will be extremely useful for consistently monitoring surface water dynamics in cloudy conditions, i.e. during the monsoon season in Bangladesh.
- The large footprint size of the Sentinel-2 data along with its short revisit time and its systematic acquisition policy allows rapid changes in ecosystems to be precisely monitored and is ideally suited to monitor dynamic ecosystems.

- The Sentinel-2 mission ideally complement the longest continuously acquired collection of optical observations at high resolution made by the family of Landsat imagers (operational since 1972) which are freely accessible and offer a unique opportunity to assess the historical conditions of the coastal areas and river deltas.

The availability of the growing volume of environmental data from the Sentinels, combined with data from long-term Earth Observation archives (e.g. ERS, ENVISAT and Landsat) represents a unique opportunity for the operational usage of EO to monitor the dynamics of the coastal and river delta areas in Bangladesh and derive early warning indicators of morphological changes and climate change impacts.

Remote sensing of vegetation and water

Vegetation is one of the most important components of ecosystems, and a significant part of the information recorded by satellite sensors is being used to map vegetation types and extract vegetation biophysical properties. The reflection of optical radiation by plants tends to have a number of common features different from other surface objects, including a low reflectance in the visible spectral region (0.5 to 0.7 µm), a steep increase in reflectance (around 0.7 µm) and a high reflectance in the near-infrared (0.7 to 1.1 µm) parts of the spectrum (Figure 1). The low reflectance in the visible region is due to chlorophyll absorption by green leaves, while the high reflectance of the near-infrared (NIR) region is due to the strong reflectance of plant mesophyll tissue. At the longer shortwave infrared wavelengths (1.6 to 2.5 µm) the reflectance is mainly a function of the plant water content. The exact spectral reflection for a given vegetation surface, across the whole optical region, will vary with respect to pigments, leaf structure, and leaf water content as well as to the density and structure of the canopy. It also changes in time according to the changes in phenology and canopy roughness.

Water has a completely different influence on the reflected energy (Figure 1), measured at the sensor. It has very strong absorption in the longer wavelengths, thus almost no light is reflected in the near or shortwave infrared parts of the spectrum. Surface water areas can, therefore, be easily discriminated from vegetated areas. To a certain extent, light is also penetrating into water bodies and here, it reacts with water constituents such as chlorophyll from microalgae or with sediment floating in the water column. The scattering and absorption on these particles allow the assessment of the concentration of the respective water constituents from the measured signal.

Active microwave remote sensing can penetrate the canopy at varying depths depending on the frequency, polarization and incidence angle of the sensor. As it was shown about the optical spectrum, varying microwave wavelengths have been found particularly useful for certain purposes. Radar in the shorter wavelengths (e.g., X-band) is sensitive to the uppermost parts of the canopy (i.e., twigs and leaves) and provide information on canopy roughness whereas the medium wavelengths (e.g., C-band) interact with the volume of the canopy and provide information on leaf biomass. The longer wavelengths (e.g. L-band) penetrate deeper into the canopy, and provide information on wood biomass. Sensitivity to soil moisture have been registered in both X-, C- and L-band SAR imagery. Moreover, different polarizations can be used to accentuate the backscatter from features with particular orientations, thus, providing further information on surface roughness and geometric arrangement.

Image classification

Image classification and analysis operations are used to digitally identify and classify pixel values into information. The most frequent objective of image classification is to transform the acquired satellite data into a land cover map. Classification is usually performed on multispectral or hyperspectral data sets (A), and this process assigns each pixel in the image to a particular class or theme (B) based on statistical characteristics of the pixel values (Figure).

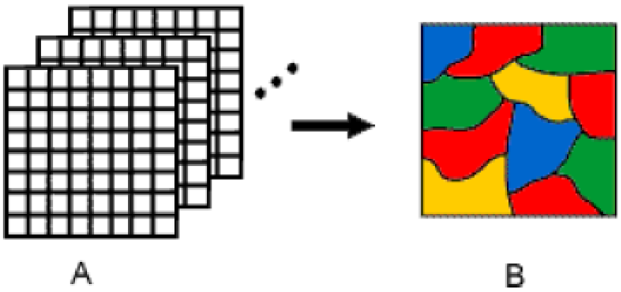

Figure 2: From multi-spectral imagery (A) to classified thematic map (B)

Image thresholding and decision trees

One of the most simple, effective and efficient methods for deriving information from remotely sensed data is to use image thresholding. The theory is simple: Using selected spectral bands or indices, some features appear brighter or darker in the image than the others, and these features or classes can be separated by applying a value threshold that separates the bright and dark areas. Classic examples include water appearing dark in the visible bands compared to the brighter land cover classes, or vegetation having higher normalized difference vegetation index (NDVI) values compared to non-vegetated areas. Image thresholding involves deciding upon an appropriate 'threshold' value to separate the class of interest – for example, all pixels in an image with NDVI values below 0.2 could be classified as non-vegetated.

Sometimes it is beneficial or necessary to refine the mapping of the target class by using multiple bands and associated thresholds. For example, an NDVI threshold could give you non-vegetated areas, but to separate this class into non-vegetated land and water, you could combine it with a threshold on the normalized difference water index (NDWI). Using multiple bands and thresholds can be considered a user-defined decision tree approach.

