



**HAL**  
open science

## Preparing the assimilation of the future MTG-IRS sounder into the mesoscale NWP AROME model

Olivier Coopmann, N. Fourrié, P. Chambon, J. Vidot, P. Brousseau, M.  
Martet, C. Birman

► **To cite this version:**

Olivier Coopmann, N. Fourrié, P. Chambon, J. Vidot, P. Brousseau, et al.. Preparing the assimilation of the future MTG-IRS sounder into the mesoscale NWP AROME model. Quarterly Journal of the Royal Meteorological Society, In press, 10.1002/qj.4548 . hal-04190181

**HAL Id: hal-04190181**

**<https://hal.science/hal-04190181v1>**

Submitted on 29 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



---

# Preparing the assimilation of the future MTG-IRS sounder into the mesoscale NWP AROME model

O. Coopmann\*, N. Fourri , P. Chambon, J. Vidot, P. Brousseau, M. Martet and C. Birman  
*CNRM, Universit  de Toulouse, M t o-France and CNRS, Toulouse, France*

\*Correspondence to: O. Coopmann, CNRM/GMAP, 42 Avenue Gaspard Coriolis, 31057 Toulouse cedex, France.  
E-mail: olivier.coopmann@meteo.fr

---

The IRS (InfraRed Sounder) instrument is an infrared Fourier transform spectrometer that will be on board the Meteosat Third Generation series of the future EUMETSAT geostationary satellites. It will measure the radiance emitted by the Earth at the top of the atmosphere using 1960 channels. IRS will provide high spatial and temporal frequency 4D information on atmospheric temperature and humidity, winds, clouds, surfaces, as well as on the chemical composition of the atmosphere. The assimilation of these new observations represents a great challenge and opportunity for the improvement of Numerical Weather Prediction (NWP) forecast skill, especially for mesoscale models such as AROME at M t o-France (Brousseau et al. 2016). The objectives of this study are to prepare for the assimilation of IRS in this system and to evaluate its impact on the forecasts when added to the currently assimilated observations.

By using an Observing System Simulation Experiment (OSSE) constructed for a mesoscale NWP model. This OSSE framework makes use of synthetic observations of both IRS and the currently assimilated observing systems in AROME, constructed from a known and realistic state of the atmosphere. The latter, called the Nature Run, is derived from a long and uninterrupted forecast of the mesoscale model. These observations were assimilated and evaluated using a 1 h update cycle 3D-Var data assimilation system over two-month periods, one in the summer and one in the winter.

This study demonstrates the benefits that can be expected from the assimilation of IRS observations into AROME NWP system. The assimilation of only 75 channels over oceans increases the total amount of observations used in the AROME 3D-Var by about 50 %. The IRS impact in terms of forecast scores was evaluated and compared for the summer and winter periods. The main findings are that (i) over both periods the assimilation of these observations lead to statistically improved forecasts over the whole atmospheric column, (ii) for the summer season experiment, the forecast ranges up to +48 h are improved, (iii) for the winter season experiment, the impact on the forecasts is globally positive but is smaller compared to the summer period and extends only to 24 h. Based on these results, it is foreseen that the addition of future IRS observations in the AROME NWP systems will significantly improve mesoscale weather forecasts.

*Key Words:* MTG-IRS, OSSE, Numerical Weather Prediction, Radiative Transfer Model, Data Assimilation, Mesoscale model, Hyperspectral Infrared sounder, Geostationary satellite

*Received ...*

## 1. Introduction

The emergence of infrared sounders such as the Atmospheric InfraRed Sounder (AIRS), Infrared Atmospheric Sounding Interferometer (IASI), Cross-track Infrared Sounder (CrIS), onboard polar or low earth orbit (LEO) satellites, has led to

tremendous developments and improvements in the quality of forecasts of NWP (Numerical Weather Prediction) models. Because such data provide information on atmospheric temperature and water vapour at high vertical resolution (McCarty et al. 2021). However, these observations have a limited coverage in space and time because they are only onboard a few

platforms in sun-synchronous orbits.

The evolution of the spatial and temporal resolutions of the models in meteorological centres is mirrored by new observational requirements. This is especially true for mesoscale models such as AROME (Applications de la Recherche à l'Opérationnel à Méso-Echelle) at Météo-France (Seity *et al.* 2011), which is now characterized by a horizontal resolution of 1.3 km (Brousseau *et al.* 2016). This fine-scale NWP model used in operations since 2008 was designed to improve the short-term forecasting of extreme events such as heavy Mediterranean rainfall, severe thunderstorms, and fog or urban heat islands during heat waves. The majority of the data assimilated for AROME forecasts are provided by ground-based radars (Martet *et al.* 2022) and conventional data. These data provide very useful information for the model (Fourrié *et al.* 2015), however they do not cover the maritime areas of the AROME geographical domain such as the Atlantic Ocean, the North and Mediterranean Seas. AROME also makes use of SEVIRI data from the Meteosat geostationary satellite but its limited number of channels provide a rather small amount of information in the vertical; and as mentioned above the hyperspectral infrared sounders onboard LEO satellites are also used but their regional coverage is rather limited. Overall, the AROME initial conditions are under-constrained, with  $2.0 \times 10^9$  degrees of freedom but  $1.6 \times 10^7$  observations used in each analysis. To overcome this issue, a significant increase in the number of assimilated sounding data would be required.

Orbiting an infrared sounder onboard a geostationary satellite (GEO) could satisfy this requirement. The China Meteorological Administration (CMA) was the first to launch two hyperspectral sounders into geostationary orbit aboard FY-4A satellites over East Asia (the first in December 2016): the Geostationary Interferometric InfraRed Sounder (GIIRS-1 and 2) (Yang *et al.* 2017; Yin *et al.* 2020). Within the framework of this project, research conducted by (Burrows 2019) on GIIRS, has shown encouraging signs of getting impact from real hyperspectral GEO data in the European Centre for Medium-Range Weather Forecasts (ECMWF) system. The European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) proposes in its future geostationary satellite programme, the Meteosat Third Generation series, to launch two types of instruments on board: an imaging platform (MTG-I) for 2022 and a sounding platform (MTG-S) with the hyperspectral InfraRed Sounder (IRS) for 2024 (Holmlund *et al.* 2021). Finally, the Japan Meteorological Agency (JMA) is considering launching a follow-up to the Himawari programme including a hyperspectral infrared sounder from geostationary orbit in 2029 (Okamoto *et al.* 2020; Bessho *et al.* 2021) and the USA has announced the GeoXO (Geostationary eXtended Observations) programme which may include hyperspectral infrared observations.

In the framework of the EUMETSAT programme and onboard the MTG-S satellite, the InfraRed Sounder (IRS) is a Fourier Transform Spectrometer with a spectral sampling of  $0.6 \text{ cm}^{-1}$  that will observe the earth with 1960 channels in 2 spectral bands: the long-wave between ( $680 - 1210 \text{ cm}^{-1}$ ) and the mid-wave between ( $1600 - 2250 \text{ cm}^{-1}$ ). It will measure the entire Earth's disk centred at longitude  $0^\circ$  with a high temporal frequency and a particular focus on Europe (revisited every 30 minutes) and a spatial resolution of 4 km at nadir. All the characteristics of the IRS instrument are presented by (Holmlund *et al.* 2021). A major benefit is expected from IRS for NWP and nowcasting in terms of geophysical field prediction, cloud detection and surface characterization. The high spatial and temporal frequency of the IRS observations will allow a significant progress in the

forecasting of thunderstorm trajectories (in association with MTG-I and FCI), as well as of rainfall intensity. IRS will fill an important data gap in the Mediterranean, giving the opportunity to be more efficient in forecasting intense rainfall events in the Mediterranean area. The Copernicus will provide novel and valuable observations for atmospheric composition, air quality and climate monitoring applications.

The studies conducted by Guedj *et al.* (2014) and Duruisseau *et al.* (2017) have shown the potential benefits of using IRS in mesoscale models and pave the way for the assimilation of IRS observations into the AROME NWP model. The Observing System Simulation Experiment (OSSE) method (Wang *et al.* 2013; Privé *et al.* 2014; Atlas *et al.* 2015; Hoffman and Atlas 2016; Ma *et al.* 2015) used in this study provides a robust framework for the simulation of IRS observations as well as all the other observations currently used in AROME. Here only IRS radiances are assimilated, no products such as AMVs from IRS have been considered. We have focused on the direct impact of the assimilation of IRS radiances in AROME and not the impact of the assimilation in the coupler model. This work also allows the 3D-Var data assimilation system to be prepared for the use of IRS in terms of reconstructing observations from principal components, estimating the observation error and tuning the cloud detection algorithm.

This article describes all the steps of the study as follows. The general framework of the OSSE, the construction of the observations and the quality control are detailed in Section 2. The implementation of the data assimilation experiments is described in Section 3. The results of the impact of the IRS are presented in Section 4. A summary and conclusions, including a discussion of the limitations of this study are given in Section 5.

## 2. Observing System Simulation Experiment design

### 2.1. Overview of the method

Preparing for the arrival of new satellite observations is essential in the evolution and improvement of NWP models. However, the novelty of these data requires a rigorous preparation of the assimilation systems and a careful evaluation of their impact on weather forecasts. But these future observations do not yet exist. A way has therefore been developed in several NWP centres to solve this problem. This is the OSSE method, which effectively tests the impact of future observation systems on analyses and forecasts.

A typical OSSE system assimilates synthetic observations simulated from an atmospheric model state assumed to be the truth, and then assesses the quality of weather analyses and forecasts, using that truth for verification. An OSSE framework includes the following elements:

- Simulation of atmospheric state or "Nature Run".
- Synthetic observations.
- A NWP model and data assimilation system to generate analyses and forecasts.

The aim is to obtain a simulation as consistent as possible with reality. To achieve this, each of the components of the OSSE must be as realistic as possible. Extensive development and validation of OSSE experiments have been carried out at NASA GMAO and NCEP (e.g. Errico *et al.* 2007; Masutani *et al.* 2010b; McCarty *et al.* 2012; Errico *et al.* 2013; Privé *et al.* 2013b; Privé *et al.* 2013c; Boukabara *et al.* 2016). This research has led to recommendations for future OSSEs. The following sections

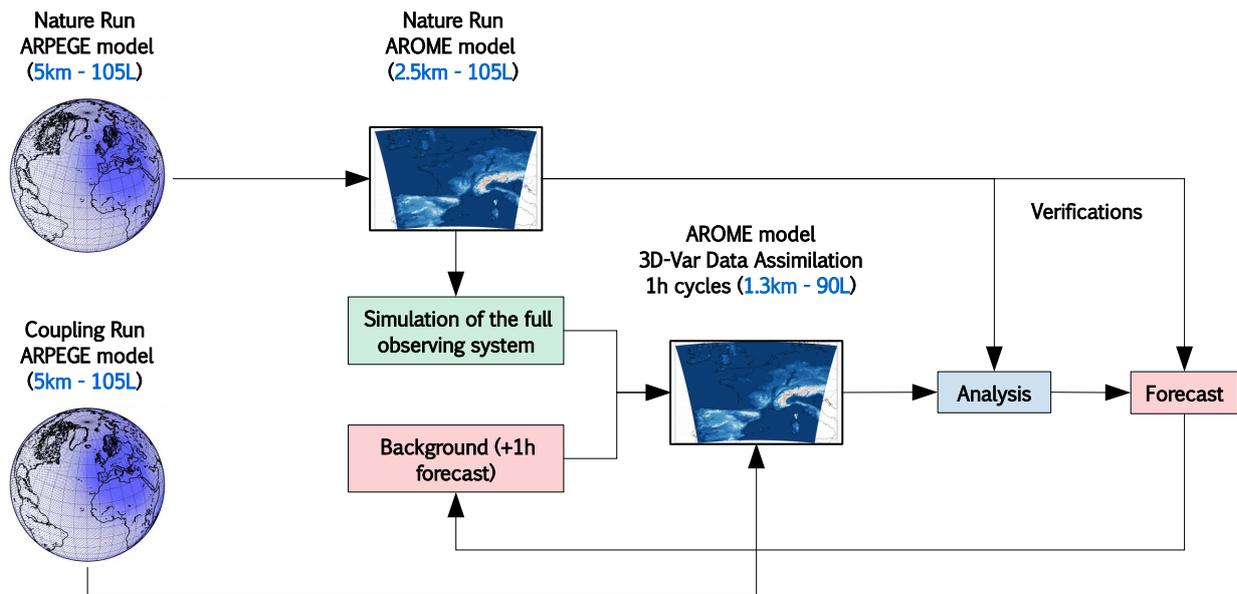


Figure 1. Observing system simulation experiment scheme used for this study.

describe the different components of the OSSE set up for this study, in line with the recommendations mentioned above.

