#### Automatic Image Data Analytics from a Global Sentinel-2 Composite for the Study of Human Settlements

C. Corbane, P. Politis, P. Kempeneers, M. Pesaresi, D. Rodriguez, V. Syrris, P. Soille

christina.corban@ec.europa.eu







# BACKGROUND



## Big Data 4 Policy: the Global Human Settlement Layer

#### **Objective of GHSL**

Produce new evidence for decision making:

- 2030 Agenda for Sustainable Development (SDGs)
- Sendai Framework for Disaster Risk Reduction
- New Urban Agenda

#### Key requirements for policy support

- Reproducible, scientifically sound, synoptic
- Sustainable information production
- Free and open access
- Facilitate information sharing and collective knowledge building



#### **GEO Human Planet Initiative**





#### Discovering new elements in the universe of cities



China, EO data vs. Cities accounted in the UN World Urbanization Prospect 2016

#### GHSL Landsat Multitemporal in 2016





First available multitemporal assessment of built-up areas

#### GHSL Sentinel-1 in 2017

https://cidportal.jrc.ec.europa.eu/services/webview/jeodpp/databrowser/



S1A & S1B >7000 GRD scenes Dec 2016 – Dec 2017

10 m pixels spacing, Polarization: VV- VH Orbit: Desc and Asc Volume: 10 TB



Corbane, C. et al. (2017) Big earth data analytics on Sentinel-1 and Landsat imagery in support to global human settlements mapping, Big Earth Data, 1:1-2, 118-144,

## The building blocks

#### Symbolic Machine Learning



# Sustainability of built-up measurements through Copernicus





# SENTINEL-2



11

# Optimized Global Coverage of Sentinel-2 Input Data (2017-2018)



P. Kempeneers & P. Soille (2017) Optimizing Sentinel-2 image selection in a Big Data context, Big Earth Data, 1:1-2, 145-158

#### World composite based on optimal Sentinel-2 tiles



92,985 DOWNLOADED TILES WERE STORED ON THE JEODPP AND ATMOSPHERICALLY CORRECTED USING THE SEN2COR L2A PROCESSOR

#### Tile-based processing workflow



### Composite-based processing workflow





# **RESULTS & WAY FORWARD**



#### Results on 6,062 S2 tiles covering China



#### Confidence in built-up areas



# Area of Qingkou

199-38a

## Results Laho

GHSL\_S2 (2017-2018)



#### Performance assessment of the two workflows

Area of 6 000 000 km<sup>2</sup> Cluster: 16 processing nodes Total RAM: 256 GB CentosOS 6.9



|                           | Tile Based workflow                | Composite based workflow       |
|---------------------------|------------------------------------|--------------------------------|
| Input                     | 1865 S2 tiles (100 x 100 km tiles) | 276 blocks (150x150 km blocks) |
| Processing time           | 15 h                               | 12 h                           |
| Number of concurrent jobs | 10                                 | 2                              |
| RAM requirements per job  | 22 GB                              | 120 GB                         |

#### Quantitative assessment and comparison



#### Status and next steps



### Status and next steps

- Global Built-up (S1 & S2 based) production under the Copernicus Global Land Component WP2020
- Global Multipurpose Validation exercise
- Integration of GHSL workflows into DIAS (Web Advanced Space Developer Interface (WASDI)- ONDA DIAS)
- Delivery of free tools for built-up areas extraction from Big Earth Data (MASADA v.2 – to be released soon)



# Thank you



#### Christina.corban@ec.europa.eu http://ghsl.jrc.ec.europa.eu/