Machine learning classification

Classification of satellite imagery using machine learning is one of the most commonly used methods to derive thematic layers from satellite imagery. Multi-spectral data sets are most commonly used for machine learning classification. The aim of this image classification method is to assign each pixel (or segment if working with image objects) to a specific class or theme based on the image characteristics, which are 'learned' by the computer.

There are two commonly used methods for machine learning classification: Supervised and Unsupervised. Supervised classification requires training a classification algorithm to recognise pre-defined land cover classes. For training a supervised classifier, a training data set is required. This training dataset consists of samples of known land cover/use within an image for all classes that want to be mapped. These samples are then used to derive idealized spectral signatures (i.e. a unique combination of reflected electromagnetic radiation in different spectral bands) for the land covers of interest. Each pixel in the image is then compared to these signatures and assigned to the class it most closely resembles. The most high-performing machine learning algorithms used in remote sensing include Random Forests, Support Vector Machines, Gradient Boosting, and Neural Networks, among others. Generally, there is not much difference in performance between these machine learning methods regarding classification accuracies[7], although some have higher levels of complexity and require varying levels of parameter tuning/adjustment.

Unsupervised classification requires no pre-training, and therefore no training data set is needed. Instead, the image is classified into several classes using an automatic clustering algorithm. In the next step, information from fieldwork or other ancillary data is used to label the resulting clusters into information classes, i.e. the required land cover classes. Commonly used unsupervised methods for remote sensing imagery include K-means or Self-Organizing Maps. Although there are some advantages to the unsupervised approach (e.g. it does not need training data), supervised classification is widely recognized for providing more accurate maps.

In-situ data requirements

For both supervised and unsupervised classification methods, a priori information about the desired classes is to be used either in the learning phase of the supervised classification, or in the post-classification phase of the unsupervised method, for the identification of classes. Even with the physically based models, the parameters of the models as well as the verification of the results need the use of in situ data. Therefore, in remote sensing applications, field observations and measurements are needed at different stages of the work, however, there is no generic rule for capturing in situ data.

[7] Fernandez-Delgado, M., Cernadas, E. & Barro, S. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems ? J. Mach. Learn. Res. 15, 3133–3181 (2014).

3 Software and cloud-based processing tools

This section describes the key tools proposed in this manual, allowing for integrated satellite data acquisition, analysis and visualization to assess coastal- and river morphological dynamics in Bangladesh.

3.1 ArcMap

State-of-art desktop suite of GIS software products

ArcGIS consists from an assortment of software products created by Environmental Systems Research Institute (Esri), including desktop, server, mobile, hosted, and online GIS products. It is a comprehensive location-based analytics and imagery and remote-sensing software platform, offering advanced data processing, analysis and visualization tools. The software package has integrated image analytics modules and tools for advanced imagery and raster analysis, aimed to accommodate data from i.e., satellites and aerial technology. The ArcGIS package allows for easy integration with other geospatial data suites and provides comprehensive documentation, step-by-step guides, tutorials and industry standard tools.

ArcMap is one of the primary desktop GIS applications within the ArcGIS desktop suite. This is the environment where a user can create maps and access most of the ArcGIS functionalities; adding and editing data, querying and symbolizing map layers, and creating map layouts and visualisations for printing. While the main software consists from a significant number of functionalities and tools, extensions can also purchased (although some are free) to extent its functionality. Some of the frequently used extensions include[8]:

- **Spatial Analyst**
  Provides a rich set of spatial modelling and analysis tools for both raster (cell-based) and feature (vector) data. The tools allow the user to create, query, map, and analyze cell-based raster data; perform integrated raster/vector analysis; derive new information from existing data; query information across multiple data layers; and fully integrate cell-based raster data with traditional vector data sources.
- **3D Analyst**
  Provides a collection of various geoprocessing tools that enable a wide variety of analytical, data management, and data conversion operations on surface models and three-dimensional vector data. This includes extruding polygons (such as parcels and building footprints) and draping surfaces (such as orthophotos) on elevation models.
- **Geostatistical Analyst**
  Provides a series of tools to create continuous surfaces from measurements stored in a point feature layer or raster layer or by using

[8] For more information on ArcMap extensions, visit https://desktop.arcgis.com/en/arcmap/latest/tools/main/a-quick-tour-of-geoprocessing-tool-references.htm

polygon centroids. Sample points could for example be soil salinity measurements, water depths or erosion rates. Geostatistical Analyst provides a comprehensive set of tools for creating surfaces that can be used to visualize, analyse, and understand spatial phenomena.

- **Spatial statistics**
  Provides a number of statistical tools for analysing spatial distributions, patterns, processes, and relationships. The tools in the Spatial Statistics toolbox allow you to summarize the salient characteristics of a spatial distribution (determine the mean centre or overarching directional trend, for example), identify statistically significant spatial clusters (hot spots/cold spots) or spatial outliers, assess overall patterns of clustering or dispersion, group features based on attribute similarities, identify an appropriate scale of analysis, and explore spatial relationships. In addition, for those tools written with Python, the source code is available to encourage you to learn from, modify, extend, and/or share these and other analysis tools with others.