Figure 1 describes the OSSE scheme used in this study. Three Nature Runs (NR) were produced: (i) a NR of the global ARPEGE (Action de Recherche Petite Echelle Grande Echelle) model (Courtier et al. 1991) to provide boundary conditions to (ii) a NR of the mesoscale AROME model used to simulate the current observing system and IRS observations and (iii) an ARPEGE Coupling Run (CR), slightly different from the global ARPEGE NR which provides boundary conditions to the AROME 3D-Var data assimilation system.

## 2.2. ARPEGE global model Nature Run

The ARPEGE Nature Run described here will not be used as the primary model for the simulation of synthetic observations, but will be used to initialise and provide the boundary conditions to the AROME Nature Run used in this OSSE.

For this global Nature Run, we have chosen the high-resolution ARPEGE NWP model based on a version operational since July 2019. It is a spectral model with a variable horizontal resolution  $T_L 1798$  (spectral truncation) C2.2 (stretch coefficient), of 5 km over France and 25 km over New Zealand. It has 105 vertical levels (from 10 m above the model surface to 0.1 hPa) (Bouysselet et al. 2022). This ARPEGE NR is design without assimilation, as an uninterrupted forecasts ; it was process for the summer (JJA) and winter (DJF) periods, each lasting 3 months, including one month of spin-up. The spin-up time represents the time needed for the simulation model to approach its own climatology after being launched from given initial conditions. The NR used here is initialized by the operational 4D-Var ARPEGE analysis of 01 December 2019 and 01 June 2020. In order to have a realistic ocean forcing, the lower boundary conditions over oceans are specified by a daily sea surface temperature (SST) forcing from OSTIA (Operational Sea Surface Temperature and Sea Ice Analysis) analyses (Stark et al. 2007).

## 2.3. AROME mesoscale model Nature Run

To conduct this OSSE, we create an AROME NR from a previous operational version of the AROME model. The model domain covers Western Europe with a horizontal resolution of 2.5 km and 90 vertical levels (Seity et al. 2011). This AROME NR is initialized and coupled to the lateral boundary conditions (LBC) of the ARPEGE NR beginning on 01 January and July 2020 and continuing uninterrupted for a 2-month period each. We used in the AROME NR the land surface and sea surface temperature fields provided by the NR ARPEGE. This AROME NR will be used to simulate the current full observing system including radars as well as the IRS. These infrared sounders are sensitive and scan the atmosphere also above 10 hPa. We adapted the AROME NR to obtain a more fine discretization above 10 hPa. To do this, we interpolated the last 15 levels of ARPEGE to specified levels of AROME. This configuration of the AROME NR simulates the meteorological parameters more accurately thanks to its additional levels, thus allowing a better simulation of the observations.

The ARPEGE NR and the AROME NR have been compared against the operational models to ensure the daily and average consistency of the different atmospheric fields (not shown). Average and standard deviation were used to evaluate temperature, humidity and wind parameters at the surface and in the atmospheric vertical as well as the proportion of cloud cover and precipitation accumulation.

## 2.4. ARPEGE global model Coupling Run

During the cycling of the 3D-Var data assimilation experiments, the AROME model will also need to be initialised and forced on its lateral boundaries. However, using the same ARPEGE NR to couple the AROME NR and the AROME 3D-Var data assimilation could lead to unrealistically small forecast errors, in particular for synoptic weather situations, driven by meteorological conditions outside the limited area domain. Using different lateral boundary conditions for the AROME NR and 3D-Var AROME would help to avoid the so-called "identical twin" problem (Masutani et al. 2010a) leading to

an overestimation of the impact of the observing system. This problem appears when the model behavior used for the data assimilation is too similar to the "truth" (the Nature Run) and when the same model is used for both the Nature Run and the data assimilation system. To avoid this problem, we have created a different atmospheric state from the ARPEGE NR which serves as the LBC for the AROME NR. We called it Coupling Run (CR) and it is initially configured on the same model as the ARPEGE NR except for. Two modifications; (i) the CR was run on the new Météo-France computer using a new compiler and new libraries bringing a slight noise to the forecasts, (ii) the CR was initialized each day by the +48 h forecasts of the ARPEGE NR valid the same day. This means that for a given day of the ARPEGE NR, its initial condition comes from a +P24 h of the previous day, while for this same day of the ARPEGE CR, its initial condition comes from a +P48 h of the ARPEGE NR without initialization of the SST in the middle of this forecast. This allows to introduce a slightly different behaviour between the atmospheric fields of the CR and NR ARPEGE. Thus, the ARPEGE NR will only be used to provide the LBCs of the AROME NR and the ARPEGE CR will only be used to provide the LBCs of the AROME 3D-Var data assimilation system.

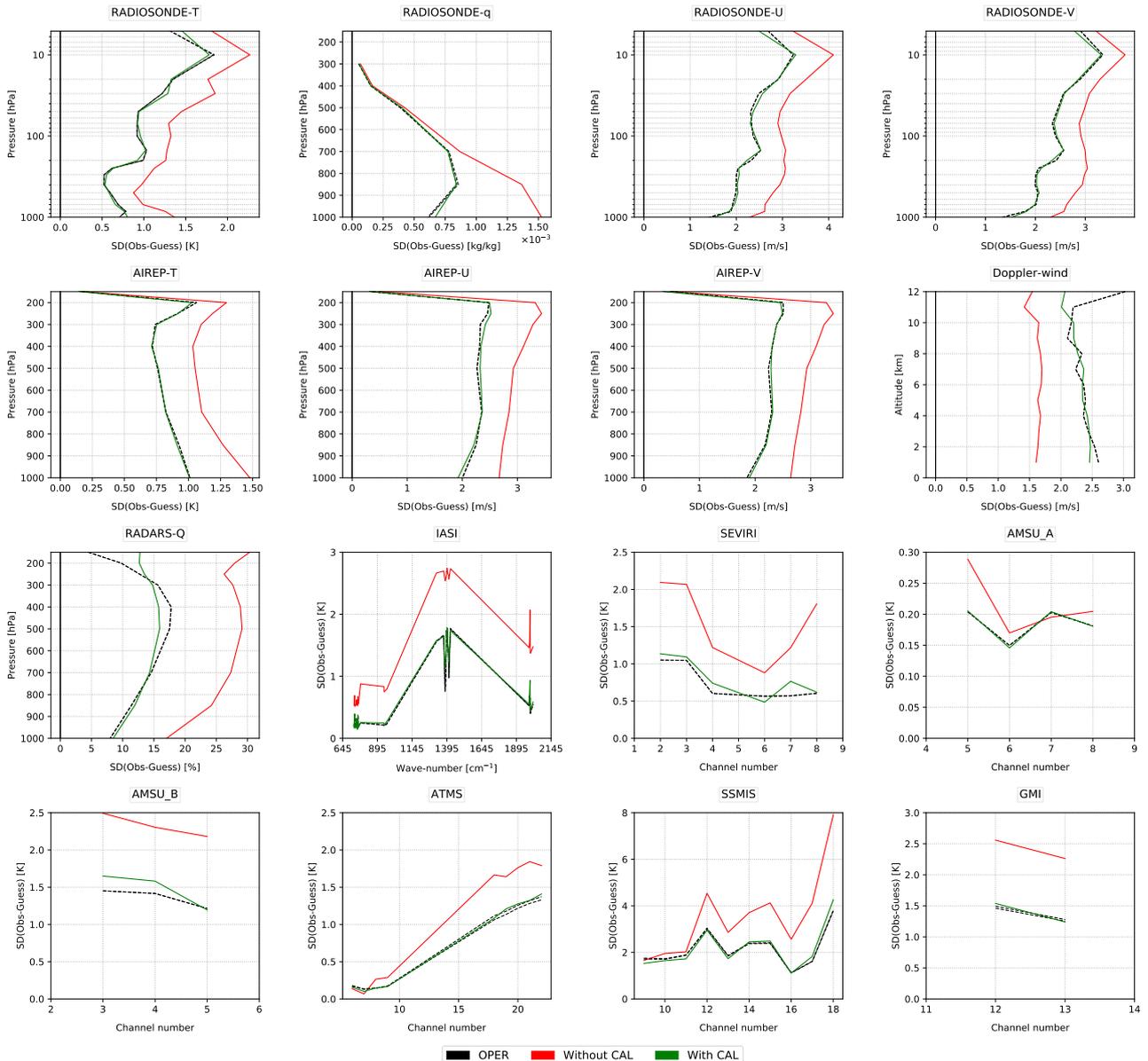
## 2.5. Data assimilation system

To perform the assimilation experiments in our OSSE framework, we used the 3D-Var data assimilation system AROME in its operational version for our study periods (January-February and July-August 2020). The operational cycle at that time was (cy43t2) on a model configuration with a horizontal resolution of 1.3 km and 90 vertical levels (from the surface to 10 hPa).

The AROME forecasts are initialized by analyses from the 3D-Var data assimilation system with a 1 h cycling (Brousseau et al. 2016). This assimilation frequency is particularly advantageous for the use of geostationary observations. Lateral boundary conditions, land surface and sea surface temperature fields are provided here by the ARPEGE CR, in order to avoid the twin problems mentioned above. Finally, the observations considered for the 3D-Var will come from synthetic observations processed and calibrated for these periods as described below.

## 2.6. Simulation and calibration of current observations

Initially, the objective is to simulate the full observing system currently assimilated in the 3D-Var AROME data assimilation



**Figure 2.** Standard deviations of the first-guess departures (SDFGD) of the main conventional and satellite observations for the operational AROME (black dashed line), the simulation experiment without calibration (red line) and with calibration (green line). The statistics were computed for assimilated observations over 44 days including 22 days of the summer period and 22 days of the winter period. Each set contains more than 10000 observations.

Type of measurement	Instrument
Surface measurements	Surface stations, ships, buoys
	Ground based Global Positioning System (GPS)
Altitude measurements	Radiosondes
	Aircraft measurements
	Wind profilers
Radar measurements	Doppler winds
	Humidity pseudo-observations
Satellite measurements	Scatterometer winds
	Atmospheric motion vectors (AMVs)
	Advanced Microwave Sounding Unit (AMSU)-A
	AMSU-B/Microwave Humidity sounder (MHS)
	Advanced Technology Microwave Sounder (ATMS)
	Special Sensor Microwave Imager/Sensor (SSM/I-S)
	GPM Microwave Imager (GMI)
Infrared Atmospheric Sounding Instrument (IASI)	
Spinning Enhanced Visible and InfraRed imager (SEVIRI)	

Table 1. List of observing systems currently assimilated in the 3D-Var AROME.

system. A wide variety of measurements are used, which can be divided into two general categories: conventional observations and satellite observations. These are listed in Table 1. In particular, as for the operational AROME NWP system, we use the Doppler winds and reflectivities from ground radars in this study. The latter observations are first transformed into a relative humidity profile with a 1D-Bayesian approach and then assimilated into a 3D-Var (Wattrelot et al. 2014).

Simulation and calibration of observations within the OSSE framework are essential steps to get as close as possible to what would happen in reality. The process is iterative and requires fine-tuning to ensure that the simulated system is similar to the operational system in order to reproduce the real impact of different types of observations. To simulate all the satellite observations assimilated in AROME, we used the RTTOV V11 radiative transfer model (<https://nwp-saf.eumetsat.int>; accessed 01 September 2022). To generate the synthetic observations, several steps were necessary. The first one consists in using the observation operators of the operational AROME 3D-Var system and applying them to the atmospheric fields of the AROME NR in order to simulate the observations. A random error is added to these simulated observations, i.e. a random and uncorrelated Gaussian perturbation is applied to the observation error. This implies that the synthetic observations of the current AROME observation system thus created are unbiased. It is therefore these added observation errors that are calibrated in the following. At the beginning of the calibration, the estimated observation errors are those specified in the AROME 3D-Var system.

The calibration of the OSSE was done in an iterative way by tuning after each cycle, the variances of the observation errors for each type of observation based on the method of (Errico et al. 2013). The simulations were first calibrated globally (a coefficient applied to all the variances) and then in a particularised manner (a coefficient specific to an atmospheric level or to a channel). To evaluate the calibrations, we used as figure of merit the standard deviations of the first-guess departures (SDFGD in the following). The objective is that the SDFGD of the observations in the OSSE converges toward those obtained with the real observations. A good agreement between the two is an indicator of the realism of our OSSE framework.

The iterative calibration process was performed as follows:

- Gaussian random and uncorrelated perturbations are added to each synthetic observation. In the first step, the initial

observation errors used for the operational in AROME are considered.

