Cloud detection using SEVIRI IR channels for the GERB processing

Alessandro.Ipe@oma.be & Luis Gonzalez Sotelino

Royal Meteorological Institute of Belgium

Overview

Motivations

Constraints

Algorithm

Results

Comparisons

Further work

Motivations

- sceneID only relying on visible SEVIRI channels (for solar ADMs selection)
  - sunglint saturates channels over ocean
  - ▶ degraded cloud mask within sunglint area
- sceneID only provided during daytime
  - ▶ users' request for cloud mask during nighttime
  - ▶ temporarily addressed by including MPEF CLM within L20 products

Constraints

- GERB aim is to deliver climate records
- GERB products must remain stable

- Limited use of *uncontrolled* ancillary data
- Independence to NWP data

- Implementation of an IR cloud detection scheme instead of using MPEF or NWCSAF

Physics

- SEVIRI IR 8.7, 10.8 & 12.0 μm channels are most sensitive to clearsky & clouds
- Clouds are characterized by lower radiances (temperatures) than clearsky surfaces (warmer) except for snow & sea ice surfaces
- Aerosols are *generally* lowering IR radiances
- IR radiances are varying with viewing zenith angle, history (precipitation, cloud shadow) and state of atmosphere (profiles)

Assumptions

- Considering time–series of pixel–based BTs
- Temporal window for time–series set to 60 days
- Samples in time–series can be grouped into 3 classes:
  1. thick cold clouds (low BTs)
  2. thin or low clouds (high BTs)
  3. clearsky conditions (highest BTs)
- Tails of upper classes are overlapping
- No realtime ancillary data such as NWP fields