For more information on ArcMap, its functions, tools and user interface please visit https://desktop.arcgis.com/en/arcmap/ or visit the below quick guides to get set up:

- **A quick tour of ArcMap**
  https://desktop.arcgis.com/en/arcmap/latest/map/main/a-quick-tour-of-arcmap.htm
  Describes the basic functionalities, concepts and functions within the ArcMap user interface. Use this guide to get familiar with ArcMap document formats; ArcMap views; Map layers; ArcMap windows and; ArcMap catalog.
- **A quick tour of ArcCatalog**
  https://desktop.arcgis.com/en/arcmap/latest/get-started/introduction/a-quick-tour-of-arccatalog.htm
  Describes the functions and data management options and workflows available in ArcCatalog. Use this guide to learn more about setting up your own workspace; Work with geodatabases; Search for data and tools; Work with data descriptions and metadata and; Learn how to organise your map documents, imagery and other data files, geoprocessing models, layers, etc.
- **A quick tour of geoprocessing**
  https://desktop.arcgis.com/en/arcmap/latest/analyze/main/a-quick-tour-of-geoprocessing.htm
  Describes the geoprocessing functions and toolboxes and explains how to work and operate these. Use this guide to get familiar with tools and toolboxes; Tool dialogue box; Results window; background processing; Models and Modelbuilder and; Python and scripting.

3.2 Digital Shoreline Analysis System (DSAS)

This section will be finalised in consultation with the national stakeholders as part of the preparations leading into the design and scope of the capacity building workshop.

3.3 Google Earth Engine

Google Earth Engine, established in 2010, is a web-based cloud processing platform and satellite data repository that provides global-time series satellite imagery and vector data and access to software and algorithm for data processing.[9] The multi-petabyte analysis-ready data catalogue, stored in the public data archive, includes historical satellite images dating back more than forty years.[10] Besides the collection of raw unprocessed satellite imagery, Google Earth Engine also provides access to various satellite-based products, including indices, composites, elevation models, land cover data, etc.

Data from the Earth Engine servers can be accessed using the JavaScript based Google Earth Engine Internet-accessible application programming interface (API). The Earth Engine (EE) Code Editor, available from code.earthengine.google.com, is a web-based interactive development environment (IDE) for the Earth Engine JavaScript API, which allow users to create and run custom algorithms to retrieve and process data rapidly in the cloud.

As illustrated in Figure below, the IDE include the following components (GOOGLE, 2019);

- JavaScript code editor;
- Map display for visualizing geospatial datasets;
- API reference documentation (Docs tab);
- Git-based script manager (Scripts tab);
- Console output (Console tab);
- Task manager (Tasks tab) to handle long-running queries;
- Interactive map query (Inspector tab);
- Search of the data archive or saved scripts;
- Geometry drawing tools

[9] Kumar, L., & Mutanga, O. (2018). Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. Remote Sensing, 10(10), 1509. https://doi.org/10.3390/rs10101509

[10] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

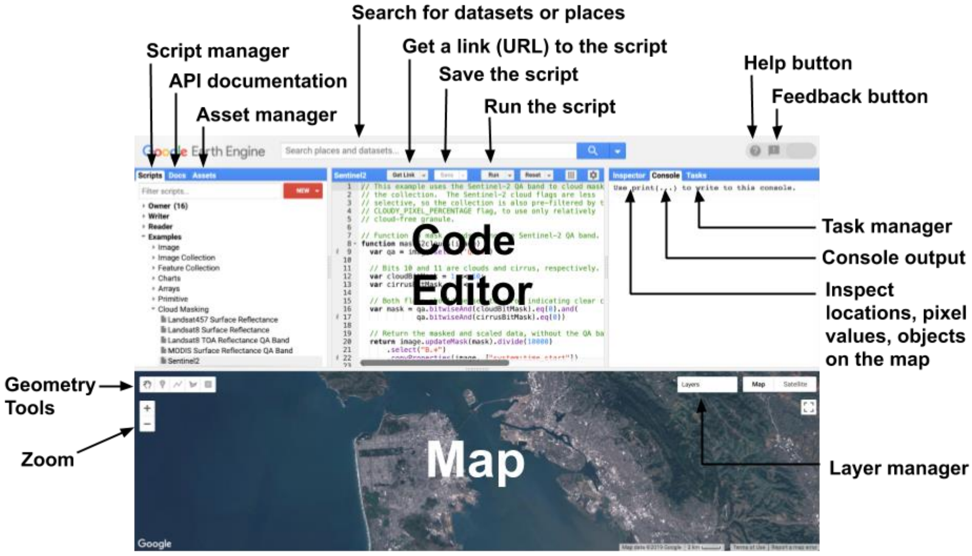

Figure 3: Components of the Earth Engine Code Editor

For more information on the Google Earth Engine environment, its functions, tools and user interface, please visit https://developers.google.com/earth-engine/tutorials#introduction-to-earth-engine-condensed. This resource provides a condensed, yet comprehensive, list of self-paced tutorials as well as video guides and introductions, ideal for a basic introduction to all the key functions of Google Earth Engine.