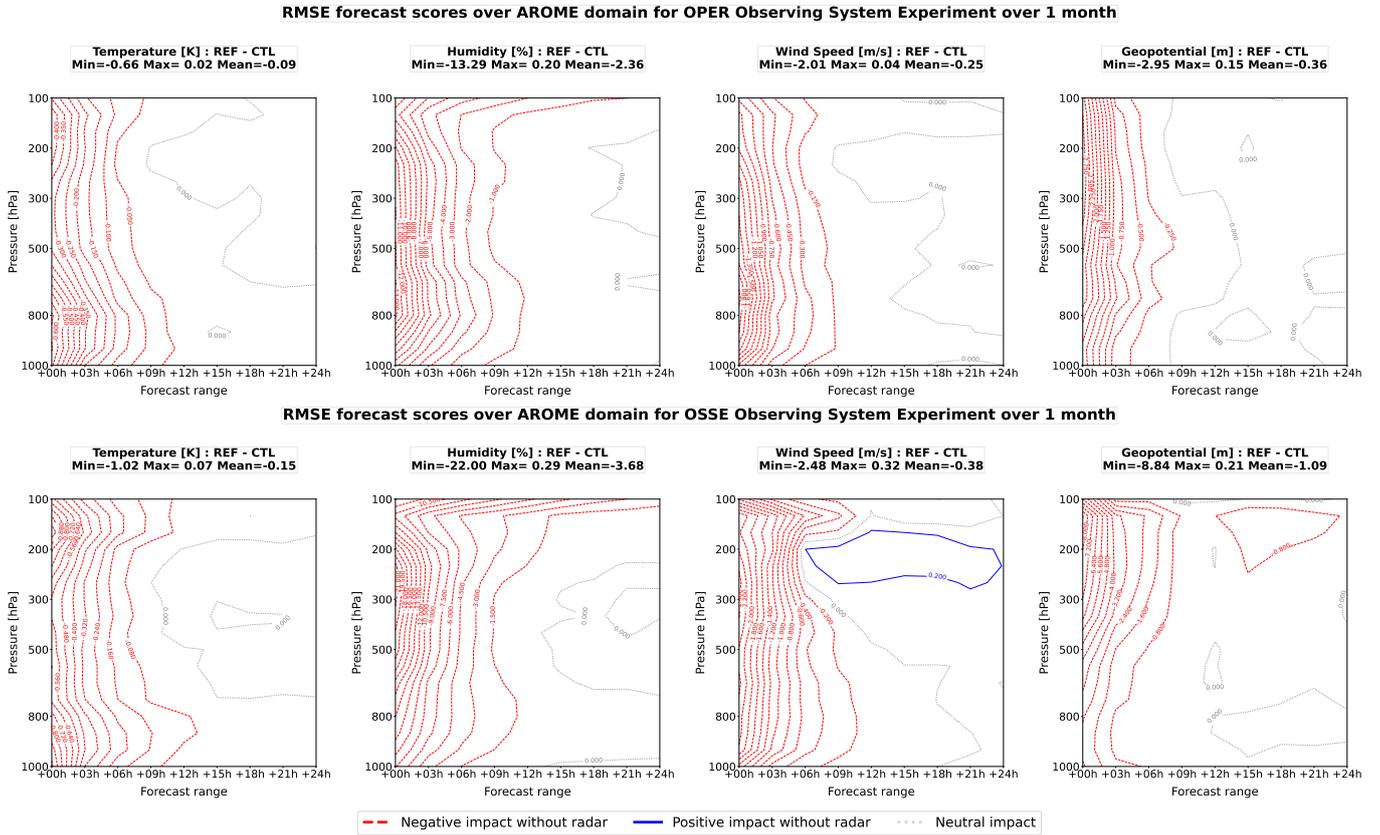
- These synthetic observations are used in a 3D-Var AROME assimilation cycle (see Section 2.4) over 22 days of the summer and winter period, allowing for a reasonable run time and a satisfactory number of statistical cases. Note that the observation errors used in the 3D-Var assimilation for each type of observation are fixed to what is used in operations. Statistics are performed on the mix of the two periods (44 days), in order to have representative results whatever the period considered. The SDFGD calculated from assimilated observations for our experiment are compared to those obtained operationally with real observations (reference) for these same periods and for each observation type.
- A new set of synthetic observations is produced from adjusted observations so that the SDFGD from this iteration is closer to the operational SDFGD. If the SDFGD are larger (resp. smaller) than the ones with real observations, then the applied noise is reduced (resp. increased) in magnitude (between 0 and 70 % depending on the observation type). For this purpose, only the observation errors used for the simulations are changed manually.
- A new assimilation cycle is thus performed with these revised synthetic observations. This process is repeated until the OSSE SDFGDs converge with those of the operational suite.

Satisfactory results of the calibration are obtained after about ten iterations. The results are summarized in Figure 2 and represent the SDFGD for the operational reference (in black dashed line), for the OSSE without calibration (in red line) and with final calibration (in green line) for the main observations assimilated in the 3D-Var AROME assimilation cycle over 44 days (22 summer and winter days). We note a very good agreement between the SDFGD obtained with operational observations and those of the calibrated OSSE, for conventional, radar or satellite observations. These results allow us to be optimistic about the quality of our synthetic observations as well as their impact on short range forecasts of the assimilation experiments in the framework of our OSSE.

## 2.7. OSSE validation

It is important to assess the behaviour of our OSSE to ensure the reliability of our results on the impact of IRS for AROME forecasts. We chose to use the Observing System Experiment (OSE) method to evaluate the effect of adding or removing an individual component of the observing system on the quality of the analyses and subsequent forecasts. NWP centers often perform this type of experiment because of the need to test upgrades to operational systems and to conduct more systematic studies of the value of individual observing system components. Several studies have been conducted using OSEs, such as at ECMWF in the context of model upgrades (Kelly and Thépaut 2007) and general observing system impact studies (Chambon et al. 2022).

In the framework of the AROME 3D-Var data assimilation, the data representing an important part of the observing system are the radars. Moreover, (Lee and Min 2021) has shown in OSE experiments that radars have a significant impact and contribute the most to the forecast of heavy rainfall. Thus, to verify the reliability of our OSSE, we have performed 2 OSE experiments within which we remove the assimilation of radar observations in the 3D-Var AROME data assimilation cycle for an operational framework (OPER) and for the OSSE. For each



**Figure 3.** Temperature, relative humidity, wind speed and geopotential root mean square error (RMSE) over the AROME domain averaged over the period from 1 July to 31 July 2020, as a function of +24h forecast range and pressure levels for OSEs in an operational context (top) and in the context of our OSSE (bottom) from the AROME forecast and verified against its own analysis (with radar).

of these experiments, we performed 1 month of assimilation to achieve forecast scores between a reference experiment 'REF' (with radar observations) and a control experiment 'CTL' (without radar observations) verified against its own analysis (from 'REF').

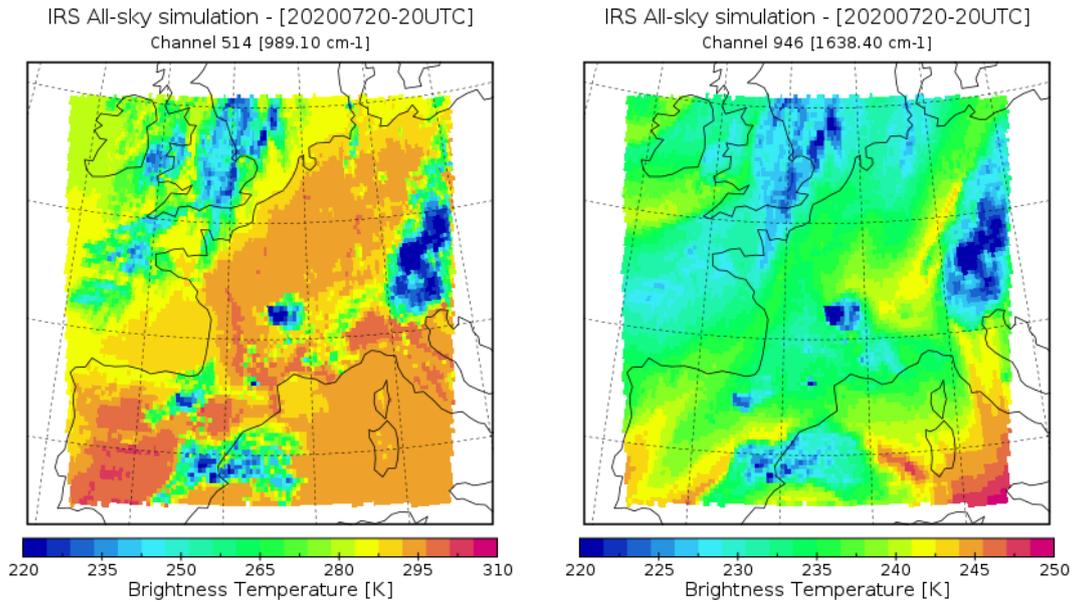
Figure 3 shows the forecast scores in terms of root mean square error (RMSE) for the OSEs for the operational NWP system (top) and the OSSE (bottom) over one month of assimilation from 01 to 31 July 2020 between experiments without assimilation of radars (CTR) and with assimilation of radars (REF). The impacts on the atmospheric vertical are globally homogeneous for all the meteorological parameters evaluated here, with an influence up to about + 09 h. The differences after + 12 h are mainly zero, probably because the influence of LBCs is dominant. For example, for the temperature at 500 hPa and +09 h of forecast range, we notice a degradation of 0.05 % for the OPER and of 0.08 % for the OSSE. This trend is visible for the other parameters studied, aiming to show an overestimation of our OSSE by a factor of almost 2. Thus, a slight overestimation is observed in the OSSE for this summer period. Nevertheless, these OSEs show similar and consistent behaviour, which is reassuring in the performance of OSSE in reproducing the real impact of the radar data. Moreover, we know that on average the initial conditions influence the AROME forecast up to 18 hours and that data assimilation improves up to these ranges compared to dynamically adapted forecasts (Gustafsson *et al.* 2015). Thus, the LBCs do not dominate that much and if we had a data assimilation problem, we would be worse than without observations in the first time range. This result is a positive indicator of the reliability of the forecast scores when adding the future IRS sounder in the next study.

### 3. IRS assimilation in OSSE

#### 3.1. Simulation of IRS observations

The system has been configured for the use of coefficients and optical properties allowing the simulation of IRS observations in cloudy sky. IRS coefficients have been generated using the channel specification (1960) which has different spectral resolutions in the long and short wave bands. This coefficient file takes into account a Hamming-type apodisation in order to obtain highly apodised simulated radiances. This apodisation was chosen because the IRS simulations from RTTOV does not work well with lightly apodised radiances, due to the negative sidelobes of the spectral response function (Atkinson 2022). When real IRS observations are available, this apodisation will be performed by the IRSP (InfraRed Sounder Pre-Processor) software, which will produce highly apodised radiances from the lightly apodised radiances that EUMETSAT will broadcast. IRSP will also be able to handle and generate different formats and will ensure the transition from principal component data to reconstructed radiances.

Several steps were necessary to create the synthetic IRS observations. The first step is similar to the previous one for the AROME observing system, i.e., perturbing the simulations by a Gaussian random function and adding the observation errors. Here, the errors are provided by the IRS instrumental noise varying between 0.18 and 1.80 K (See purple line in Fig. 5). It should be noted that for IRS there was no attempt to tune as for the other currently available radiances. The IRS observations were created for all channels (1960) and for 1 pixel out of 2 in latitude and longitude to limit the volume of data. A selection of pixels was made to be contained in the AROME domain. Note that for the following, a 70 km thinning of the IRS observations



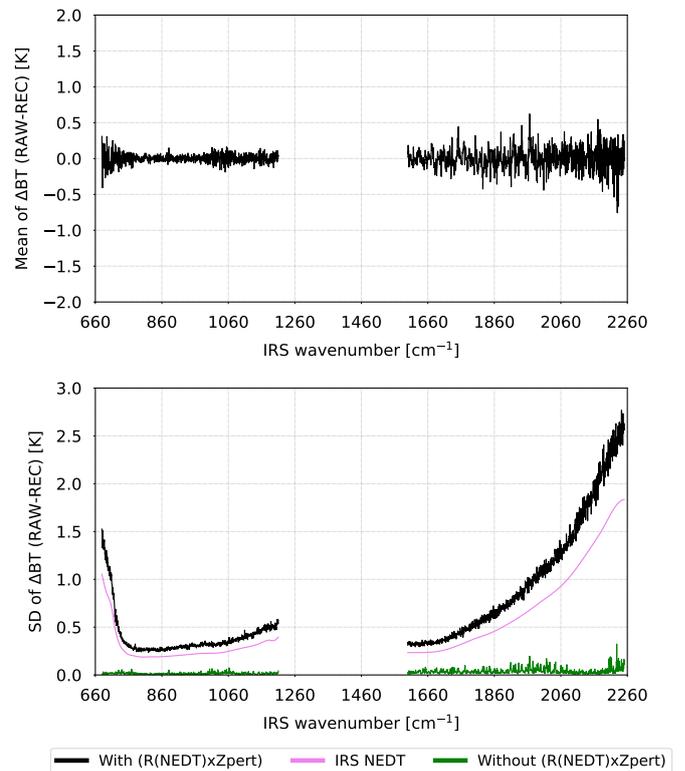
**Figure 4.** Representation of synthetic IRS observations in the AROME domain for a surface [ $989.10 \text{ cm}^{-1}$ ] (left) and water vapor [ $1638.4 \text{ cm}^{-1}$ ] (right) sensitive channel on 04 July 2020 at 20 UTC.

will be used for the assimilation experiments.

