COPERNICUS MASTER IN DIGITAL EARTH

With the support of the Erasmus+ Programme of the European Union

How reliable are SEN2 cloud detection algorithms?

Global uncertainty estimation using Deep Kernel Learning.

Cesar Luis Aybar Camacho

Erasmus Mundus Joint Master Degree Programme
Copernicus Master in Digital Earth
Specialization track GeoData Science

Vannes, France, 2022

Acknowledges

MINISTERIO DE CIENCIA E INNOVACIÓN

Université Bretagne Sud

THE LINUX FOUNDATION

sen2cube.at

Radiant MLHub
EARTH IMAGERY FOR IMPACT

Google Cloud

PARIS LODRON UNIVERSITÄT SALZBURG

R consortium

With the support of the Erasmus+ Programme of the European Union

COPERNICUS MASTER IN DIGITAL EARTH

Problem definition

```
// Map the function over one month of data and take the median.
// Load Sentinel-2 TOA reflectance data.
var dataset = ee.ImageCollection('COPERNICUS/S2')
                  .filterDate('2018-01-01', '2019-01-31')
                  // Pre-filter to get less cloudy granules.
                  .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 1));

var rgbVis = {
  min: 0.0,
  max: 0.3,
  bands: ['B4', 'B3', 'B2'],
};

Map.setCenter(-9.1695, 38.6917, 12);
Map.addLayer(dataset.median().divide(10000), rgbVis, 'RGB');
```

Valencia, Spain

Problem definition

```
// Map the function over one month of data and take the median.
// Load Sentinel-2 TOA reflectance data.
var dataset = ee.ImageCollection('COPERNICUS/S2')
                  .filterDate('2018-01-01', '2019-01-31')
                  // Pre-filter to get less cloudy granules.
                  .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 1));


var rgbVis = {
  min: 0.0,
  max: 0.3,
  bands: ['B4', 'B3', 'B2'],
};

Map.setCenter(-9.1695, 38.6917, 12);
Map.addLayer(dataset.median().divide(10000), rgbVis, 'RGB');
```

Iquitos, Peru

Problem definition

```
// Map the function over one month of data and take the median.
// Load Sentinel-2 TOA reflectance data.
var dataset = ee.ImageCollection('COPERNICUS/S2')
                  .filterDate('2018-01-01', '2019-01-31')
                  // Pre-filter to get less cloudy granules.
                  .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 1));

var rgbVis = {
  min: 0.0,
  max: 0.3,
  bands: ['B4', 'B3', 'B2'],
};

Map.setCenter(-9.1695, 38.6917, 12);
Map.addLayer(dataset.median().divide(10000), rgbVis, 'RGB');
```

Huancayo, Peru

Problem definition

Is it possible to predict cloud cover more accurately?

Problem definition

Is it possible to predict cloud cover more accurately?

How do cloud cover algorithms work?

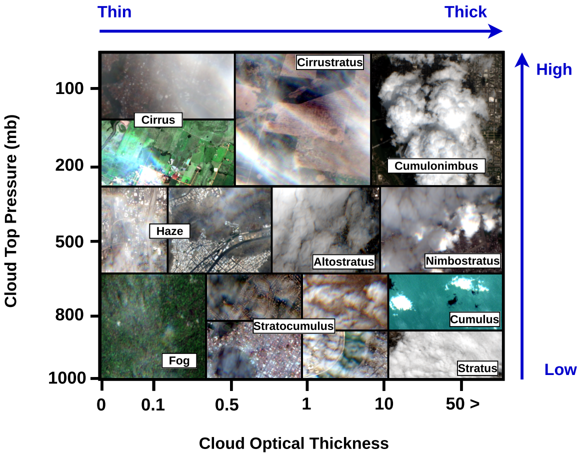
Problem definition

Is it possible to predict cloud cover more accurately?
How do cloud cover algorithms work?
What is a cloud?

What is a cloud?