4 Methodology

4.1 Automated land-water classification

In order to assess river and coastal morphology, specifically sites of erosion and accretion, the delineation of riverbanks and coastlines must be extracted. Historically, this has been digitized manually, however, land-water can also be determined in a more automated process involving unsupervised methods (McFeeters, 1996; Feyisa et al., 2014) or mapped as a class out of a category of land cover classes via supervised classification (Chen et al., 2015).

The mapping of surface water from satellite imagery has a long heritage and today high-resolution products are available globally (Peckel et al. 2016; Pickens et al. 2020). However, those global products have a 1-year latency and relies solely on optical data (cf. US Landsat imagery). As a consequence, they do not support operational applications and will underestimate the water presence in cloudy regions. SAR imagery is essential to ensure all weather operational mapping capabilities for surface water extent. Therefore, state-of-the-art surface water extent mapping from satellite remote sensing should use a combination of Optical and SAR imagery. But only recently and through the Copernicus Sentinel-1 mission, SAR data has recently become globally available free of charge and with short latency. Meanwhile, the advancements in compute power and data accessibility (e.g. Google Earth Engine) enable the rapid and automated delineation of these land-water boundaries at scale. Thus, an automated processing pipeline (Figure 4) was developed to determine land-water classification for analysis into coastal and river morphology.

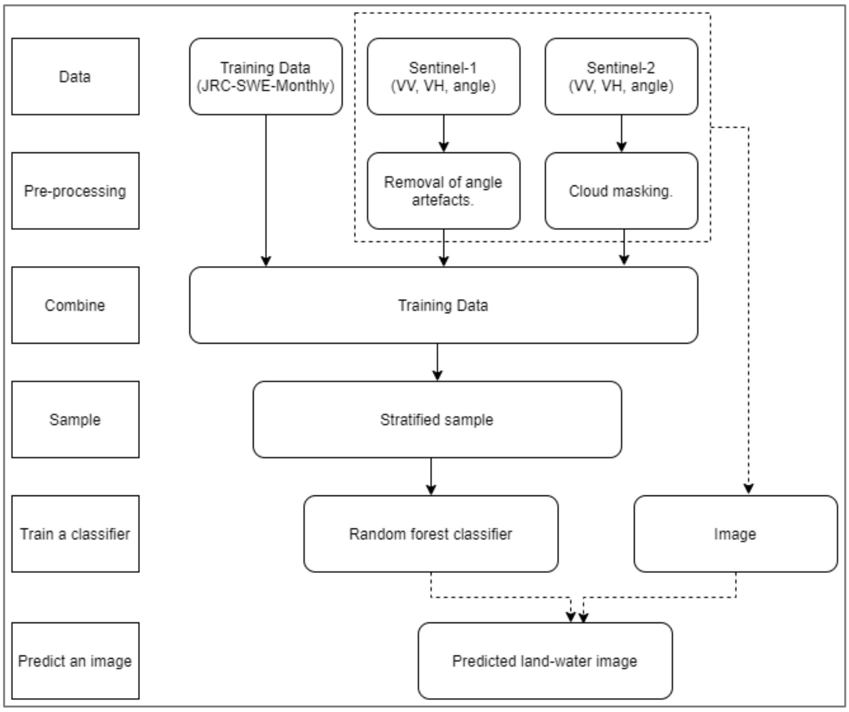

Figure 4: Supervised land-water classification method

The spectral data like Sentinel-2 must be cleaned to remove noise from clouds and cloud shadows in the optical data and to remove angle related artefacts in the radar data like Sentinel-1. A list of water extraction indices (NDVI, NDWI, MNDWI, ND1208, ND0204) derived from spectral data have been shown (Fisher et al., 2016) to enhance the separation of land-water and were thus calculated for Sentinel-2 data. For Sentinel-1 data the backscatter bands (VV and VH) as well as the incidence angle was used because higher backscatter variances resulting from variations in the incidence angle, which would be taken into account by a supervised machine learning algorithm. Figure 5 shows the Sentinel-2 derived NDVI and Sentinel-1 backscatter band of VV.

