

# At the confluence of forward and inverse remote sensing, AI and calval studies : The 2021 overview of the LMD

## (GEISA, ARSA, TIGR, ICO) databases

R. Armante, N. A. Scott (\*), J. Pernin, V. Capelle, L. Crépeau, C. Crevoisier, A. Chédin

(\*) : Corresponding author: [noelle.scott@lmd.ipsl.fr](mailto:noelle.scott@lmd.ipsl.fr) / Sea-related publications from the ARA/ABC(t) group at LMD: <http://ara.abct.lmd.polytechnique.fr/>



### Scientific update at LMD

#### GEISA in 2021

**LINE PARAMETERS SUB-DATABASE (2020 update)**

**20 molecules updated (12 new isotopologues)**  
**6 new molecules (15 isotopologues)**

**HONO, COFCl, CH<sub>3</sub>, CH<sub>3</sub>F, RuO<sub>4</sub>, H<sub>2</sub>C<sub>3</sub>H<sub>2</sub>**  
(Paper submitted to J. Mol. Spec., April 2021, Delahaye et al)

**Example of evaluation of O<sub>3</sub> with the SPARTE chain (armante et al, 2016)**

Collocations ARSA/IASI tropical/sea/night 2017

Distribution and tools on AERIS/ESPRI web site  
<https://geisa.aeris-data.fr>

Download, content information and graphical tools

### ARSA

LMD has elaborated the 41-year ARSA database, starting from observations by worldwide distributed radiosonde stations and combining them with surface and other auxiliary observations

From January 1979, and extended onwards on a monthly basis, ARSA contains 7 millions of severely quality controlled (P, T, H<sub>2</sub>O, O<sub>3</sub>) profiles on the 4A/OP 43 level pressure grid, from Surface to 0.026 hPa

**!! NEW VERSION AVAILABLE !!**

- The accuracy assessment of the ARSA database is based on the robust radiometric stability of IASI. Analysis of IASI residuals has led to improve the ERA-Interim vertical representation of the water vapor profile above 300hPa. (Confirmed by ERA-5 results)
- The ARSA database stability has been demonstrated not only on the IASI period but also over several years of operational NOAA satellites (NOAA 10, 11, 15) (Scott et al, 2015, Schröder et al, 2017).
- At the full vertical resolution of the radiosondes nominal reports from Surface to 400 hPa and on the 43 level pressure grid above;
- A newly implemented 8-airmass classification based on Pernin et al (2016) approach

**TO COME :**  
Extend + Update + Homogenize (ERA-5)

## Forward and Inverse Radiative Transfer, AI, Cal/Val activities

### TIGR

#### What are Thermodynamic Initial Guess Retrieval (TIGR) data sets ?

- Successive versions of size-limited collections of 3D representative atmospheric thermodynamic states, resulted from adequate sampling methods and classified into air mass classes, each situation described by T, H<sub>2</sub>O, O<sub>3</sub> vertical profiles and surface T/P.
- Associated with these profiles are Radiances, Transmittances, Jacobians TIGR data sets specifying the situations that satellite-based instrument (e.g., TOVS, ATOVS, AIRS, IASI, Modis, Sevir, IIR) would have observed under various in-flight conditions.
- Widely used by our scientific community (300 versions distributed worldwide).

See Talk 3.02, Pernin et al

#### A PROBABILISTIC POINT OF VIEW FOR CLASSIFICATION

1. Probabilistic input data : cumulative Distribution Functions (CDFs) to aggregate T/Tdp vertical profiles.
2. Probabilistic classification method : Gaussian Mixture Model (GMM) with the Expectation-Maximization (EM) algorithm to estimate the parameters of the mixture and the Maximum A Posteriori (MAP) to assign each situation to its nearest class

Unsupervised mode on a training data set then supervised mode

Pernin J., Vrac M., Crevoisier C. and Chédin A. : « Mixture model-based atmospheric air mass classification: a probabilistic view of thermodynamic profiles. », *Adv. Stat. Clim. Meteorol. Oceanogr.*, 2, 115–136, doi:10.5194/asclmo-2-115-2016 (2016)

#### IMPACT OF THE NEW AIRMASS CLASSIFICATION

Differences in the classification of ERA-Interim using both TIGR-2000 (top) and TIGR-2020 (bottom): January (left) and July (right) 2008

5-class supervised classification of ECMWF reanalyses, after unsupervised multivariate GMM on TIGR-2000

8-class supervised classification of ERA-Interim reanalyses, after unsupervised multivariate GMM on training dataset ERA-Interim reanalyses { 00h + 12h UTC of 15th of January, April, July, October between 2005 and 2009 }

### ICO

#### ICO (IASI Composite Observatory) : A reference dataset from space+time collocated and quality controlled satellite and non-satellite data to help identifying gaps and performing homogenization, accordingly.

So far, a 2-year demo dataset of global, day/night, land/sea, clear/cloudy MetOp (IASI, HIRS, AMSU, MHS) satellite data and auxiliary data is available.

- Analyzed RadioSoundings Archive (ARSA).
- Surface characteristics temperature and emissivity. (from Capelle et al)
- LMD Cloud/Aerosols/Clear Flag.
- GEISA-2015.
- Automatized Atmospheric Absorption Atlas (4A) forward model.

At LMD, we have verified that such a dataset is of paramount importance for the L1 to L3 satellite data QC, for feedbacks to Spectroscopic databases (e.g. GEISA) or for improved descriptions of water vapor or ozone, ..., profiles. As such, ICO appears to be a major contributor to GSICS, GEWEX G-Vap, GAIA-Clim activities.

Planned work:

- Update and expand the demo dataset with the newly available auxiliary models and databases (4A/OP, ARSA, GEISA-2020)
- Refine of the CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O, ..., mixing ratios distribution WRT time and geolocation of MetOp pixels
- Make ICO available to the International Community through AERIS data Services.