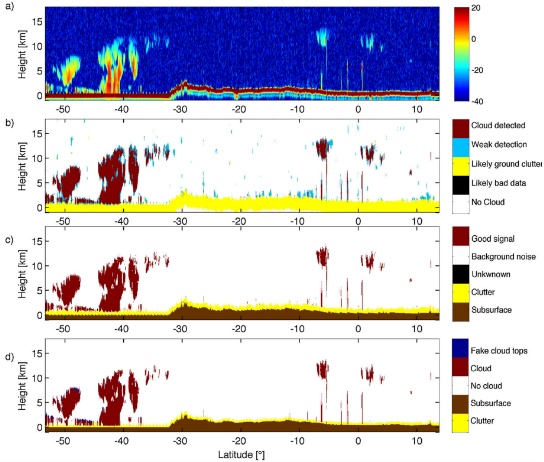
A cloud is a mass of water drops or ice crystals suspended in the atmosphere.

What is a cloud?



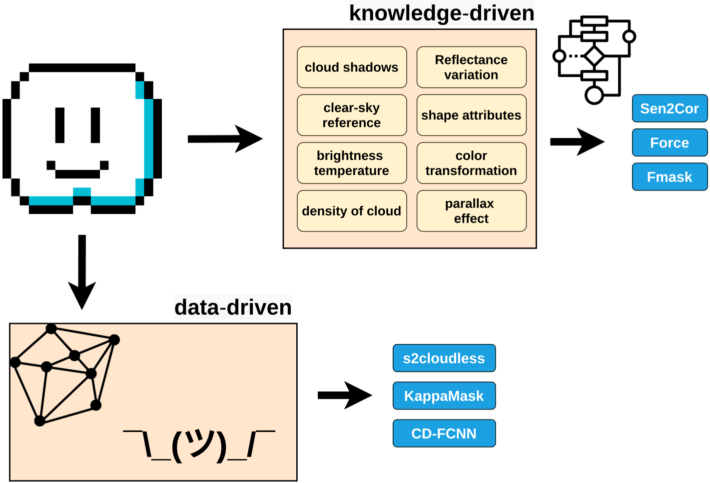
A cloud is a mass of water drops or ice crystals suspended in the atmosphere.

What is a cloud?

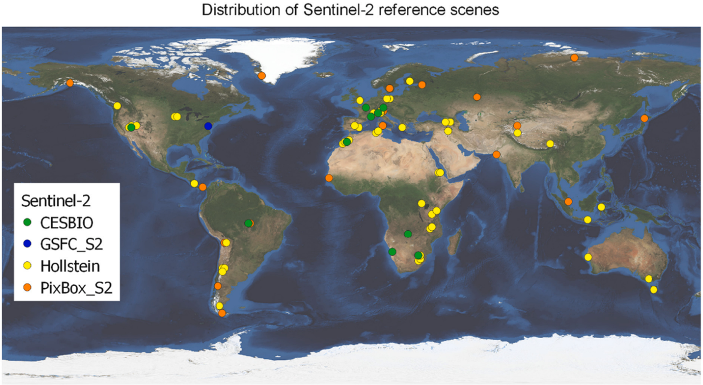


Cloud classification - CloudSat (Ceccaldi, et al. 2020)

How do cloud cover algorithms work?



Context



Geographical distribution reference cloud detection datasets for Sentinel-2 (Skakun et al. 2022).

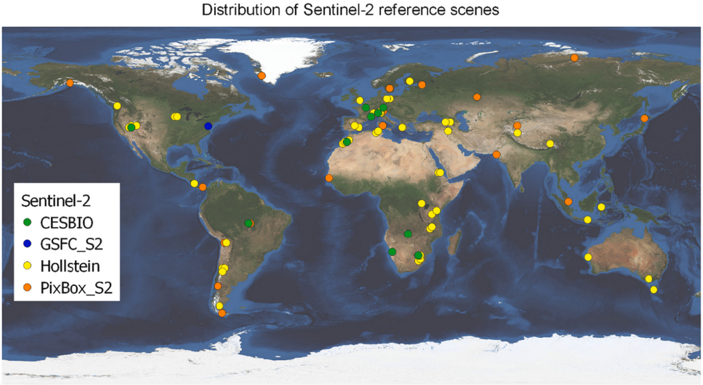
https://cloudsen12.github.io/

CloudSEN12 - Team ¡3

 Herrera Fernando	 Loja Jhomira	 Ysuhuaylas Luis	 Gonzales Karen
 LLactayo Valeria	 Bautista Lesly	 Diaz Lissette	 Flores Angie
 Cuenca Nicole	 Inga Joselyn	 Espinoza Wendy	 Esquivel Antony
 Fernando Prudencio	 Yali Roy	 Aybar Cesar	 Mateo-García Gonzalo
 Gomez-Chova Luis	 Tiede Dirk	 Sudmanns Martin	