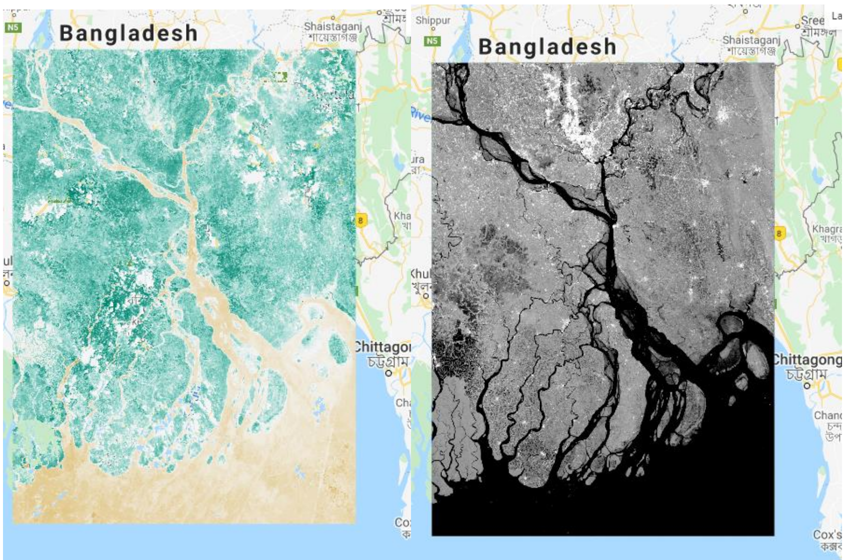

Figure 5: Sentinel-2 NDVI & Sentinel-1 VV

Supervised classification requires labelled training samples. For this, the highly appraised global surface water map (30 m spatial resolution and monthly temporal resolution) by Pekel et al. (2016) is used (Figure 6). The JRC surface water layer can be combined with Sentinel-2 and Sentinel-1 data from which a classifier can be trained to predict land-water. Further, they must be mosaiced and composited into a single image for that time period. For the scope of this training, a single monthly image composite was generated.

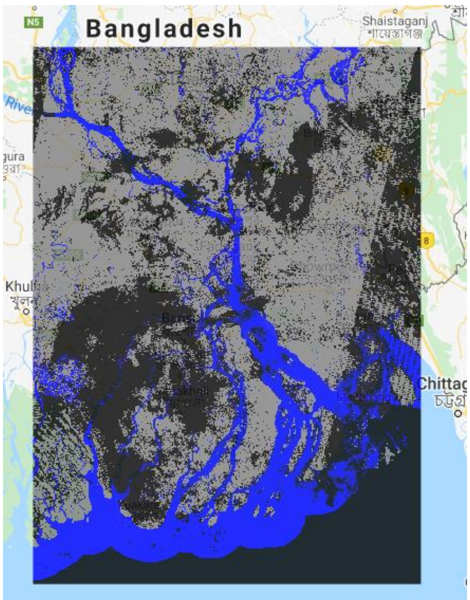

Figure 6: Training data, JRC surface water extent

The Random Forest (RF) (Breiman 2001) is a non-parametric machine learning classifier widely used for image classification due to its simple parameterization and high classification accuracy (Pelletier et al., 2016). . RF randomly split the inputs into user-defined number of trees (=50) as larger values are known to have little influence on the overall classification accuracy (Breiman 2001) and water is not a tricky class to map. RF assign the class labels based on the majority vote among all bootstrapped classification trees. By splitting the training data into an internal training and testing sets, information such as out of bag error rate can be extracted to quickly evaluate the classifier's performance. A more robust accuracy assessment should utilise fully independent validation data if available. The model is reliable once the error rate is determined to be acceptable. Figure 7 illustrates the land-water probability derived from above approach.

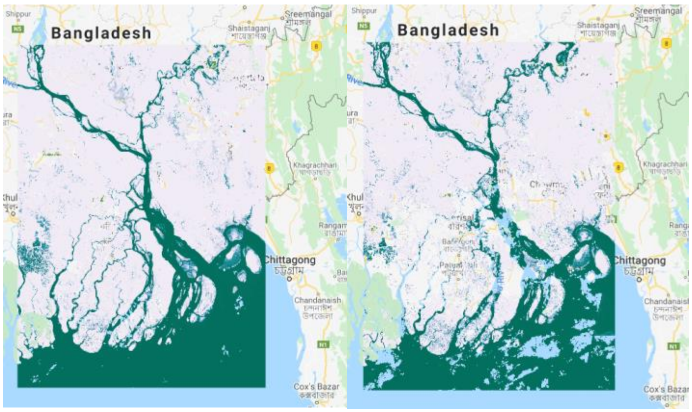

Figure 7: The original data 2018-06 (left) and new data 2018-07 (right) land-water probability

Land-water probability are predicted using Sentinel-2 and Sentinel-1, separately. To yield a monthly land-water product, a fusion of two types of predictions is needed. In the predictions, NoData values exist due to cloud existing in the Sentinel-2 images for that month. This can render our predictions 'incomplete' and could be supplemented with a Sentinel-1 only prediction in these areas, and likewise when there is no sentinel-2 data we could use Sentinel-1 only. Examining the error rates of each classifier it is obvious that the combined approach (Sentinel-1 + Sentinel-2) is most accurate and the Sentinel-1 only approach is the least accurate. Thus, we need to apply logic such that the best available classifier is used, and a 'complete' prediction can be created for each time period. The logic is shown below in Figure 8.

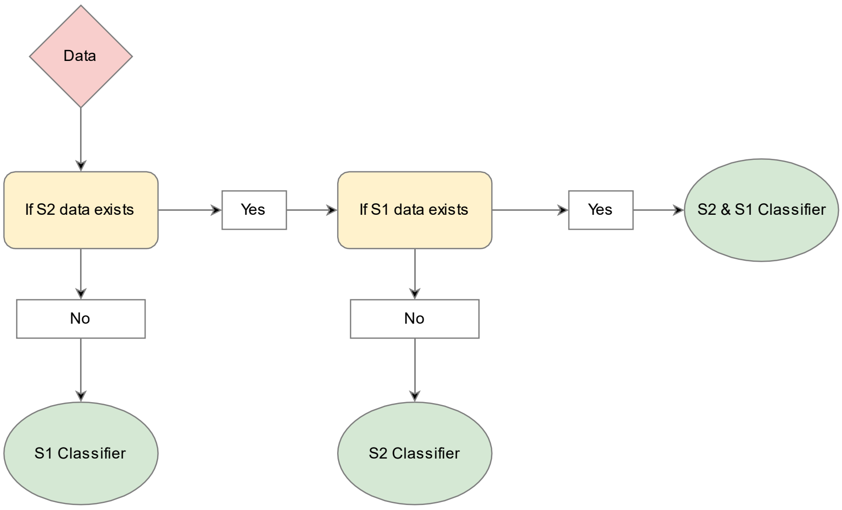

Figure 8: Flowchart demonstrating a rule-based system for generating full-coverage predictions