► Cannot be applied to snow & sea ice surfaces

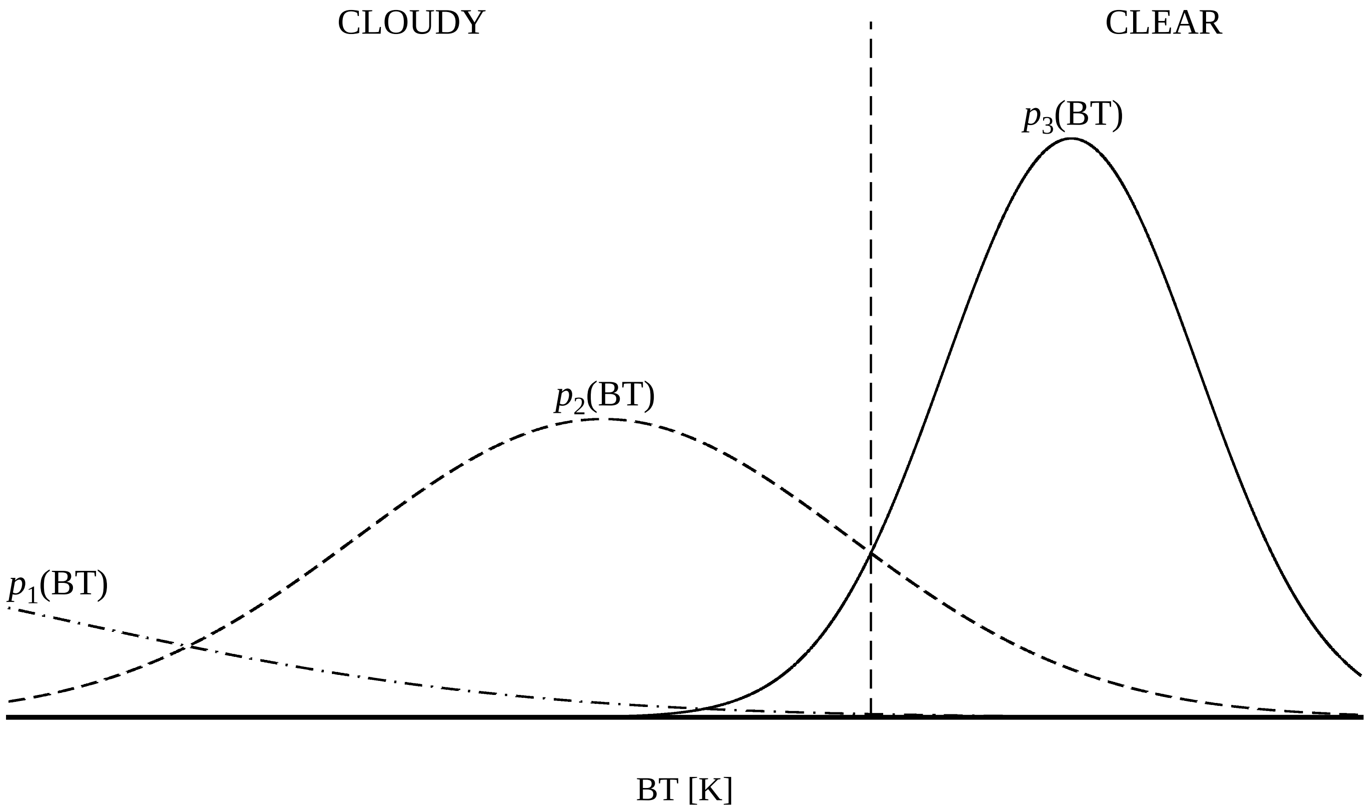
Scheme

- Perform a *modified* k–means clustering:
  1. Initialize the μ_n and σ_n for the 3 clusters
  2. If initialization fails goto step 1 with 2 clusters and so on…
  3. Classify all 60 BTs according to their nearest cluster with $d(T, \mu_n, \sigma_n)$
  4. Update μ_n and σ_n
  5. Repeat from step 3 until all μ_n do not significantly change ($\Delta\mu_n < 0.01$ K)

- Metric $d(T, \mu_n, \sigma_n) = (T - \mu_n)^2 / 2\sigma_n^2 + \ln \sigma_n^2$ if values in each class follow $p_n(T) = N(\mu_n, \sigma_n)$
- Initialization driven by physics (climatology)

Scheme

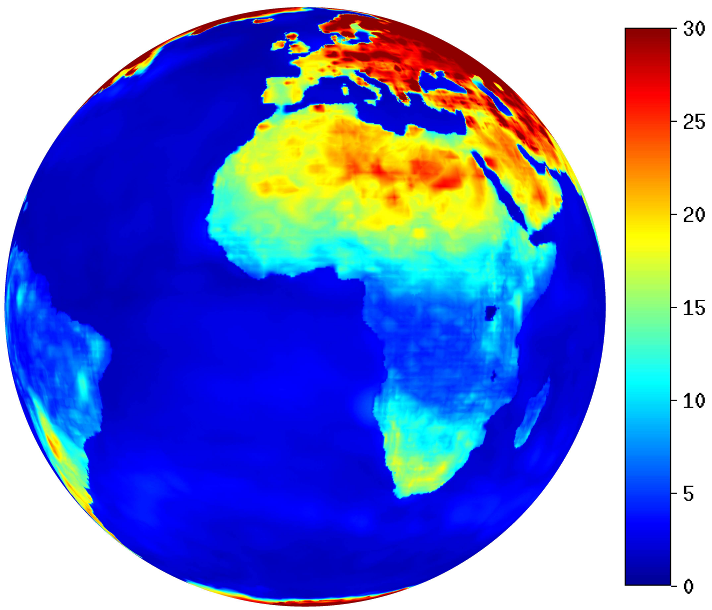
- Final classification (of the most recent sample):