On Figure 4, the synthetic IRS observations simulated in all-sky in the AROME domain are shown for a surface (left) and water vapor (right) sensitive channel for 20 July 2020 at 20 UTC. On the left panel we find the cloudy signal characterized by the dark blue areas. We also find the thermal contrast between the continent and the ocean. On the right panel, we observe the water vapor fluxes as well as colder values in the cloudy areas previously identified. Several evaluations were conducted to validate the realism of the IRS simulations with respect to the atmospheric conditions of our OSSE, such as with the cloud cover, temperature and humidity fields. These comparisons have shown a consistency between the atmospheric fields and the IRS simulations (not shown).

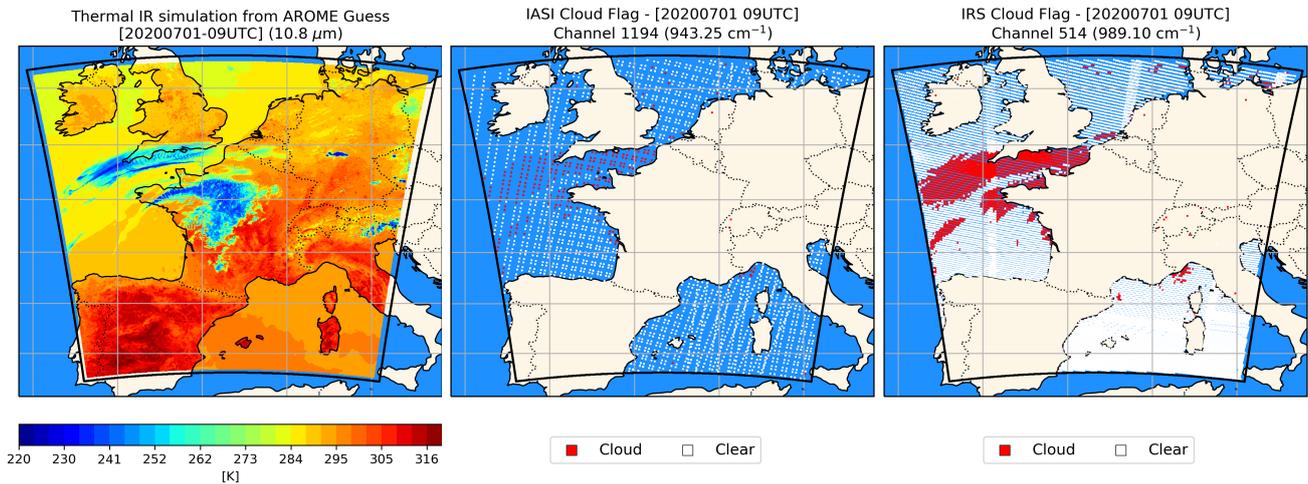
One of the main steps in the creation of synthetic IRS observations is the consideration of principal components. Indeed, EUMETSAT will disseminate the IRS observations in the form of principal components in order to reduce the volume of data in view of the high temporal and geographical sampling that this instrument will generate. To take into account the Principal Components (PCs) in our study, we projected the simulated raw IRS radiances into the 300 PCs space. The radiances were then reconstructed and random uncorrelated Gaussian perturbations (error from IRS instrumental noise) were added. Ideally, the perturbations should be correlated to represent the instrument errors and the additional correlations included with the principal component analysis. However, the AROME 3D-Var does presently not make use of inter channel correlation, this is one of the limitations of the study.

Figure 5 shows the statistics (mean at the top and standard deviation at the bottom) of the differences between the raw and reconstructed synthetic IRS brightness Temperature (BT) over a day of hourly simulations. On average, the differences remain below 0.5 K, which is consistent with what can be obtained with other infrared sounders (Matricardi 2010). On the lower figure, we find the standard deviations of the differences (in black) between the raw and reconstructed synthetic observations. We notice high values mainly due to the addition of the instrumental



**Figure 5.** Statistics of differences between raw and reconstructed brightness temperatures from IRS synthetic observations over one day of data, mean (top) panel and standard deviation (bottom) panel. The purple curve represents IRS instrumental noise, the black and green curve represents respectively, BT with and without perturbations of the simulations.

noise of IRS (in purple) in the creation of synthetic observations. Indeed, without perturbations of the observations (in green) the values of the standard deviations of the differences are much more reasonable, lower than 0.1 K.



**Figure 6.** Map representing a thermal infrared simulation from an AROME guess (left), IASI pixels (middle) and IRS pixels identified as clear or cloudy over sea on July 01, 2020 at 09 UTC.

### 3.2. 3D-Var AROME configuration for IRS

#### 3.2.1. Cloud detection

For the assimilation experiments of the synthetic observations of IRS in the 3D-Var data assimilation system AROME, several adjustments were necessary in order to use these data. Thus, a new configuration of the system has been built in order to give it the capacity to assimilate the IRS observations. To be as realistic as possible, we have simulated IRS data for all-sky conditions. This ensures a realistic sampling of the atmosphere in the infrared spectrum which is characterized by an important sensitivity to cloud ice and cloud liquid water.

As in most of the other NWP centers, the capability to assimilate IR observations affected by clouds has not been developed yet. These contaminated observations represent unfortunately a significant fraction of the observations and need to be filtered out with a dedicated technique. For example with the observations from AIRS, (McNally and Watts 2003) has shown that about 90 % of its pixels are totally or partially contaminated by cloudy signals.

Among these methods, the AROME 3D-Var data assimilation system uses the ECMWF Aerosol and Cloud Detection Software which uses the method given by (McNally and Watts 2003) available on the NWPSAF website and which detect all the channels affected by the presence of a cloud for a given meteorological scene. To reveal the radiative effect of the cloud and to separate it from other contributions, the cloud detection algorithm works by taking the background departures (i.e., the difference between the observed and simulated in clear-sky BT) and looking for the signature of opacity that is not included in the clear-sky calculation (i.e. cloud). To do this, the channels are first ranked in the vertical, according to their height assignments (Eresmaa 2020).

At Météo-France, a selection of channels is used for the cloud detection scheme. For IASI, we have kept the one used in operations. For IRS, a new configuration of the cloud detection scheme has been made with a new list of channels. Several tests of assimilation of IRS observations have been conducted to obtain consistent results on cloud identification and the relative amount of rejected observations compared to what is done for IASI. On Figure 6, a thermal infrared simulation of SEVIRI at  $10.8 \mu\text{m}$  from an AROME guess (left) is represented on maps for July 01,

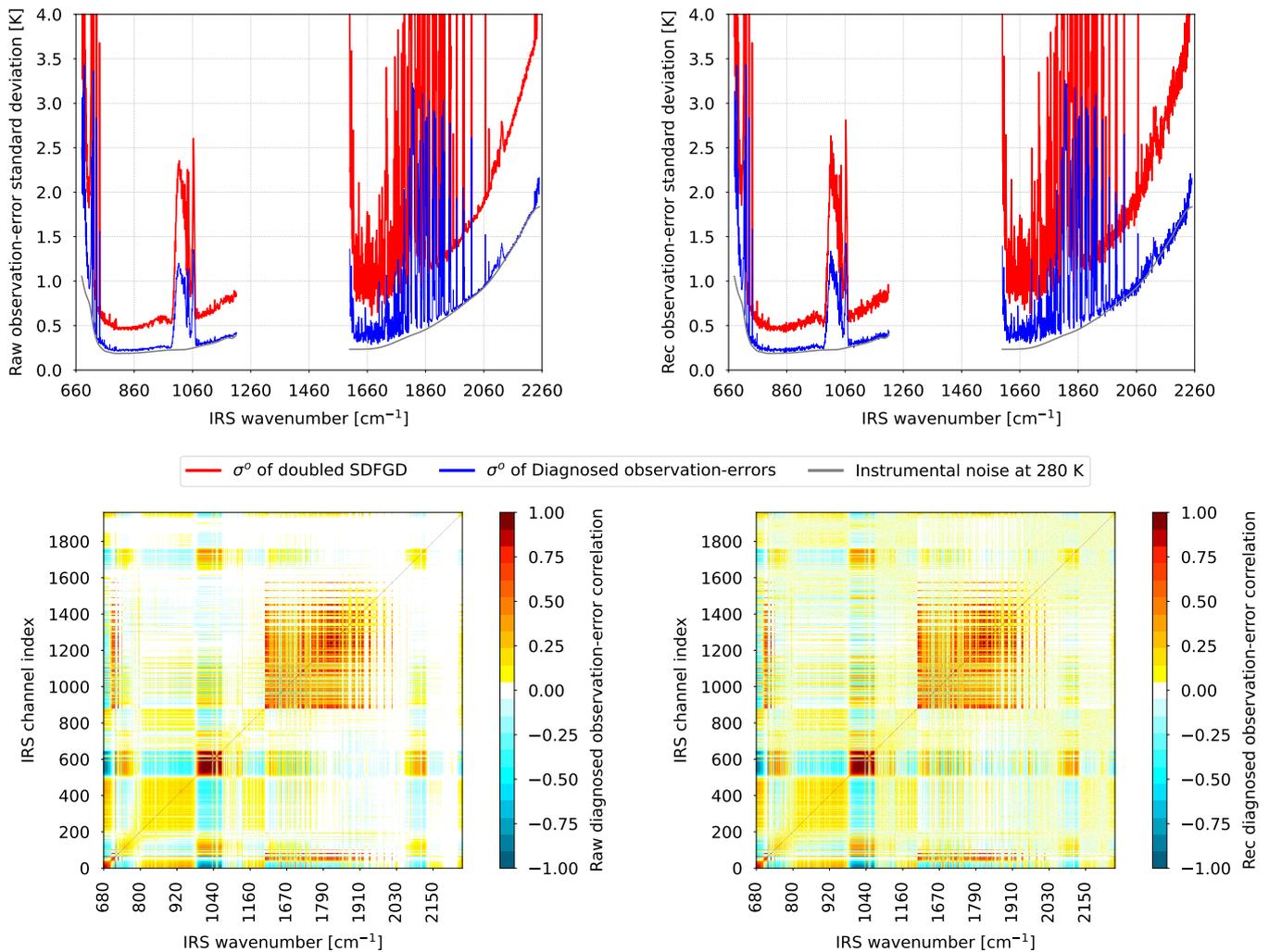
2020 at 09 UTC. In the middle is represented the IASI pixels identified as clear (in white) and cloudy (in red) over sea and the same applies on the right figure for IRS. We notice on this example that where we observe clouds (in dark blue) on the left map, the cloudy detection scheme correctly identifies the pixels contaminated by a cloud for surface channels from IASI and IRS. There are also some cloudy flagged pixels on the Mediterranean coast that do not appear cloudy on the thermal infrared simulation. This is probably due to very low clouds over the sea which are not represented in the ISP. These results have allowed us to validate the correct functioning of the cloud detection for our synthetic IRS observations, which will be a guarantee of reliability during the 3D-Var AROME data assimilation experiments. A more detailed study is provided also in (Coopmann *et al.* 2022).

#### 3.2.2. IRS observation-error

An accurate estimation of observation errors is necessary because they have a significant impact on the weight given to the observations within the assimilation and therefore on the resulting analyses. Fortunately, several methods have been developed over the years to calculate these observation errors and to take into account each of the uncertainties that make them up, including diagonal and non-diagonal terms.

One of the most widely used techniques was introduced by (Desroziers *et al.* 2005). By deriving estimates of optimal observation errors from the departure statistics of the assimilation systems, this method makes it possible to diagnose both the variances and covariances of the observation error, including error correlations between channels, allowing an accurate estimate of the total errors that characterize the observation (instrumental noise, spatial representativeness error, error in the calculation of the radiative transfer, etc.).