4.2 River erosion and early warning

Stacking a series of land-water classifications across time facilitates the interpretation of trends and ability to infer future changes. Applying the land-water classification methods in section 4.1 to an entire time-series can enable us to determine areas of change (erosion or accretion) and the rate at which that occurs.

Applying sections 4.1 and the logic explained above creates a time-series image collection of even interval predictions (e.g. bi-weekly or monthly). Using this monthly time series land-water product, a surface water frequency layer (SWF) can be created simply by summing the number of water observations and dividing by the total number of observations for each pixel. This product can be useful for indicating areas of permanent water vs seasonal inundation (Figure 9).

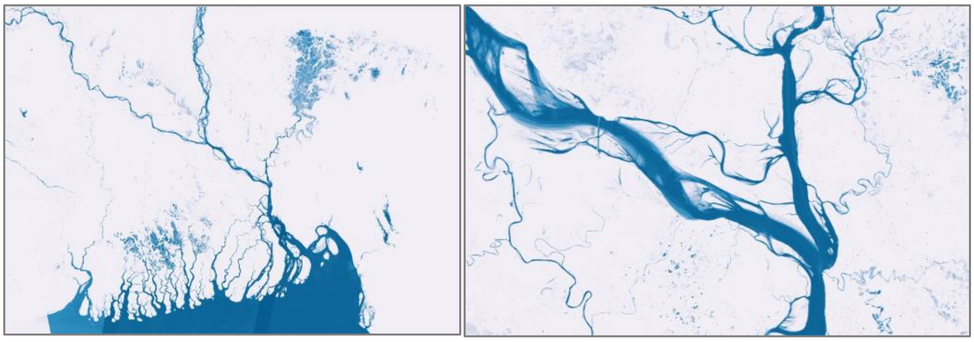

Figure 9: Surface water frequency for Bangladesh (left) and to the South of Dhaka (right)

SWF whilst interesting does not consider time, thus we cannot infer if areas are getting wetter or drier (erosion or accretion). Linear regression is a method to model the relationship between a dependent variable (e.g. time) and the independent variable (e.g. land-water probability) assuming a linear relationship between the variables. Consider the following linear model (equation 1), where β_1 is the slope, β_0 is the offset and e_t is a random error. The slope indicates the direction and altitude of changes during time, so called as trends.

$$p_t = \beta_0 + \beta_1 t + e_t \quad (1)$$

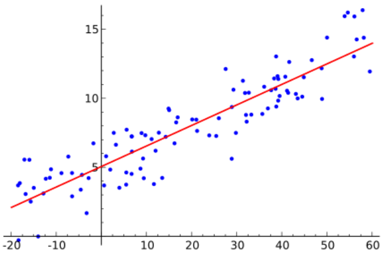

Figure 10: Example of a linear relationship between two variables. The dots are the observations and the line is the trendline modelled from the observations.

When visualizing the slope (Figure 11), we can identify areas of positive change (erosion), no change (stability), negative change (accretion). Further, we can estimate the rate of change per day (since that is our dependent variable). Note that does not normalize for seasonal variations for example due to non-constant rainfall and snow meltwater.

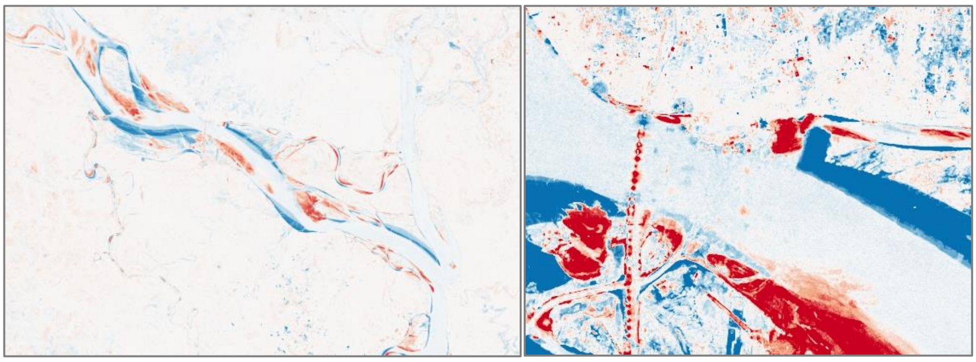

Figure 11: Visualisation of the slope from the linear regression of land-water probability. Red is accretion (negative), Blue is erosion (positive). The left image shows an area South of Dhaka and the right image shows the construction of the Padma bridge.