- Can be seen as dynamical thresholding

Initialization

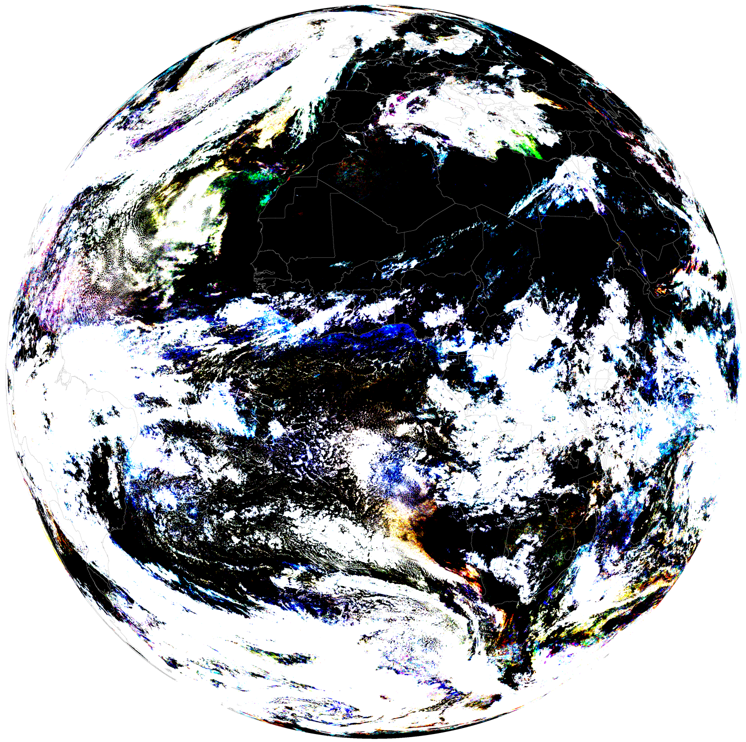
- Assume that clearsky class is Δ wide
- Δ is only needed for starting the clustering
- Δ is estimated from last 10 years of 6–hourly ERA–INTERIM surface skin temperatures



Δ [K] for March 15 at 0:00 UTC

Results

- High correlation between 8.7, 10.8 & 12 μm channels
  - Clustering separately applied to each channel



March 11 2007 at 0:00 UTC

Datasets

Multispectral threshold schemes:

- MPEF CLM: broadcast together with SEVIRI data and only for $\theta < 75^\circ$
- ▶ NWCSAF CMa: considered as truth

for hourly March 11–17 2007

- Both use ancillary NWP data
- Both use spatial texture filtering as post–processing

- Reprocessed SEVIRI data for GERB cloud mask: effective IR radiances (ED02)

Merging scheme

Geotype	Band [μm]	11	12	13	14	15	16	17	mean
ocean	8.7	84.47	86.47	86.38	85.96	85.77	86.13	86.38	85.94 ± 0.52
	10.8	84.50	86.49	86.39	85.86	85.87	86.33	86.51	85.99 ± 0.54
	12	83.36	85.52	85.57	84.74	84.80	85.37	85.32	84.95 ± 0.59
vegetation	8.7	88.49	87.93	88.34	88.19	86.49	86.70	85.44	87.36 ± 0.88
	10.8	89.06	88.66	88.98	89.16	87.63	87.93	86.60	88.28 ± 0.71
	12	89.45	89.10	89.42	89.83	88.43	88.84	87.33	88.91 ± 0.63
desert	8.7	94.35	94.60	95.19	95.03	91.84	90.85	89.28	93.00 ± 1.78
	10.8	95.19	95.50	95.82	95.62	93.11	92.52	91.40	94.15 ± 1.35
	12	95.59	95.82	96.20	95.90	93.76	93.46	92.75	94.77 ± 1.07

Weighted daily means according to the number of night pixels of the hourly pixels' agreement (in %) between NWCSAF CMa & GERB IR cloud masks

- No channel is statistically suitable for each surface type
  - ▶ Selection of the channels associated to the highest mean agreement and lowest uncertainties: 8.7 μm for ocean, 12 μm for land

Results

- Better than MPEF CLM for nighttime
- Worse for daytime since no use of visible bands
- Better anytime when supplemented with NWCSAF spatial filtering

Cloud mask	Geotype			
	ocean	vegetation	desert	all
MPEF CLM	85.73	88.70	91.19	87.20
GERB IR	85.94	88.91	94.77	87.84
GERB IR+	91.26	90.57	96.07	91.72

