Automated Extraction of Time-Variable Water Surfaces based on Google Earth Engine

Christian Schwatke, Daniel Scherer

Deutsches Geodätisches Forschungsinstitut Technische Universität München (DGFI-TUM)

ESA-ESRIN | Frascati, Italy | 27-28 March 2018, 2nd Mapping Water Bodies from Space Conference
Motivation

• Inland waters play an essential role for monitoring the hydrological water cycle
• Number of in-situ stations has been decreasing since 1980 (see GRDC)
• Remote sensing can be used to monitor storage changes also in remote areas
• Satellite altimetry can provide height information
• Radar or optical images can be used to retrieve surface information
• DGFI-TUM already developed a “Database for Hydrological Time Series of Inland Waters” which provides more than 750 water level time series from satellite altimetry for inland waters
• A new “Automated Water Area Extraction Tool” (AWAX) using Google Earth Engine has been developed in order to extract monthly water masks and area extents

http://dahiti.dgfi.tum.de
Google Earth Engine

• Planetary-scale platform for Earth science data and analysis (Gorelick et al., 2017)
• Data catalog includes more than 200 public available datasets (imagery, geophysical, climate and demographic)
• Advantages of Google Earth Engine:
  • Cloud-based data storage
  • Cloud-based geospatial data processing
• Online development in the Earth Engine Code Editor (JavaScript)
• Local development using Python-API

General Processing Scheme of Google Earth Engine
Google Earth Engine - Data

- **USGS Landsat 4 Surface Reflectance** (*ImageCollection ID: LANDSAT/LT04/C01/T1_SR*)
  
  1982-08-24 – 1993-12-14, 22152 Scenes, 30m,
  
  Masks: CFMASK (C Function of Mask) for cloud, shadow, water and snow/ice,
  
  Atm. Corr.: LEDAPS (Landsat Ecosystem Disturbance Adaptive Processing System)

- **USGS Landsat 5 Surface Reflectance** (*LANDSAT/LT05/C01/T1_SR*)
  
  1984-01-01 – 2012-05-05, 1510066 Scenes, 30m, Masks: CFMASK, Atm. Corr.: LEDAPS

- **USGS Landsat 7 Surface Reflectance** (*LANDSAT/LE07/C01/T1_SR*)
  
  1999-01-01 – active, 1677441 Scenes, 30m, Masks: CFMASK, Atm. Corr.: LEDAPS

- **USGS Landsat 8 Surface Reflectance** (*LANDSAT/LC08/C01/T1_SR*)
  
  2013-04-11 – active, 679140 Scenes, 30m, Masks: CFMASK, Atm. Corr.: LEDAPS
Google Earth Engine - Data

- **USGS Landsat 4 Surface Reflectance** *(ImageCollection ID: LANDSAT/LT04/C01/T1_SR)*
  1982-08-24 – 1993-12-14, 22152 Scenes, 30m,
  Masks: CFMASK (C Function of Mask) for cloud, shadow, water and snow/ice,
  Atm. Corr.: LEDAPS (Landsat Ecosystem Disturbance Adaptive Processing System)

- **USGS Landsat 5 Surface Reflectance** *(LANDSAT/LT05/C01/T1_SR)*
  1984-01-01 – 2012-05-05, 1510066 Scenes, 30m, Masks: CFMASK, Atm. Corr.: LEDAPS

- **USGS Landsat 7 Surface Reflectance** *(LANDSAT/LE07/C01/T1_SR)*
  1999-01-01 – active, 1677441 Scenes, 30m, Masks: CFMASK, Atm. Corr.: LEDAPS

- **USGS Landsat 8 Surface Reflectance** *(LANDSAT/LC08/C01/T1_SR)*
  2013-04-11 – active, 679140 Scenes, 30m, Masks: CFMASK, Atm. Corr.: LEDAPS

- **Sentinel-2 MSI: MultiSpectral Instrument, Level-1C** *(COPERNICUS/S2)*
  2015-06-23 – active, 4010742 Scenes, 10m/20m,
  Mask: ESA for clouds only *(No mask for shadow, water, snow/ice)*
  Atm. Corr.: Not applied

Sentinel-2 not used in combined approach
Methodology

Start

Initialization of AOI (Polygon, Location)

Compute Probability Mask

Fill Data Gaps of Monthly Mask

Read Input Data

Calculate Threshold

Fill Gaps

Export Results

No

All months processed?