Computing linear regression across a subset of the time-series can enable greater insights into short-term changes that would otherwise be ‘averaged’ out in the long-term regression analysis (Figure 12). On the other hand, it may be skewed more by erroneous values (e.g. due to cloud).

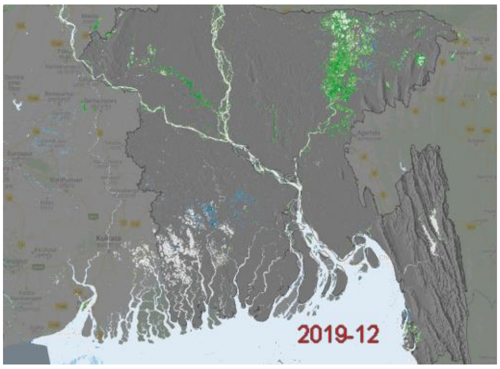

Figure 12: Short term linear regression for the 2 months prior to 2019-12-01. Green is negative change (getting drier) and blue is positive (getting wetter).

The short-term regression can be a useful data-source for an early warning system for areas at risk of erosion. Thresholding values that are getting wetter (positive trend, Figure 13) and then buffering them by an 'at-risk' distance is a simple yet effective method to determine this since the direction of erosion is likely to continue. Coupling these at-risk areas with land-cover maps could enable an economic cost analysis of these areas.

Figure 13: Short-term linear regression (left) applying a threshold and buffer (right)

4.2.1 Coastal stability and dynamics

The same methods (section 4.1 and 4.2) can be applied to analysis in coastal regions; however, the impacts of tidal variations must be considered, especially in lowland areas. Categorization of predictions into low tide, no tide and high tide requires an accurate timestamp of the image acquisition and a tide-timetable. Some descriptive tide height statistics can be generated, and the results shown below may be useful to determine appropriate thresholds for what constitutes high or low tide.

```
max: 1.406540036201477
mean: 0.1699419798748426
min: -1.0861330032348633
sample_sd: 0.644862255248537
sample_var: 0.41584732824422926
sum: 682.6569331572427
sum_sq: 1786.0549410248548
total_count: 4017
total_sd: 0.6447819836038309
total_var: 0.4157438063800908
valid_count: 4017
```

Figure 14: Descriptive statistics on tidal data recorded at 10.30am

The Sentinel-2 timestamp is stored properly in the image metadata, however, while this is useful to determine which day the overpass was, the overpass time is always 10.30am. This is to ensure the angle of sunlight, and thus shadow and ground illumination is consistent as well as minimizing cloud cover[11]. Knowing the time, we can join our image collection to the corresponding tide value at that time and filter based upon the tidal value. With the categorized data linear regression (both long-term and short-term) can be computed to determine any trends.

[11] https://sentinel.esa.int/web/sentinel/missions/sentinel-2/satellite-description/orbit

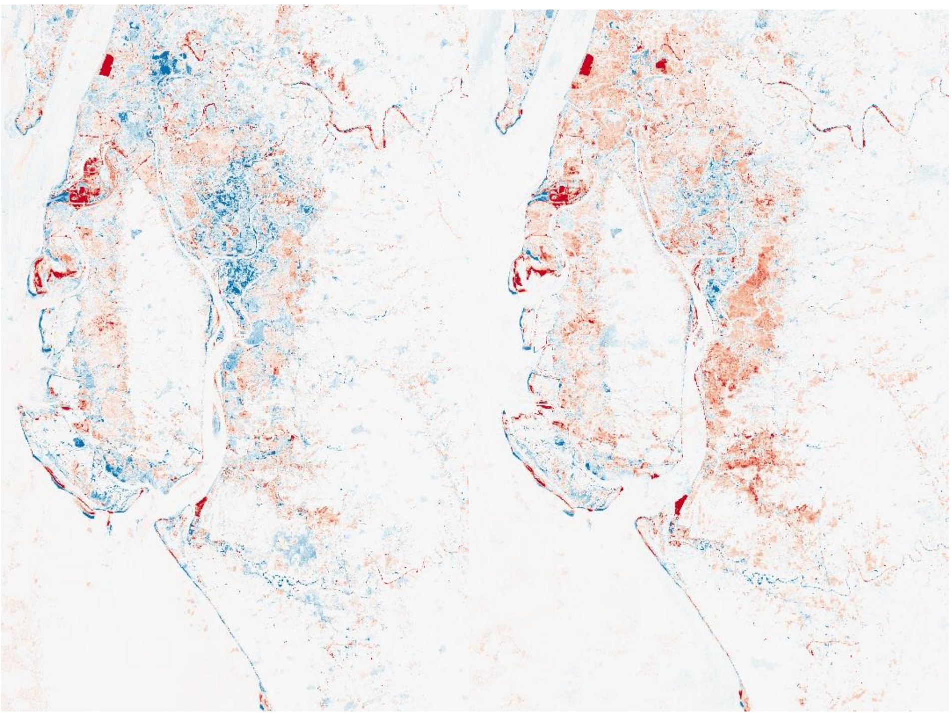

Figure 15: Linear regression for low (left) and high (right) tides around Cox's Bazar

4.2.2 Soil Salinity Analysis

Soil salinity is one of most severe environmental hazards, particularly in the coastal areas. The consequences of salinization include land degradation and desertification, which results in the reduction of agricultural production if occurred to arable land. Therefore, soil salinity monitoring is crucial to make decisions on land rehabilitation for the affected area (Hoa et al., 2019). The assessment of soil salinity can be done via field measurements on electrical conductivity (EC) of soil at field levels and via mapping soil salinity using remote sensing techniques based on the correlation between satellite observations and soil characterises such as reflectance spectra and dielectric constant. In addition, remote sensing techniques have its known advantages on large scale monitoring at regular temporal intervals.