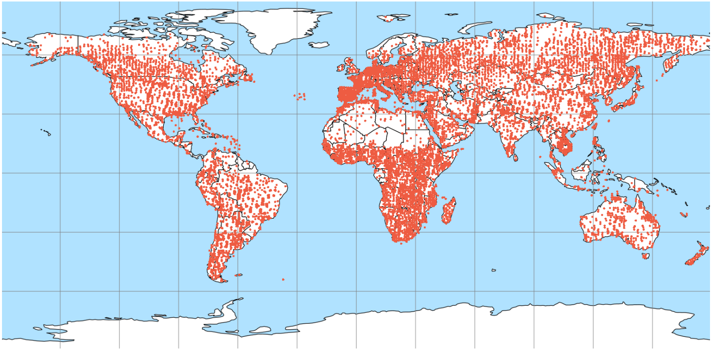
CloudSEN12 team

CloudSEN12 - Map



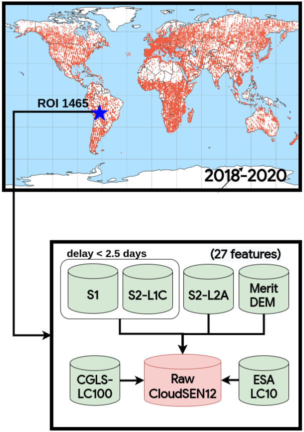
Geographical distribution reference cloud detection datasets for Sentinel-2 (Skakun et al. 2022).

CloudSEN12 - Map



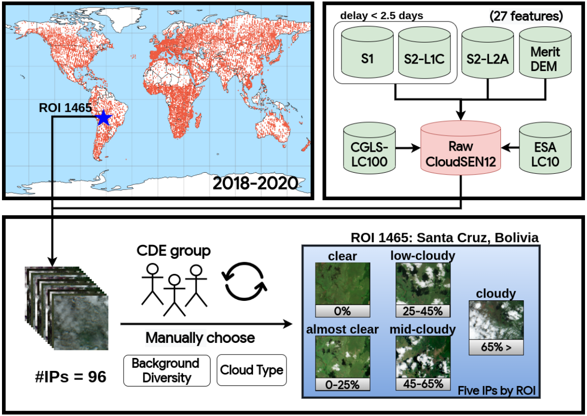
CloudSEN12 spatial distribution

CloudSEN12 - Data preparation



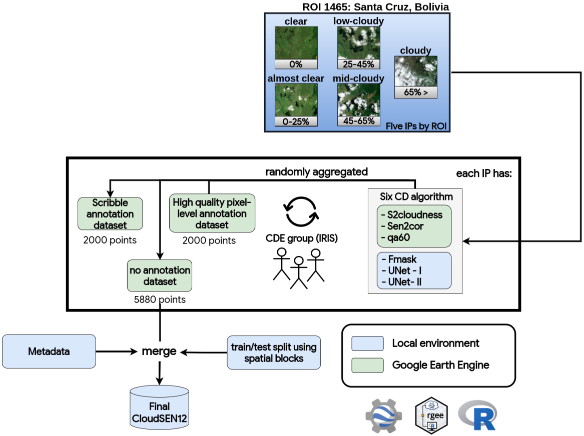
- Merging of different spatio-temporal datasets to use as predictors.
- Semi-automatic selection of ROIs (5090x5090 m.)

CloudSEN12 - Data Selection



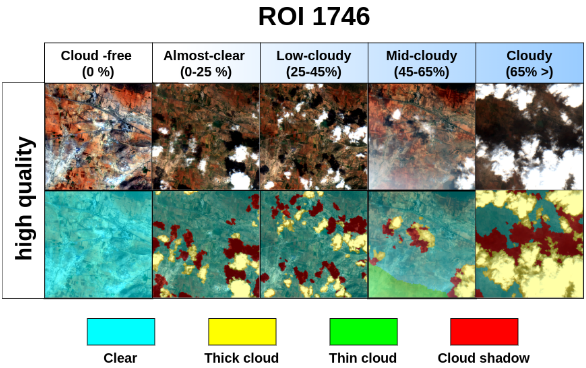
- Each ROI have multiple IPs. We manually select five IPs considering the background, cloud type, and cloud coverage.