Yes

Compute Probability Mask

Fill Data Gaps of Monthly Mask

Read Input Data

Calculate Threshold

Fill Gaps

Export Results

No

All months processed?

Yes

Calculate Time Series

Surface Area Time Series

Shape File

Probability Mask

Filled Monthly Mask

Filled Monthly Shape File

Surface Area in km²
Study Area

Lake Tawakoni, Texas, USA
Surface Area: ~149 km²
Type: Reservoir
Methodology

Start

Initialization of AOI (Polygon, Location)

Computation of Monthly Mask with Data Gaps

Create GEE Composite

Calculate Indexes

Land/Water Classification

Land/Water Masking

Monthly Mask with Gaps

Compute Probability Mask

Fill Data Gaps of Monthly Mask

Read Input Data

Calculate Threshold

Fill Gaps

Export Results

No
All months processed?

Yes

Monthly Mask with Gaps

Shape File

Probability Mask

Filled Monthly Mask

Filled Monthly Shape File

Surface Area in km²

All months processed?

Yes

Calculate Time Series

Surface Area Time Series

C. Schwatke | 2nd Mapping Water Bodies from Space Conference | Frascati, Italy | 27-28 March 2018
Extract Composite

- Create composite image with Google Earth Engine
- Pixel-based merging is performed by using a cloud-score to get best cloud-free image
- AWAX uses six bands (Red, Green, Blue, NIR, SWIR1, SWIR2) and three masks (Voids, Clouds, Snow)
- For Lake Tawakoni in 12/2014 are 5 scenes of Landsat 7 respectively Landsat 8 available
Calculate Indexes

- **Modified Normalized Difference Water Index** (Xu, 2006)
  \[ MNDWI = \frac{Green - SWIR_1}{Green + SWIR_1} \]

- **New Water Index** (Li, 2016)
  \[ NWI = \frac{Blue - (NIR + SWIR_1 + SWIR_2)}{Blue + (NIR + SWIR_1 + SWIR_2)} \times C \]

- **Automated Water Extraction Index for Non-Shadow Areas and Shadow Areas** (Feyisa et al, 2014)
  \[
  AWEI_{nsh} = 4 \times (Green - SWIR_1) - (0.25 \times NIR + 2.75 \times SWIR_2)
  
  AWEI_{sh} = Blue + 2.5 \times Green - 1.5 \times (NIR + SWIR_1) - 0.25 \times SWIR_2
  \]

- **Tasseled Cap – Wetness** (Crest, 1985)
  \[ TC_{wet} = 0.0315 \times Blue + 0.2021 \times Green + 0.3102 \times Red + 0.1594 \times NIR - 0.6806 \times SWIR_1 - 0.6109 \times SWIR_2 \]
1. Compute Indexes

- **MNDWI**
- **NWI**
- **AWEI_{sh}**
- **AWEI_{nsh}**
- **T_{C, wet}**
Calculate Histograms

1. Compute Indexes
2. Calculate Histograms
Apply Thresholds

1. Compute Indexes
2. Calculate Histograms
3. Apply Thresholds
1. Compute Indexes
2. Calculate Histograms
3. Apply Thresholds
4. Accumulate Masks
1. Compute Indexes
2. Calculate Histograms
3. Apply Thresholds
4. Accumulate Masks
5. Create Monthly Mask with Gaps
   - 0, 1 → No water
   - 2, 3 → Data gap
   - 4, 5 → Water

Accumulate Masks
Time Series with Data Gaps

Time Series of Surface Area for Lake Tawakoni with Data Gaps

\( R^2 : 0.37 \)
Methodology

Start

Initialization of AOI (Polygon, Location)

Computation of Monthly Mask with Data Gaps

Create GEE Composite

Calculate Indexes

Land/Water Classification

Land/Water Masking

No

All months processed?

Yes

Monthly Mask with Gaps

Compute Probability Mask

Shape File

Probability Mask

Fill Data Gaps of Monthly Mask

Read Input Data

Calculate Threshold

Fill Gaps

Export Results

No

All months processed?

Yes

Calculate Time Series

Surface Area Time Series

Filled Monthly Mask

Filled Monthly Shape File

Surface Area in km²
1. Accumulate all water/land pixels for all available month between 1984-2017
1. Accumulate all water/land pixels for all available month between 1984-2017
2. Calculate water probability mask
1. Accumulate all water/land pixels for all available month between 1984-2017
2. Calculate water probability mask
3. Label all connected surfaces
1. Accumulate all water/land pixels for all available month between 1984-2017
2. Calculate water probability mask
3. Label all connected surfaces
4. Select area using reference location
5. Create AOI Mask
Methodology

Start

Initialization of AOI (Polygon, Location)

Computation of Monthly Mask with Data Gaps

Create GEE Composite

Calculate Indexes

Land/Water Classification

Land/Water Masking

No

Yes

All months processed?