This work was conducted in an OSSE context with synthetic IRS observations. This has already been done by (Vittorioso *et al.* 2021) in a similar study context with synthetic IASI-NG observations. Privé *et al.* (2013a) has shown that it is possible to apply an accurately quantified error to synthetic observations. However, in an OSSE, it is important to keep in mind that the observation errors are much smaller than those of real observations. To estimate the observation error covariance matrix IRS as accurately as possible, we proceeded in three steps:



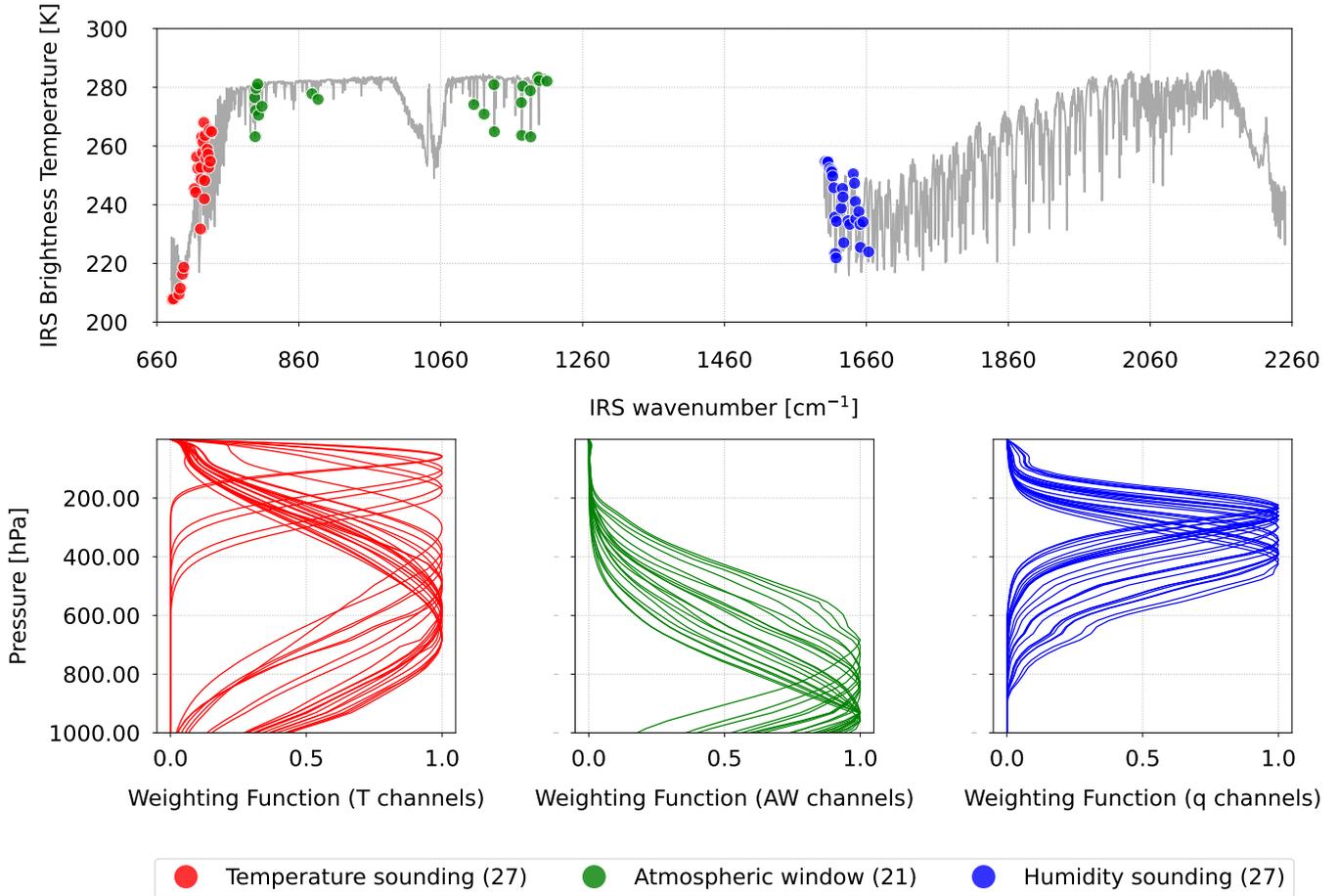
**Figure 7.** Observation-error standard deviation (top) and correlation matrix (bottom) for IRS from raw (left) and reconstructed brightness temperature diagnosed for 4 days. On the top graphs; the standard deviations of observation errors from the doubled SDFGD (in red), the diagnosed observation errors (in blue) and the instrumental noise of IRS at 280 K (in grey). At the bottom are represented the correlation matrices of diagnosed observation errors for the 1960 IRS channels.

- First, 3D-Var AROME experiments were conducted to assimilate the 1960 IRS channels (raw and reconstructed radiances) on two days of the summer and winter period. For these experiments, a simple diagonal matrix containing variances from the IRS instrumental noise was considered. At the end of these runs, SDFGDs are calculated (raw and reconstructed radiances) over 4 days including the 2 summer and winter days in order to take into account the seasonal variability of errors.
- Then a second run of assimilation experiments is performed as in the first step, but this time using the SDFGDs as variances, inflated by 2 to take into account the lack of correlation (Stewart et al. 2014). These inflated  $\sigma^o$  are represented by the red curve on the top graphs of Figure 7, on the left with raw and on the right reconstructed brightness temperatures.
- Finally, at the end of these second assimilation runs, we applied the Desroziers method to diagnose the observation error covariance matrix ( $\mathbf{R}$ ) in the observation space with the deviations of the observations from the background and the analysis.  $\sigma^o$  of diagnosed observation errors are represented in blue on the top graphs of Figure 7 for raw (left) and reconstructed (right) brightness temperatures. Similarly, the statistics were calculated over 4 days (2 days in summer and winter). The diagnostic of  $\mathbf{R}$  was also performed separately between the summer and winter

period and slightly different  $\sigma^o$  values are observed (not shown). However, for the following study we have chosen to use an  $\mathbf{R}$  matrix diagnosed on the basis of 4 days in order to make it applicable to any period of the year.

In the top graphs of Figure 7, we see larger SDFGD values for some channels in the  $\text{CO}_2$  absorption band (680 - 700  $\text{cm}^{-1}$ ) and the water vapour absorption band (1600 - 2020  $\text{cm}^{-1}$ ). This is due to a sensitivity of these channels to the model top. Indeed, the 3D-Var AROME data assimilation system is configured operationally with a model top at 10 hPa, but some IRS channels are sensitive above this limit. The synthetic IRS observations were simulated from a version using a model top of 0.1 hPa in order to increase their realism. Finally, the higher values observed for the channels of the ozone absorption band (1000 - 1060  $\text{cm}^{-1}$ ) result from not taking into account the realistic ozone field in the model. As expected, the diagnosed  $\sigma^o$  values (in blue) converge towards the instrumental noise curve (in grey). We find the same structures as before with higher values for the channels sensitive to the model top and to ozone. Regarding the differences between raw and reconstructed observations, we notice noisier values for the reconstructed observations especially on the last part of the spectrum between 2000 and 2250  $\text{cm}^{-1}$ .

Finally, Figure 7 shows at the bottom the diagnosed observation error correlation matrices for the raw (left) and reconstructed (right) observations. Even if we do not introduce correlated noise



**Figure 8.** Location of the 75 assimilated IRS channels on a typical spectrum (top) by application type; temperature-sensitive channels (red), atmospheric window channels sensitive to surface, temperature and humidity in the lower troposphere (green) and water vapour sensitive channels (blue). The normalized weighting functions (bottom) for these 75 channels are separated by these same types of application as a function of pressure.

on the channels when creating the synthetic observations, we still diagnose error correlations induced by the intrinsic spectroscopy error and also the model bias via the observation operator when simulating the IRS observations. Finally, the difference in simulation between the AROME NR at 2.5 km and the AROME 3D-Var at 1.3 km should also induce representativeness errors in the diagnosis of the observation error covariance matrix.

One can see different significant correlation block structures centred around the diagonal. These channel blocks have typical signatures in the spectra of infrared sounders (Matricardi 2010). We find the temperature-sensitive channel block from the top of the atmosphere to the lower troposphere between 680 - 800  $\text{cm}^{-1}$ , the two atmospheric window channel blocks between 800 - 1000  $\text{cm}^{-1}$  and 1070 - 1210  $\text{cm}^{-1}$ , the ozone-sensitive channel block between 1000 - 1070  $\text{cm}^{-1}$  and the water vapour-sensitive channel block between 1600 - 2020  $\text{cm}^{-1}$ . Strongly correlated errors are observed within the upper-air temperature, ozone and water vapour sensitive channels. This is mainly due to the sensitivity of some of these channels to the top of the AROME model. As expected, the use of reconstructed brightness temperatures introduces higher inter-channel correlations. This error diagnostic step with the reconstructed brightness temperatures is necessary since in the following we will use these reconstructed synthetic IRS observations for the assimilation experiments. This ensures that the proper correlations are taken into account within the data assimilation, even though they are enhanced with the PC compression (Lupu *et al.* 2021).

## 4. Assimilation experiments and results

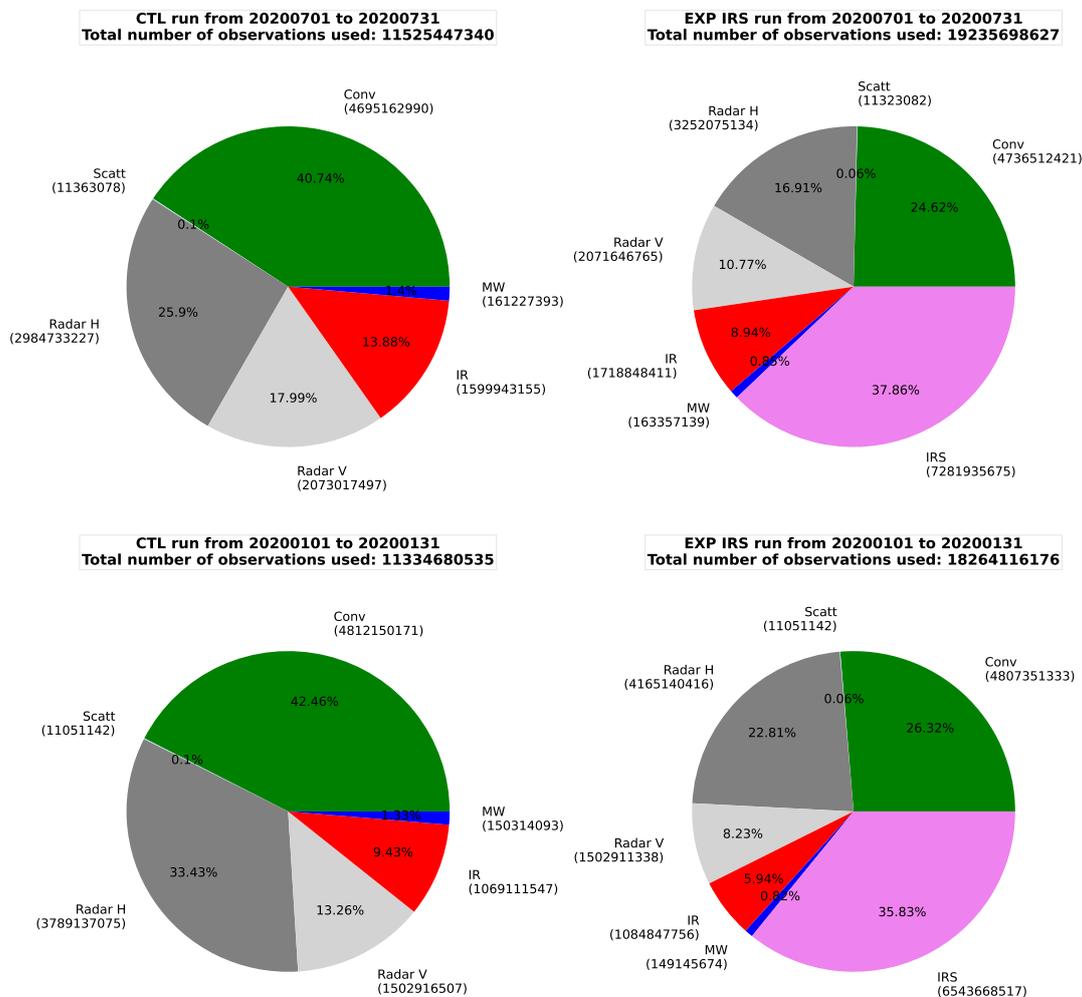
### 4.1. Experimental setup

We performed data assimilation experiments using an operational configuration of the AROME 3D-Var NWP system. An analysis is provided every hour, followed by a 1 h forecast that serves as a background for the next analysis. Every day at 00 UTC, a +48 h long-range forecast is produced. The evaluation will therefore be carried out over two winter months from 01 January to 29 February 2020 and two summer months from 01 July to 31 August 2020.

For each of the study periods, two different data assimilation experiments were conducted:

- A control experiment 'CTL' mimicking the operational AROME context, assimilating the full synthetic observing system (see Table 1),
- An IRS experiment 'EXP IRS' assimilating all the synthetic observations of the 'CTL' experiment as well as the synthetic observations from IRS.

For the experiments assimilating IRS reconstructed BT, we chose to use a sub-selection of 75 channels from the 300-channel selection made in (Coopmann *et al.* 2022). The 75-channel set includes 27 temperature-sensitive channels, 27 water vapour-sensitive channels and 21 atmospheric window channels. The location of these channels and their weighting function is shown in Figure 8. The temperature-sensitive channels cover the atmosphere between 100 and 700 hPa, while the channels of the atmospheric window sensitive to temperature and



**Figure 9.** Ratio of the number of assimilated observations per instrument type for the summer period (top) and the winter period (bottom) for the control experiments (left) and the IRS experiments (right) over one month of data.

humidity cover the lower troposphere from 700 hPa to the surface. Finally the water vapour sensitive channels cover part of the atmosphere between 200 and 500 hPa. The assimilated synthetic IRS observations are derived from reconstructed brightness temperatures converted from principal components (see Section 3.1.). In this study, we applied a thinning of 70 km for the IRS pixels and limited the assimilation to oceanic surfaces. A diagnosed observation error covariance matrix is used for the 75 IRS channels, derived from the previously estimated variances and correlations (see Section 3.2.2.). In order to assess the impact of adding the IRS instrument in addition to the current observing system, the forecasts from our assimilation experiments will be compared to the Nature Run AROME considered as our atmospheric 'reality' in our OSSE.