CloudSEN12 - Labeling

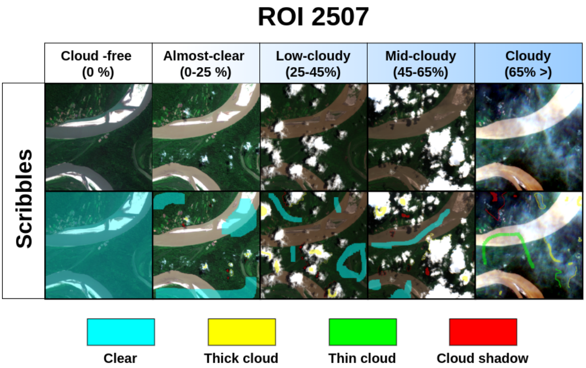


- Add to each IP the results of 6 different CD algorithms. In addition, each ROI counts with manual labeling that can be of type: high-quality, scribble, or no annotation.

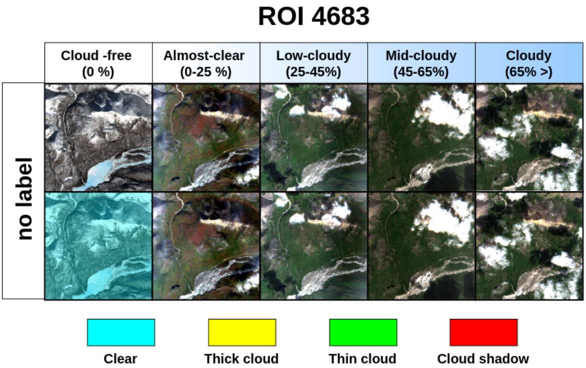
CloudSEN12 - Labels



CloudSEN12 - Labels



CloudSEN12 - Labels



Methodology

1. The **cloudSEN12 high-quality dataset** is used only for this experiment D; pixel-by-pixel cloud masking is converted to cloud cover percentages in order to perform a regression task.
2. The dataset of S2 L1C $\{x_1, x_2, ..., x_n\}$ and corresponding cloud cover percentages $\{y_1, y_2, ..., y_n\}$ is splitted in train (D_{train}), validation (D_{val}), and test (D_{test}) dataset.
3. **Experiment 01**: Convolutional Neural Network (CNN) for regression.

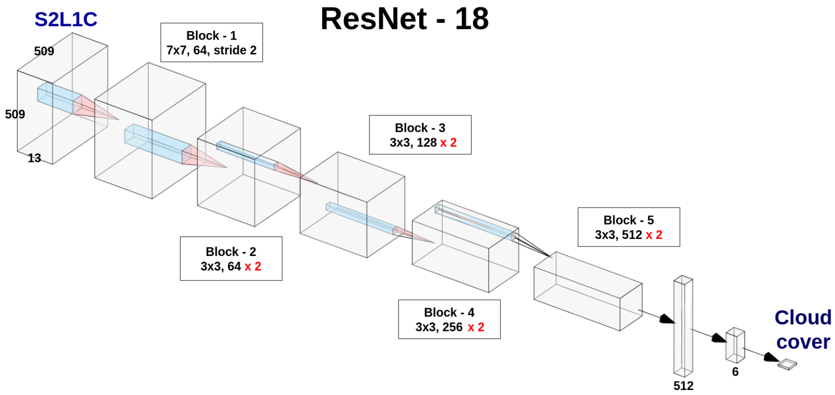
Methodology

1. The **cloudSEN12 high-quality dataset** is used only for this experiment D; pixel-by-pixel cloud masking is converted to cloud cover percentages in order to perform a regression task.
2. The dataset of S2 L1C $\{x_1, x_2, ..., x_n\}$ and corresponding cloud cover percentages $\{y_1, y_2, ..., y_n\}$ is splitted in train (D_{train}), validation (D_{val}), and test (D_{test}) dataset.
3. **Experiment 02**: CNN with a GP header layer.
   - Residual CNN $f_\theta : x \to R^J$ with feature space dimensionality J and parameters θ.
   - Exact GP header with parameters $\phi = \{l, s\}$ where l and s are kernel hyperparameters.