Researchers has been investigated various machine learning/data mining techniques to retrieve soil salinity from the correlation between satellite observations (such as spectral bands, transformed spectral bands, vegetation indices and backscatter intensity) and EC field measurements using different types of satellite imagery (Taghadosi et al., 2019). However, when without field measurements, vegetation indices (VIs) (or salinity index by naming differently when multiplying the VIs with a minus sign) can be adopted as an indirect salinity indicator because VIs can indicate soil in its lower value range (e.g 0-0.2) and reflect the negative impacts of saline soil on vegetation. Amongst the VIs, Enhanced Vegetation Index (EVI) is proved to be better than NDVI, which has saturation issue in dense vegetation and is sensitive to soil background (Lobell et al., 2010). By using Landsat imagery available since the 1970s, one can assess the temporal development of vegetation at 30 meters resolution over nearly 40 years though time series analysis, as introduced in the above session. Assuming when vegetation cover was not changed and environment was not significantly changed over time, then a negative trend of vegetation indices would indicate an increase of soil salinity, vice versa.

We conducted the Landsat-based trend analysis from 1988 to 2019 on Enhanced Salinity Index (ESI), which is the minus EVI (Tran et al., 2019). Firstly, we prepared the annual ESI during the dry season in Bangladesh (January to May) using the medoid composite approach (Flood, 2013). Furthermore, due to the shift of sensors, the between sensor calibration was applied (Roy et al., 2016). A linear regression was then conducted on the annual time series to identify the significant increase in soil salinity (Figure 16).

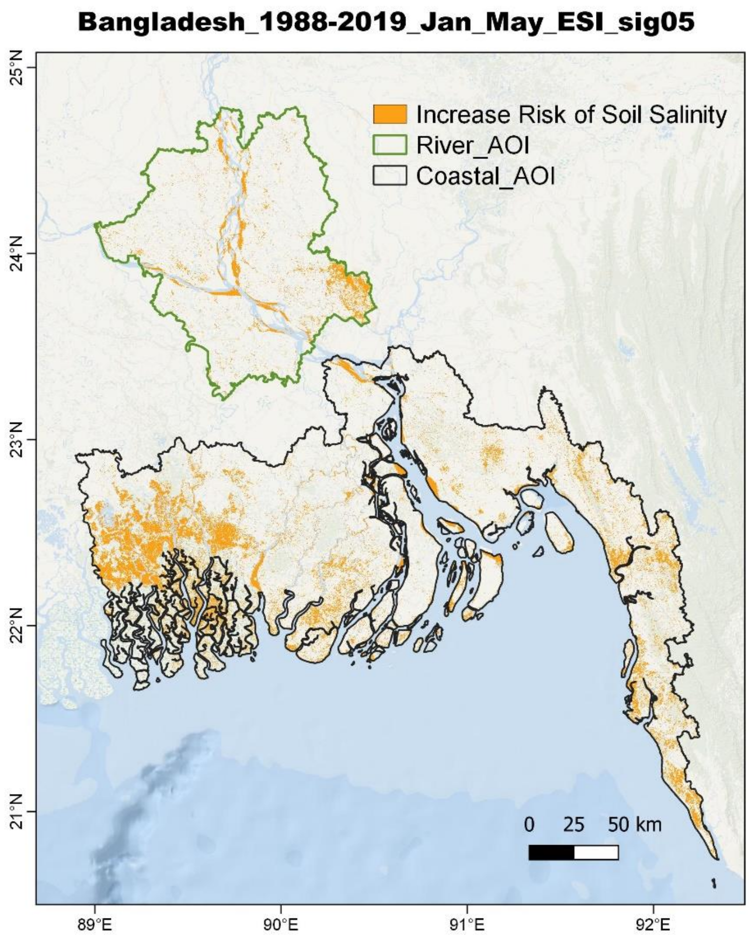

Figure 16: Significant increase of soil salinity (p<0.05) in Bangladesh over 1988-2019

5 Workflows

This section is divided into two levels of analysis, where sub-section 5.1 at a beginner-level introduces the step-by-step basics about how to search, find, visualize and preprocess remotely sensed imagery in Google Earth Engine (GEE) and sub-sections 5.2-5.5 at a pro-level illustrate the methodology developed for land-water classification, river erosion and early warning, coastal stability and dynamics and lastly soil salinity in Bangladesh. Both of the analysis requires GEE account, please make sure to sign up here: https://signup.earthengine.google.com/#!/

This section will be finalised in consultation with the national stakeholders as part of the preparations leading into the design and scope of the capacity building workshop.