#### 4.2. IRS impacts on the AROME NWP system

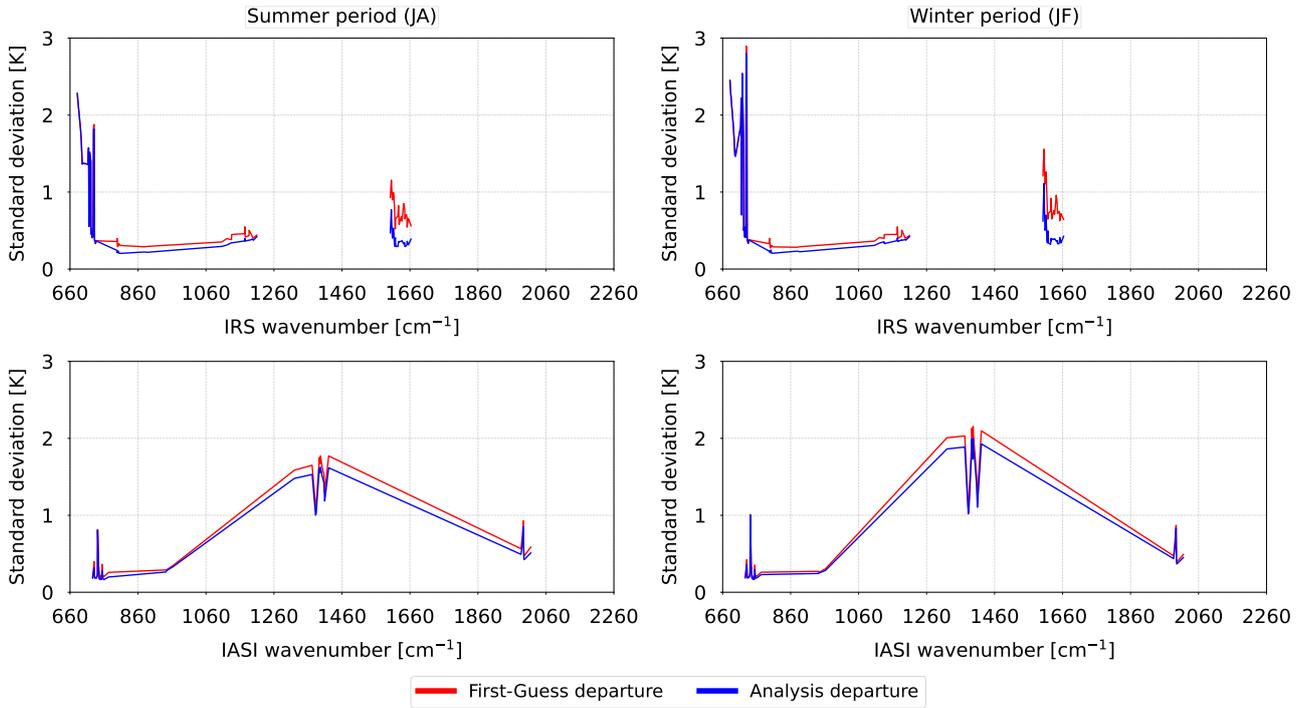
##### 4.2.1. Proportion of assimilated observations

The first impact of the addition of the future IRS sounder in the AROME NWP system is on the proportion of assimilated observations. Figure 9 shows the fraction of the number of assimilated observations per instrument type for the summer period (top) and the winter period (bottom) for the control experiments (left) and the IRS experiments (right) over one month of data.

The graphs of the number of observations assimilated within the control experiments (left plots) show similar behaviour to

that observed for the operational for these same periods. There are also well identified trends such as the increase in the number of conventional observations and radar humidity data used in winter periods and at the same time the decrease in the number of assimilated infrared observations. This is mainly due to a decrease in anticyclonic conditions and a higher frequency of synoptic perturbations and precipitation. This increases the use of radar data and reduces the assimilation of infrared observations due to cloud contamination and also due to the fact that radars see through rain which is more frequent in winter. Indeed, the calculation of the average cloudiness and rainfall over the full domain of the AROME NR for both periods shows significant differences between the two seasons; with an average of about 0.5 mm of rainfall, 12 % of low clouds, 13 % of medium clouds and 24 % of high clouds in the summer. While the winter season averaged about 2.5 mm of rain, 48 % low cloud, 34 % medium cloud and 42 % high cloud.

For the IRS experiments (right plots), the assimilation of the 75 IRS channels over oceans changes the fraction of different types of assimilated observations in the 3D-Var AROME. Conventional and radar data represent 85 % of the observations assimilated in the control experiment but with the addition of IRS, they represent only half of the observations used. Indeed, the addition of IRS represents almost 50 % additional assimilated observations in the NWP system over the two periods. As with the other infrared observations, the amount of IRS data assimilated is less for the winter period. It can also be seen that for the summer period, the addition of IRS has an impact on the number of



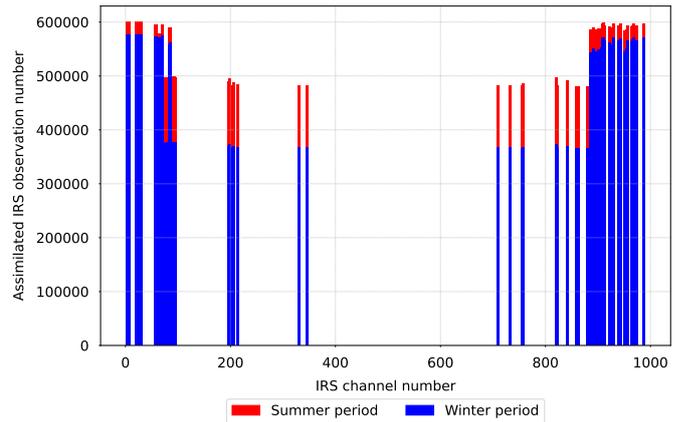
**Figure 10.** Standard deviations of First-Guess departure and Analysis departure for IRS (top) and IASI (bottom) over the summer (left) and winter (right) period of the channels assimilated in the 3D-Var AROME.

other assimilated observations, with an increase for conventional observations, for humidity radar and for infrared. For the winter period, only the humidity radar data are significantly increased. For the other observations, the ratio remains relatively the same or even decreases. Table 2 summarises for a summer and winter month the percentage difference in the number of observations assimilated between the EXP IRS and CTL experiments. The (+) or (-) indicate more or less observations in the EXP IRS experiment compared to CTL.

Observations	Summer (EXP IRS - CTL)	Winter (EXP IRS - CTL)
Conv	+ 0.9 %	- 0.1 %
Scatt	- 0.4 %	0 %
Radar H	+ 8.2 %	+ 9.0 %
Radar V	- 0.1 %	0 %
IR	+ 6.9 %	+ 1.5 %
MW	+ 1.3 %	- 0.8 %

Table 2. Percentage difference in the number of observations assimilated between the EXP IRS and CTL experiment.

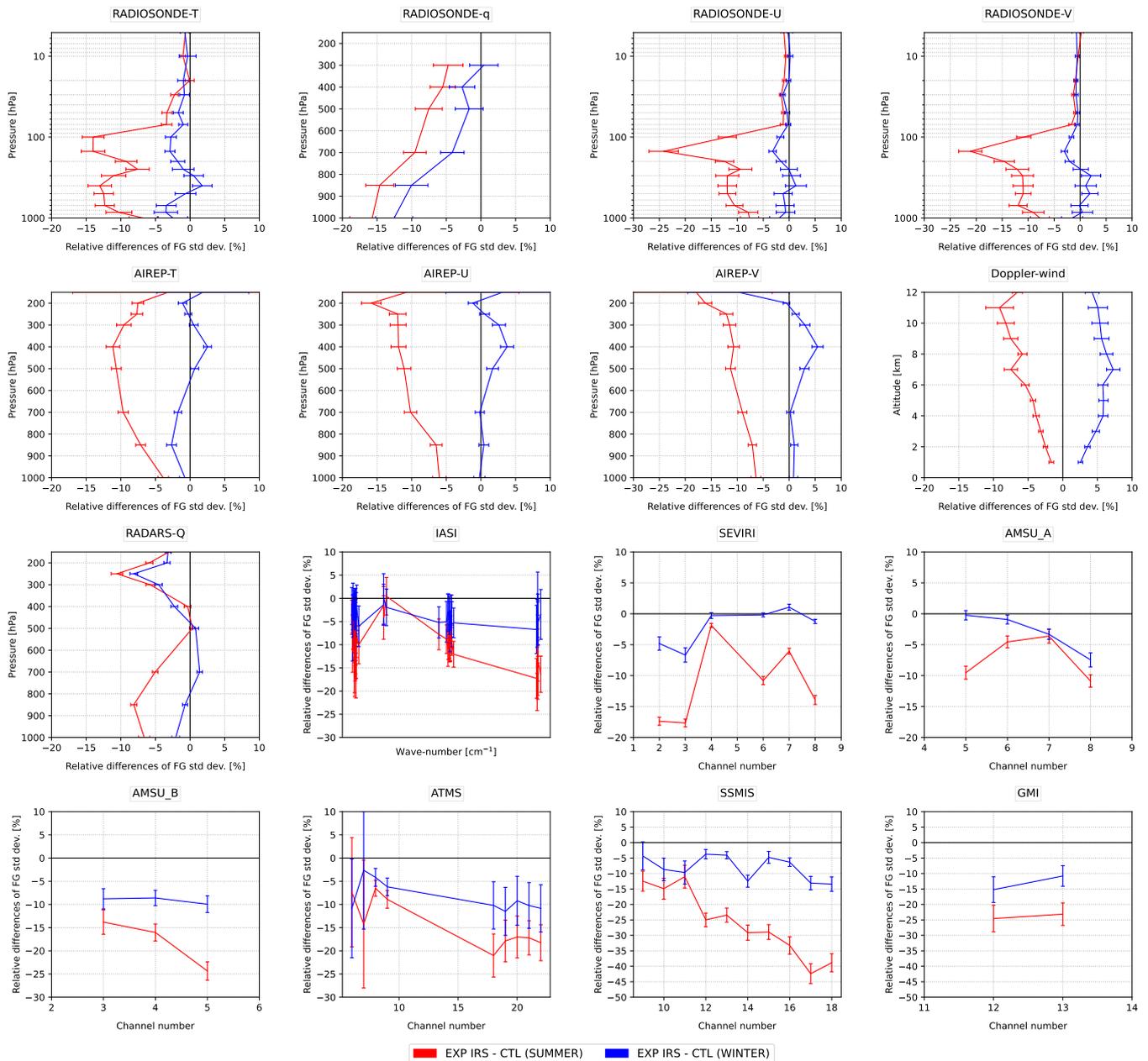
In Figure 10 we have evaluated the standard deviations of the First-Guess departure and Analysis departure for IRS and IASI of the assimilated channels in the 3D-Var AROME over each period. Even if it is complicated to compare the results between IRS and IASI because we do not assimilate the same channels, similarities in minimisation behaviour can be observed. Indeed, for both instruments, we notice a smaller gap between the First-Guess departure and the Analysis departure for the channels sensitive to CO<sub>2</sub> absorption [660 - 760 cm<sup>-1</sup>] and larger gaps for the channels sensitive to the atmospheric window [760 and 1260 cm<sup>-1</sup>] and the water vapour absorption band [1260 - 2060 cm<sup>-1</sup>]. We also observe higher values between [726 and 738 cm<sup>-1</sup>] for the IRS channels than for the IASI channels whatever the period. Finally, more globally, we notice that the values tend to be higher for the winter period, particularly for the channels in the CO<sub>2</sub> and water vapour absorption bands for both IRS and IASI.



**Figure 11.** Number of assimilated IRS observations per channel for the summer period (in red) from 01 July to 31 August 2020 and the winter period (in blue) from 01 January to 29 February 2020.

Looking more specifically at the behaviour of the assimilation of IRS observations over the two study periods, we notice in Figure 11 that the number of data assimilated in summer is greater than in winter, mainly for the certain channels located towards the end of the CO<sub>2</sub> absorption band that are sensitive to temperature in the troposphere [721 - 736 cm<sup>-1</sup>] and those of the atmospheric window [798 - 1210 cm<sup>-1</sup>] that are sensitive to temperature and water vapour in the lower troposphere. For these channels, up to one hundred thousand fewer observations are assimilated in winter as a consequence of a larger cloud cover for this season and therefore a stronger rejection of cloud contaminated pixels.