Nighttime

Cloud mask	Geotype			
	ocean	vegetation	desert	all
MPEF CLM	84.13	89.24	90.27	86.21
GERB IR	82.03	82.98	93.19	83.71
GERB IR+	88.01	85.87	94.06	88.26

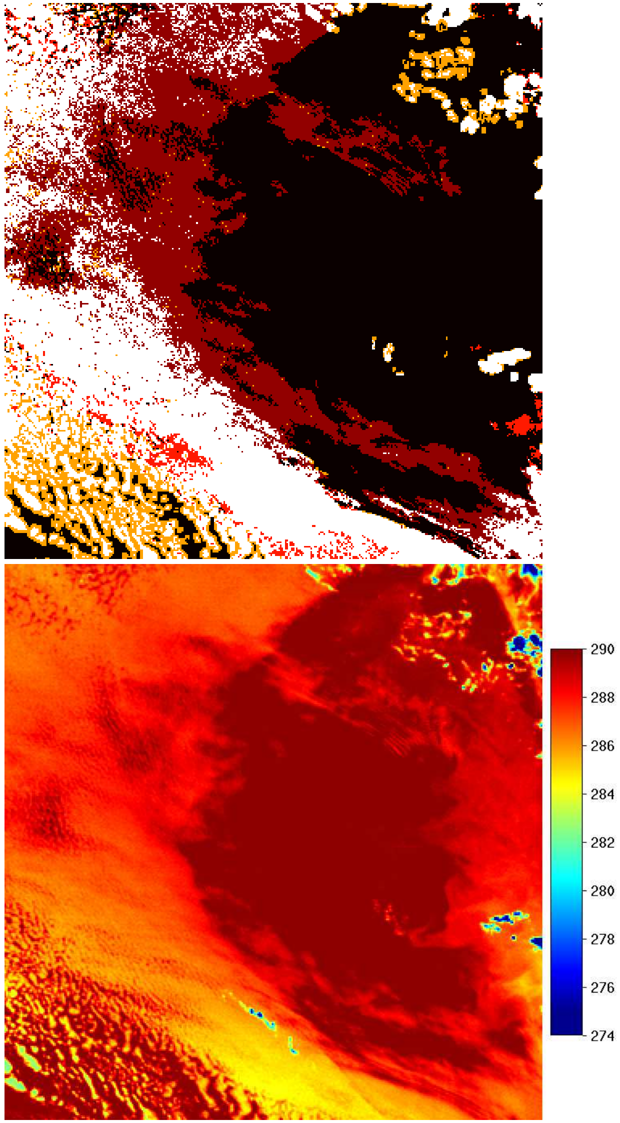
Daytime

Cloud mask	Geotype			
	ocean	vegetation	desert	all
MPEF CLM	84.93	88.97	90.73	86.70
GERB IR	83.98	85.94	93.99	85.78
GERB IR+	89.63	88.22	95.07	89.98

Alltime

Limitations

- Decrease of performance for $\theta > 70^\circ$ (limb darkening)
- Low warm clouds over ocean due to low BT contrast (≈ 1 K)
- Cloud edges in broken cloud fields



March 11 2007 at 00:00 UTC

Further work

- Low clouds over ocean detected with $BTD_{12-3.9} > 4.25$ K (night)
  - Use of 3.9 μm or $BTD_{12-3.9}$ in *temporal* clustering scheme ?
- Δ should vary according to θ:
  - SEVIRI BTs with NWCSAF cloud mask climatology

Cloud mask	Geotype			
	ocean	vegetation	desert	all
MPEF CLM	85.73	88.70	91.19	87.20
GERB IR	85.94	88.91	94.77	87.84
GERB IR+	91.26	90.57	96.07	91.72

Nighttime

Cloud mask	Geotype			
	ocean	vegetation	desert	all
MPEF CLM	85.73	88.70	91.19	87.20
GERB IR	86.34	88.91	94.77	88.09
GERB IR+	91.49	90.57	96.07	91.86

Nighttime with $BTD_{12-3.9}$ threshold test

- Length of time–series varying from pixel to pixel ?