Methodology

1. The **cloudSEN12 high-quality dataset** is used only for this experiment D; pixel-by-pixel cloud masking is converted to cloud cover percentages in order to perform a regression task.
2. The dataset of S2 L1C $\{x_1, x_2, ..., x_n\}$ and corresponding cloud cover percentages $\{y_1, y_2, ..., y_n\}$ is splitted in train (D_{train}), validation (D_{val}), and test (D_{test}) dataset.
3. **Experiment 03**: Variational Deep Kernel Learning.
   - Residual NN $f_\theta : x \rightarrow R^J$ with feature space dimensionality J and parameters θ.
   - Approximate GP header with parameters $\phi = \{l, s, \omega\}$ where l and s are kernel hyperparameters, ω GP variational parameters (including m inducing point locations Z).
   - Using a random subset of p points of our training data, $X^{init} \subset X$, compute; **Initial inducing points:** Use found centroids as initial inducing point locations Z in GP. **Initial length scale:** set as 1.

Experiment 01



Experiment 02 - Gaussian processes

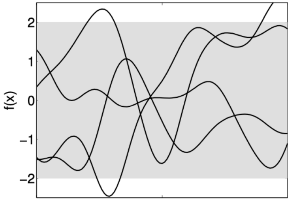
Definition

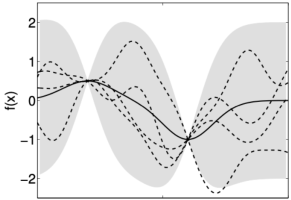
A Gaussian process (GP) is a collection of random variables, any finite number of which have a joint Gaussian distribution.

Nonparametric Regression Model Prior:

$f(x) \sim \mathcal{GP}\left(m(x), k\left(x, x'\right)\right)$, meaning $\left(f\left(x_1\right), \ldots, f\left(x_N\right)\right) \sim \mathcal{N}(\boldsymbol{\mu}, K)$, with $\mu_i = m\left(x_i\right)$ and $K_{ij} = \operatorname{cov}\left(f\left(x_i\right), f\left(x_j\right)\right) = k\left(x_i, x_j\right)$.

$$\overbrace{p(f(x) \mid \mathcal{D})}^{\text{GP posterior}} \propto \overbrace{p(\mathcal{D} \mid f(x))}^{\text{Likelihood}} \overbrace{p(f(x))}^{\text{GP prior}}$$