Monthly Mask with Gaps

Compute Probability Mask

Shape File

Probability Mask

Fill Data Gaps of Monthly Mask

Read Input Data

Calculate Threshold

Fill Gaps

Export Results

Yes

All months processed?

No

Calculate Time Series

Surface Area Time Series

Filled Monthly Mask

Filled Monthly Shape File

Surface Area in km²

C. Schwatke | 2nd Mapping Water Bodies from Space Conference | Frascati, Italy | 27-28 March 2018
Filling Monthly Data Gaps

Mask with Gaps (2014-12)  
Probability Mask ∩ AOI  
Probability Mask ∩ AOI ∩ Water Mask (2014-12)

0.5% Percentile Fill-Limit: 97 %
Filling Monthly Data Gaps

Mask with Gaps (2014-12)

Probability Mask ∩ AOI

Probability Mask ∩ AOI ∩ Water Mask (2014-12)

Probability Mask > 97%

0.5% Percentile Fill-Limit: 97 %

C. Schwatke | 2nd Mapping Water Bodies from Space Conference | Frascati, Italy | 27-28 March 2018
Filling Monthly Data Gaps

- Mask with Gaps (2014-12)
- Probability Mask ∩ AOI
- Probability Mask ∩ AOI ∩ Water Mask (2014-12)
- Probability Mask > 97%
- Fill Mask: Prob97 ∩ Gaps

0.5% Percentile Fill-Limit: 97 %
Filling Monthly Data Gaps

Mask with Gaps (2014-12)

Probability Mask ∩ AOI

Probability Mask ∩ AOI ∩ Water Mask (2014-12)

0.5% Percentile Fill-Limit: 97 %

Probability Mask > 97%

Fill Mask: Prob97 ∩ Gaps

Combined Mask

75.67 km²
29.98 km²
Filling Monthly Data Gaps

- **Mask with Gaps (2014-12)**
- **Probability Mask ∩ AOI**
- **Probability Mask ∩ AOI ∩ Water Mask (2014-12)**
- **Probability Mask > 97%**
- **Fill Mask: Prob97 ∩ Gaps**
- **Combined Mask**
- **Final Mask**

- **0.5% Percentile Fill-Limit: 97 %**

- **Confidence-Fill: 29.98 km²**
- **Fill Mask: Prob97 ∩ Gaps**
- **Combined Mask**
- **Final Mask**

C. Schwatke | 2nd Mapping Water Bodies from Space Conference | Frascati, Italy | 27-28 March 2018
Results – Tawakoni, Lake

$R^2 : 0.37$

$R^2 : 0.93$
Results – Sam Rayburn, Reservoir

- R² : 0.24
- R² : 0.87

C. Schwatke | 2nd Mapping Water Bodies from Space Conference | Frascati, Italy | 27-28 March 2018
Results – Toledo Bend, Reservoir

- **R²**: 0.19
- **R²**: 0.87
Results – Claiborne, Lake

C. Schwatke | 2nd Mapping Water Bodies from Space Conference | Frascati, Italy | 27-28 March 2018
Google Earth Engine is a powerful tool for data extraction and data processing to create monthly composites.

The combination of five water indexes ($MNDWI, NWI, AWEI_{sh}, AWEI_{nsh}, TC_{wet}$) has been used to create reliable monthly land-water masks with remaining data gaps.

More than 30 years of monthly masks are merged to calculate a long-term water probability mask.

Remaining data gaps are filled successfully by using the water probability mask.

Resulting surface area time series show a strong improvement of the correlation coefficients with respect to in-situ water levels:

- Tawakoni, Lake: $0.37 \rightarrow 0.93$
- Sam Rayburn, Reservoir: $0.24 \rightarrow 0.87$
- Toledo Bend, Reservoir: $0.19 \rightarrow 0.87$
- Claiborne Lake: $0.11 \rightarrow 0.77$

Sentinel-2 data will be added if Level 2 data including reliable masks (clouds etc.) are available.
Data Access

• Surface area time series for lakes and reservoirs will be successive available on DAHITI within the next weeks

• All data are freely available

• See you on DAHITI!

[Image of DAHITI website]

http://dahiti.dgfi.tum.de