Thus, depending on the period, the assimilation of IRS leads to a different number of observations used and globally to higher standard deviation values of first-guess departure and analysis departure in winter. This behaviour can explain the results of Figure 9, which shows a greater number of other observations assimilated in summer than in winter. The assimilation of IRS has a rapid impact on the analysed weather fields, so in summer more IRS observations are assimilated than in winter, the weather



**Figure 12.** Global relative differences in First-Guess departure standard deviation between CTL and EXP IRS experiments on the main current assimilated observations with 1 h assimilation cycles from the summer period (in red) between 01 July and 31 August 2020 and the winter period (in blue) between 01 January and 29 February 2020. Values below 0 indicate an improvement with respect to the CTL experiment. Error bars give statistical significance internals for differences at the 95 % level.

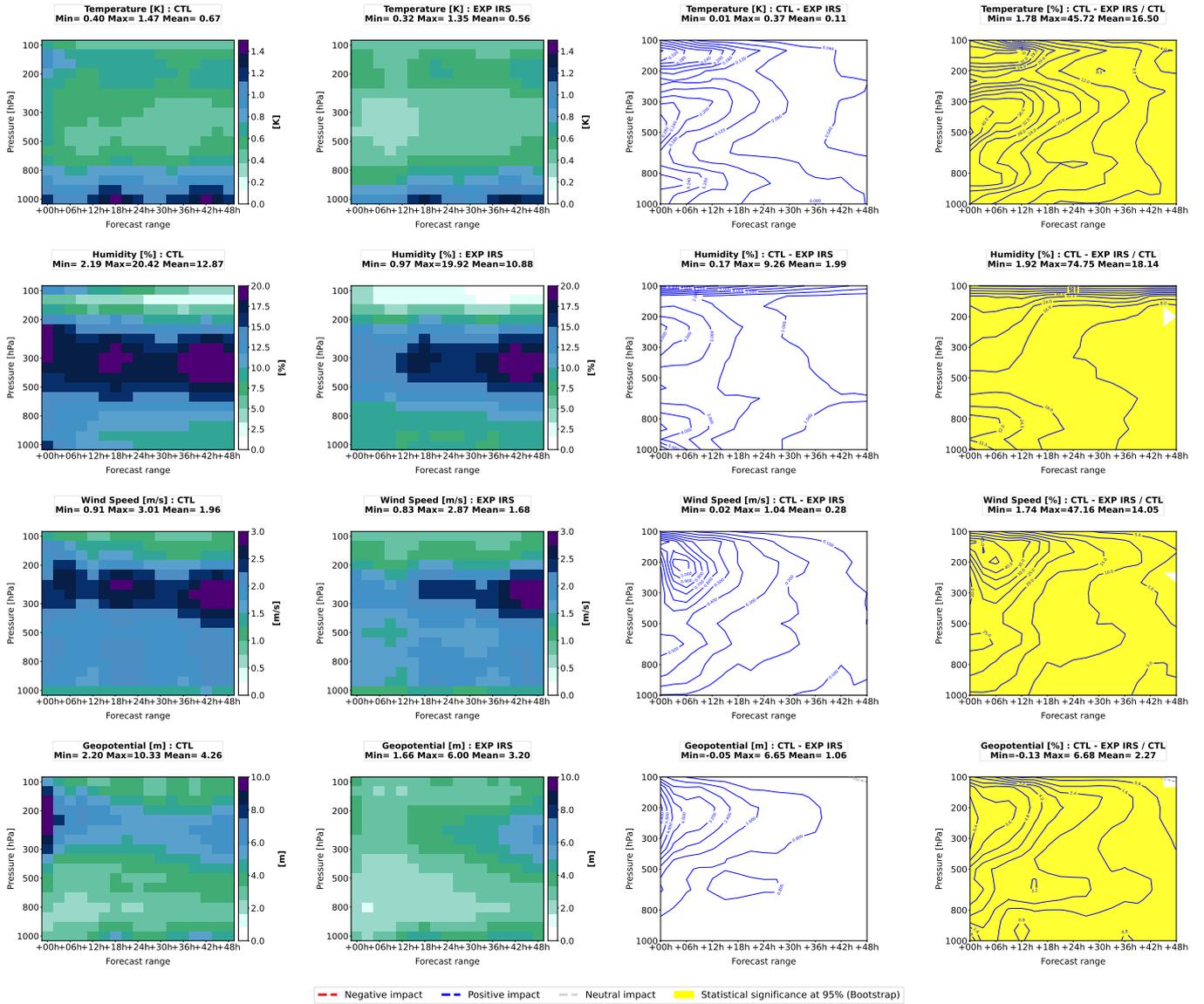
fields are therefore more accurate in summer, which leads to a more advantageous First-Guess-Check for the observations for this period. These differences will probably have an impact on the quality of the analyses and forecasts in winter.

#### 4.2.2. Impact on other observations

The impact of IRS assimilation is first analysed in the short term within the 1h 3D-Var cycle. The addition of such a large amount of IRS observations to the AROME NWP system will have a significant influence on the first guess. As this new atmospheric field is used for the evaluation of the observation function of the full observing system, a significant impact on the behaviour of the other types of observations is expected compared to the control experiment. Figure 12 shows the impact of the IRS assimilation in terms of relative differences in First-Guess departure standard deviation between CTL and EXP IRS experiments on the main currently assimilated observations with 1 h assimilation cycles from the summer period (in red) and the winter period (in blue).

Values below 0 % indicate an improved fit of the observation and the error bars give statistical significance internals for differences at the 95 % level.

It can be seen that the IRS assimilation has a substantial impact on the other observations, mainly in the summer period, with significant generalised improvements for all the observations evaluated here. The improvements for the summer period can reach up to 25 % for radiosondes mainly in the troposphere and up to 15 % for aircraft in this same part of the atmosphere. Radar data are also better simulated for this period, as well as infrared and microwave observations with improvements ranging from 10 to 40 %. The impact of IRS assimilation on other observations in winter is less important but nevertheless globally neutral to positive with reduction of 5 to 15 %. However, we note some significantly negative impacts for aircraft between 300 and 500 hPa, for radar winds over the entire atmospheric vertical and for radar humidity at 700 hPa. These degradations during the winter period could be an indicator that the ingredients of the IRS assimilation may need to be refined with real observations,



**Figure 13.** From left to right; the root mean square error for the control experiment (CTL), the IRS experiment (EXP IRS), the differences between (CTL - EXP IRS) and the normalized differences [%] for the AROME forecasts with respect to the Nature Run (NR) as a function of the forecast range (+48 h) and pressure levels over the period from 1 July to 31 August 2020. From top to bottom; for temperature, relative humidity, wind speed and geopotential. The blue isolines (respectively red) indicate that EXP IRS experiment have improved (resp. degraded) the forecasts. Yellow shaded region indicate statistical significance at 95 % computed from bootstrap test.

including the cloud detection which may need to be set to filter more scenes and avoid degradations in winter. These refinements are left for future investigations with real data.

It is important to bear in mind that we have made the choice to calibrate the synthetic observations of the current system by taking into account the two periods combined for the statistics. The objective of this method is to capture the variability of the impact of the observations over both study periods, so that the OSSE is applicable at any time of the year. This technique provides a general calibration of the synthetic data but may potentially over- or underestimate the impact of these observations during a particular season. The impact of the calibration is referred to in the further part of the paper as the "Universal Error Calibration Effect".

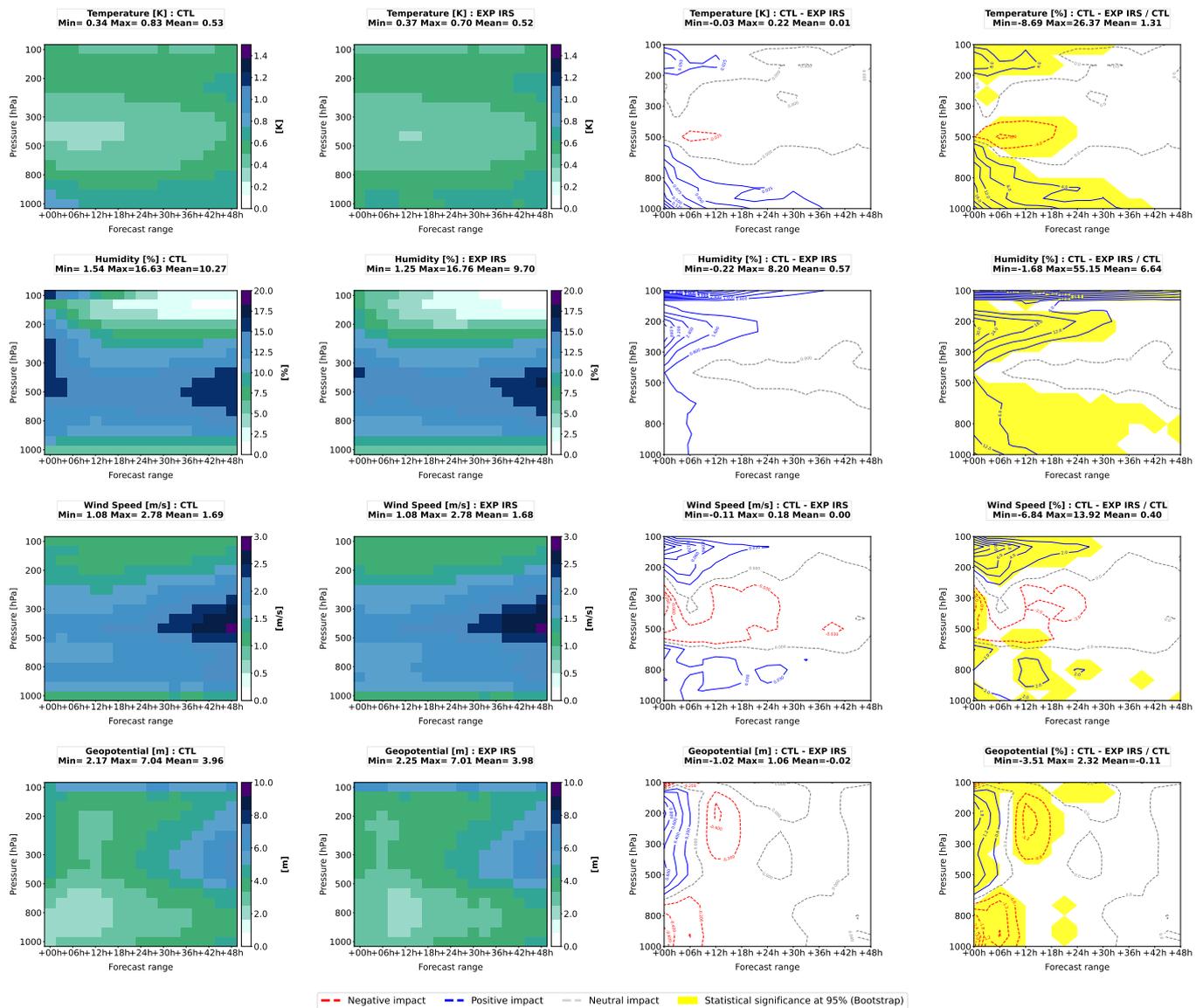
#### 4.2.3. Forecast scores

The objectives of this study are multiple; implementation of an OSSE framework, evaluation of the Nature Run, process and calibration of current and IRS synthetic observations, preparation of the AROME NWP system to assimilate these future observations, etc. This experimental framework also allows

a first evaluation of the impact of IRS assimilation on AROME weather forecasts. To achieve this goal, a statistical evaluation of the forecast errors (forecast scores) was carried out over two months of each study period by calculating the root mean square error (RMSE) between the different experiments.

We have thus calculated for the summer period from 01 July to 31 August 2020 (Figure 13) and for the winter period from 01 January to 29 February 2020 (Figure 14) the root mean square error of forecasts compared to the NR as verification data. From left to right; for the control experiment (CTL), the IRS experiment (EXP IRS), the difference between CTL - EXP IRS and the normalized difference between CTL - EXP IRS, as a function of the forecast range (+48 h) and the pressure. The same metrics are shown for the summer period on Figure 13).

It can be seen that, for the summer period, the assimilation of the IRS results in a significant reduction of the forecast error for all parameters and over a long forecast range. This trend is impressively reflected in the normalized forecast scores with an almost universal improvement of EXP IRS forecasts compared to CTL for all parameters and forecast ranges. In the framework of our OSSE, the IRS assimilation leads to a



**Figure 14.** From left to right; the root mean square error for the control experiment (CTL), the IRS experiment (EXP IRS), the differences between (CTL - EXP IRS) and the normalized differences [%] for the AROME forecasts with respect to the Nature Run (NR) as a function of forecast range (+48 h) and pressure levels over the period from 1 January to 29 February 2020. From top to bottom; for temperature, relative humidity, wind speed and geopotential. The blue isolines (respectively red) indicate that EXP IRS experiment have improved (resp. degraded) the forecasts. Yellow shaded region indicate statistical significance at 95 % computed from bootstrap test.

significant average improvement of 16.5 % in temperature, 18 % in relative humidity, 14 % in wind speed and 2 % in geopotential forecasts for this period. Nevertheless, it is important to put the results obtained for relative humidity between 100 and 150 hPa into perspective. Indeed, the low concentrations and the variability that water vapour can have at this altitude between the different experiments, increases the errors. However, the RMSE calculation gives more weight to large errors than to small ones.