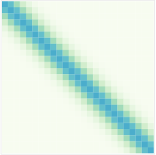




Experiment 02 - Gaussian processes

RBF KERNEL

$$\sigma^2 \exp\left(-\frac{||t-t'||^2}{2l^2}\right)$$

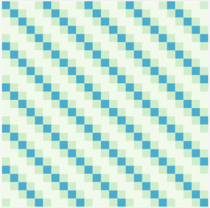


Variance σ = 0.8

Length l = 0.8

PERIODIC

$$\sigma^2 \exp\left(-\frac{2\sin^2(\pi|t-t'|/p)}{l^2}\right)$$



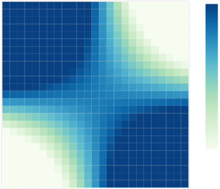
Variance σ = 0.8

Length l = 0.8

Periodicity p = 0.5

LINEAR

$$\sigma_b^2 + \sigma^2(t-c)(t'-c)$$



Variance σ = 0.3

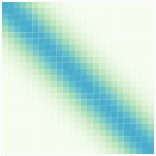
Variance σ_b = 0.8

Offset c = 0

Experiment 02 - Gaussian processes

RBF KERNEL

$$\sigma^2 \exp\left(-\frac{||t-t'||^2}{2l^2}\right)$$



Variance σ = 0.8

Length l = 1.58

PERIODIC

$$\sigma^2 \exp\left(-\frac{2\sin^2(\pi|t-t'|/p)}{l^2}\right)$$

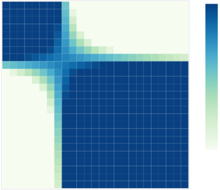
Variance σ = 0.4

Length l = 0.8

Periodicity p = 0.5

LINEAR

$$\sigma_b^2 + \sigma^2(t-c)(t'-c)$$

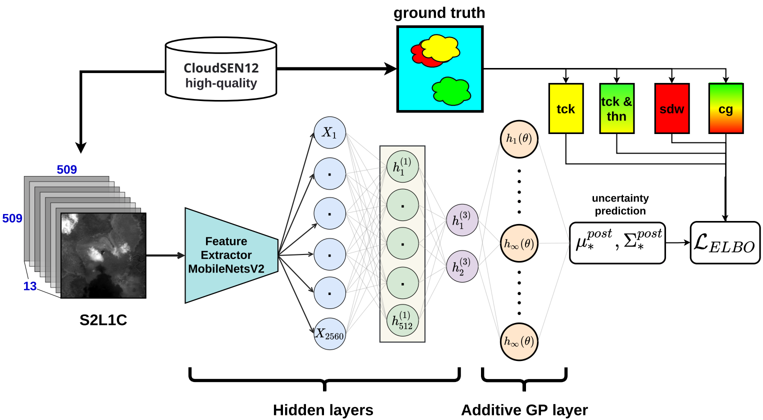


Variance σ = 0.67

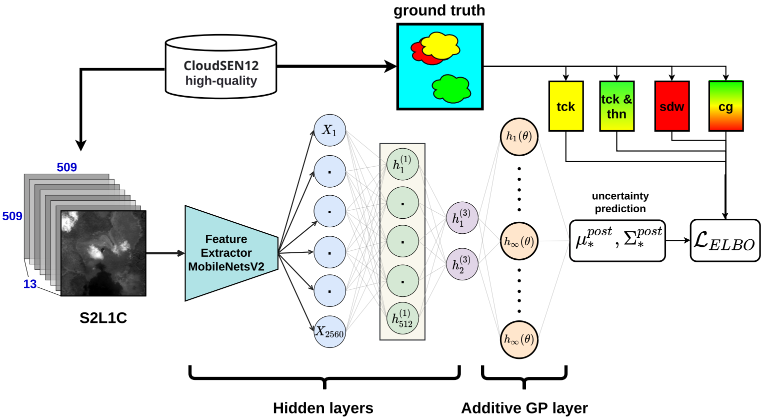
Variance σ_b = 0.8

Offset c = -2

Experiment 02 - Gaussian processes

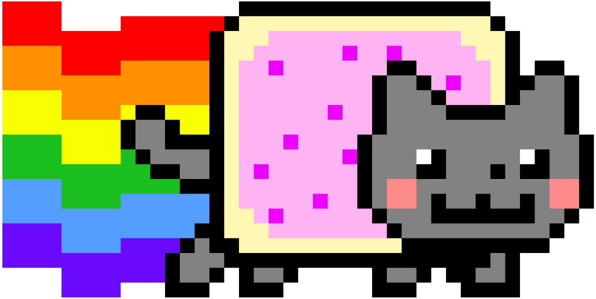


Experiment 03 - Stochastic Variational DKL



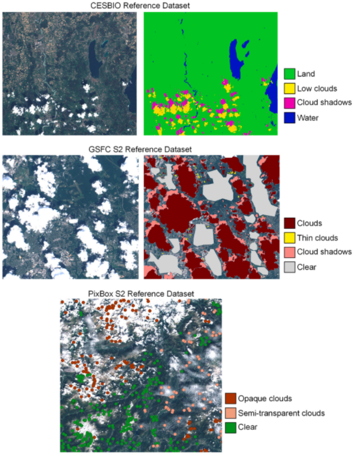
References

- Fomin FV, Grandoni F & Kratsch D, 2009, *A note on the complexity of minimum dominating set* , Journal of Discrete Algorithms, **4(2)**, pp. 209–214.
- Grobler PJP & Mynhardt CM, 2009, *Secure domination critical graphs*, Discrete Mathematics, **309**, pp. 5820–5827.
- Van Rooij JMM & Bodlaender HL, 2011, *Exact algorithms for dominating set*, Discrete Applied Mathematics, **159**, pp. 2147–2164.



Muchas gracias

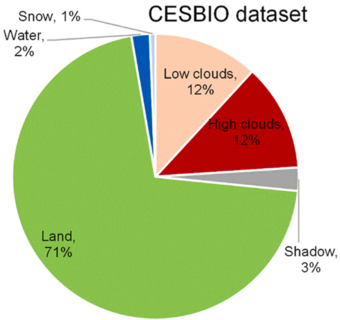
Context - II



Skakun et al. 2022

- Cloud labels created by human photo-interpretation and ground-based camaras.

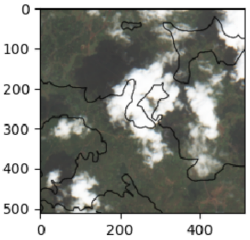
Context - II

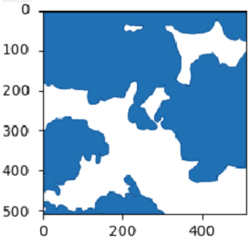


Skakun et al. 2022

- Cloud labels created by human photo-interpretation and ground-based camaras.
- High class imbalance.

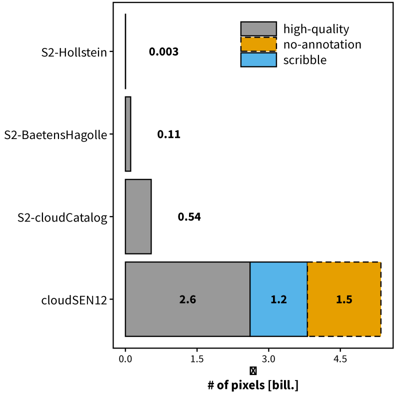
Context - II





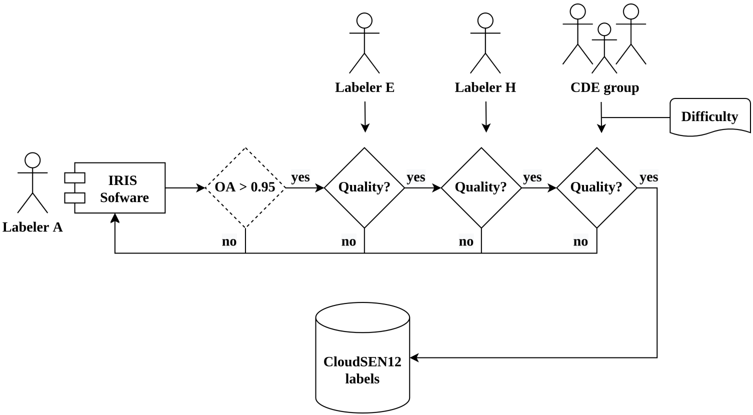
- Cloud labels created by human photo-interpretation, active learning and ground-based camaras.
- High class imbalance.
- **The quality of some datasets is poor.**

Context - II

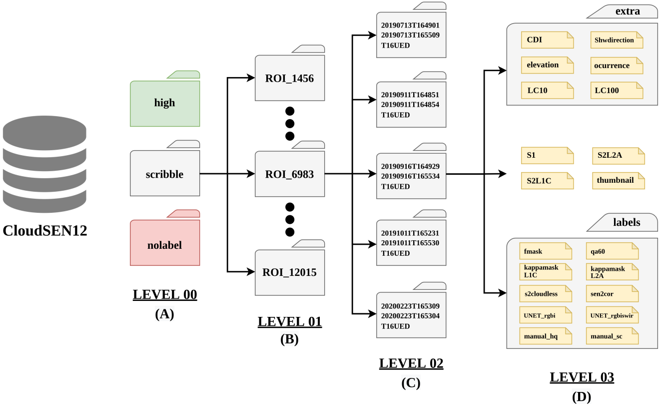


- Cloud labels created by human photo-interpretation and ground-based camaras.
- High class imbalance.
- **The quality of some datasets is poor.**
- Created by ***closed science practices***.
- **No temporal features.**

Quality Control



Dataset Structure



Technical validation