The results are more mixed for the winter period (Figure 14). Significant improvements in temperature forecasts are obtained for the EXP IRS experiment in the lower troposphere up to +42 h and in the upper troposphere up to +24 h. Nevertheless, a significant degradation is observed around 500 hPa up to +24 h. For relative humidity, there is a significant improvement in the forecasts, particularly in the upper and lower troposphere up to about +36 h and a neutral impact is observed around 500 hPa. A similar behaviour can be observed for the wind speed forecast scores, with a significant improvement in the lower and upper troposphere up to +24 h and a degradation between 300 and 500 hPa, mainly significant over the first 3 hours. Finally, the geopotential forecast shows a significant improvement between 150 and 700 hPa up to +06 h, then a degradation between

+12 h and +18 h and also a significant degradation in the lower troposphere up to +12 h. Globally, the addition of IRS is also positive in terms of the forecast scores over the winter period, but somewhat less so compared to the summer period.

How can we explain the strong positive impact of the IRS on the forecasts for this period? Firstly, this result was relatively expected for the summer period as it is a season with less cloud cover overall. This allows more IRS observations to be assimilated down to the surface (see Figure 11), around 10 % more IRS observations to be used compared to the winter period. In addition, this time of year is less subject to synoptic perturbations and therefore to precipitation, thus reducing the impact of the radar data. Finally, the assimilation of IRS has a direct impact on the assimilation of other observations as shown in Figure 12. Thus, the combination of; a high volume of assimilated information from IRS, a more realistic analysis and a better assimilation of other types of observations, provides the necessary conditions for a significant improvement of weather forecasts. Of course, it should be kept in mind that the OSSE is an idealized case that may overestimate the impacts. Indeed, as the previous OSE study has shown (Section 2.6), the OSSE provides similar results to the operational in terms of atmospheric

behaviour over the vertical and over the forecast ranges with nevertheless a slight overestimation of the impact. Several reasons can explain this; (i) simplified observation error characteristics, (ii) simplifying assumptions such as the use of a "perfect" surface, (iii) an over-optimistic simulated observation coverage due to an underestimation of the cloud contamination (Masutani *et al.* 1999). Finally, larger errors are observed in summer than in winter in the CTL and EXP IRS experiments. This is often observed in AROME and is explained by the fact that it is more complicated to predict convection in summer than perturbations in winter. Thus, the summer period offers a much greater potential for improved forecasts for IRS observations than the winter period. Even if the percentage improvement is slightly amplified, the contribution of IRS to the reduction of forecast error is very clear.

Moreover, these results can be supported by a similar study conducted by (McGrath-Spangler *et al.* 2022). This work was also carried out using an OSSE and concerns the evaluation of the impact of assimilating infrared observations from a hyperspectral sounder on board a geostationary satellite on the NWP, within the framework of the GeoXO programme in the USA. The results of this paper show that the assimilation of 85 channels (70 channels for temperature sounding and 15 channels sensitive to water vapour) from their GeoXO sounder into their NWP mesoscale system with a thinning of 180 km over a period from 01 July to 31 August 2006, leads to improved initial conditions used for subsequent forecasts and in particular to a significant improvement of the 24 to 48 hour forecasts for temperature and zonal wind over the full atmospheric column. An even greater improvement is seen in the specific humidity of the troposphere over a longer forecast range. This work is consistent with our results for the assimilation of a future GEO sounder of the impact on weather forecasts.

This difference in impact between summer and winter can be explained in several ways. From the IRS point of view, we have seen that cloudy conditions in winter decrease the amount of assimilated observations by about 10 % compared to the summer period. Especially for sensitive channels in the mid-troposphere, which can limit the modification of the analyses at this altitude, thus influencing the impact on the forecasts for this part of the atmosphere. Moreover, Figure 8 shows a lower density of sensitive IRS channels between 400 and 600 hPa. It is specifically at this location that degradation or more limited impacts on the forecast scores are observed. However, this cannot explain a degradation but rather a neutral behaviour. Then, we have shown in Figure 12 that the impact on current observations is not necessarily positive for the winter period, particularly on conventional and radar observations. More specifically, we notice in Figure 14 that the altitude level where there is a degradation of the temperature forecast (between 400-500 hPa) corresponds to the same atmospheric level where we identify a degradation of the impact of radiosonde and aircraft in temperature. Similarly, the degradation of the wind speed forecast between 300 and 500 hPa corresponds to the same level as the degradations for radiosonde and aircraft observations of zonal and meridional winds, as well as the general degradation for wind radar.

Finally, another reason may also explain this difference in behaviour between both periods, i.e. "Universal Error Calibration Effect". Indeed, when evaluating the SDFGDs on current observations by separating the two periods, it is noted that the fixed calibration tends to slightly overestimate the errors in summer, especially for radiosondes, aircraft and some infrared and microwave observations. This has the effect of giving less weight to these instruments and leaving the analysis open to modification

by the IRS, since its observation errors are underestimated. Conversely, for the winter period, the evaluation of the SDFGDs shows that the fixed calibration tends to underestimate the errors of radiosondes, aircraft, wind radars, infrared observations and some microwaves. In this case, more weight is given to these observations during the minimisation and the influence of the IRS observations on the modification of the analysis is thus limited. Nevertheless, taking into account these two behaviours, we can clearly expect that the "real" impact of IRS is between these results obtained for the two periods and that the overall influence of IRS observations is positive for the weather forecast.

## 5. Summary and Conclusions

The objective of this study was to evaluate the potential impact of adding the observations of the future IRS sounder in a mesoscale model such as AROME at Météo-France. To achieve this, we built an OSSE framework for a summer and winter period as follows:

- A Nature Run (NR) has been constructed from the ARPEGE global model corresponding to a long and uninterrupted forecast for 3 months of each period taking into account 1 month of spin-up. Analyses from OSTIA are used to force the sea surface temperature each day. This ARPEGE NR is then used to force the lateral boundary conditions of an AROME NR. The latter has been configured to have vertical levels down to 0.1 hPa and a horizontal resolution of 2.5 km coarser than the assimilating model. In parallel, an ARPEGE Coupling Run (CR) was created with the same configuration as the ARPEGE NR but for which errors have been introduced allowing a different weather behaviour compared to the ARPEGE NR. This ARPEGE CR is used to force the LBCs of the AROME NWP model in its operational configuration for the 3D-Var data assimilation system. This avoids twin model problems.
- The NR AROME is then used to simulate the full current observing system which, once calibrated using an iterative error modification method, provides a set of synthetic observations that can be assimilated. The aim of this method is to obtain equivalent first-guess departure between the OSSE framework and the operational AROME data assimilation. In our case, the calibration takes into account the combined statistics of both periods (summer and winter).
- The OSSE framework was evaluated using the Observing System Experiment (OSE) method. This approach verifies the behaviour of our OSSE against the operational model by comparing forecasts from experiments with and without an observation type. In our case, we compared with and without radar data assimilation. The results show similar behaviour of the OSE forecasts between the operational and our OSSE framework.
- The NR AROME was also used to create the synthetic IRS observations for one out of two pixels (pre-thinning) with the 1960 channels that make up its spectrum. The RTTOV coefficients use pseudo-Hamming apodisation and the simulations were performed for all-sky conditions. The raw radiances (in terms of brightness temperatures) were converted using principal component analysis to 300 pieces of information in spectral space. This resulted in two databases of IRS observations: raw and reconstructed brightness temperatures.

In order to use the new IRS observations, the configuration of the AROME 3D-Var data assimilation system has been adapted:

- The (McNally and Watts 2003) cloud detection scheme has been modified to take into account IRS observations and to identify cloud-contaminated pixels during quality control and to exclude them from the assimilation process. Its detection capability has been evaluated and is found to be similar to that achieved for IASI.
- Independent experiments were conducted to assimilate all 1960 IRS channels in order to diagnose its observation errors. The estimation of the observation error covariance matrix takes into account both periods and was produced from raw and reconstructed brightness temperatures. As expected in this OSSE framework, the variances are slightly underestimated but the behaviour is similar to that of other infrared sounders.
- For the two periods from January 01 to February 29, 2020 and from July 01 to August 31, 2020, a control experiment was carried out assimilating the entire current observation system (CTL) and an IRS experiment (EXP IRS) assimilating the observations from CTL plus IRS observations. For these experiments we use the reconstructed brightness temperatures of 75 channels consisting of 27 temperature sensitive channels, 27 water vapor sensitive channels and 21 from the atmospheric window sensitive to surfaces and to temperature and humidity in the lower troposphere. In this study as a first step we choose to carry out a horizontal thinning of 70 km between the IRS pixels and to assimilate the observations only over sea in clear sky conditions.

3D-Var AROME data assimilation experiments were carried out for CTL and EXP IRS over the study periods with the following main results:

- EXP IRS experiments has about 50 % additional assimilated observation from IRS in addition to current observations. The more cloudy and rainy conditions in winter decrease the number of assimilated infrared observations (IRS, IASI and SEVIRI) and increase the quantity of radar humidity and conventional observations.
- More specifically, in winter we notice about one hundred thousand (over 2 months) fewer IRS observations compared to summer for the channels of the atmospheric window sensitive to surfaces and to temperature and humidity in the lower troposphere.
- The impact of adding IRS has a different impact on the other observations depending on the study period. The impact is greatly positive and significant for all of the current observation system in summer, while it is less so in winter or even negative for certain types of observations, mainly for conventional observations around 300 to 500 hPa and for the entire radar wind profile. The quantity of IRS observations assimilated in winter being lower due to cloudy conditions (difference in quality control with a less advantageous first-guess check), can be an explanation for the difference in impact between the two periods, but not only as we will see later.
- The addition of IRS and the impact on the current observing system modifies the analyses and therefore the initial conditions of subsequent forecasts. Forecast

scores were compared between the CTL and EXP IRS experiments in terms of RMSE according to the AROME NR as verification data. In summer, IRS improves temperature, relative humidity and wind speed forecasts by more than 10 % significantly over the full tropospheric column and up to +48 h. This largely positive impact in summer is less pronounced in winter with a significant improvement in temperature forecasts in the lower and upper troposphere up to +36 h and a degradation around 500 hPa up to +24 h. The relative humidity forecast is significantly improved over the same atmospheric layers and ranges. There is also an improvement in the forecasts of the wind speed in the lower and upper troposphere up to +30 h and a degradation between 300 and 500 hPa.

- The differences in impact on assimilated observations and forecast scores between the summer and winter period can be explained by the difference in cloud cover and precipitation that can occur between summer and winter, thus influencing the amount of infrared observations assimilated and a greater use of radar humidity in case of rain. The production of the current observation system in this OSSE framework can also explain these differences, i.e. "Universal Error Calibration Effect". This has the effect of increasing the impact of the IRS on the forecast in summer and minimising it in winter. We justify this choice of calibration with the aim of being able to apply it to any month of the year. Finally, the OSSE radar study showed that our OSSE tends to globally overestimate the impacts by a factor of 2. In any case, this study shows both the maximum (in summer) and minimum (in winter) impact that IRS will have in our AROME NWP system on weather forecasts, with an overall significantly positive impact.

This study provides an overview of the assimilation of the future IRS infrared sounder into our NWP systems and particularly for mesoscale models with very encouraging results for the significant improvement of our weather forecasts.

This work is carried out in an idealized framework and is subject to several limitations that may influence the results:

- One of the first limitations in this study is the use of "perfect" land surface and sea surface temperature conditions from the ARPEGE CR to force the AROME 3D-Var data assimilation system. This choice simplifies the construction of our OSSE but slightly limits the differences between the AROME forecast and the NR considered here as the real atmospheric state. In future studies we may investigate the use of sea surface temperatures from other conditions.
- In this OSSE, the creation of synthetic observations combines simulations for each type of observation and the perturbation of their errors by a random and uncorrelated Gaussian function. In reality, the observation errors are correlated but further research is needed to take them into account in the data assimilation system.
- The calibration of observations is a tedious and essential step in the development of an OSSE. In this study we have chosen to build it taking into account the two periods in order to make it more widely applicable. Nevertheless, we have noticed that this choice tends to overestimate or underestimate the impact of the observations evaluated. Despite this, this method allows us to consider the maximum and minimum impacts that IRS may have in

our NWP system. In future work we can focus on a single season in order to particularise the calibration to that period.

Finally, some short and medium term perspectives on the development of IRS assimilation in our AROME NWP system are the following:

- Further experiments will be carried out on the basis of this OSSE in order to evaluate the impact of IRS assimilation over land in addition to over oceanic surfaces. For this purpose, a specific selection of IRS channels not sensitive to the surface will be carried out in order to assimilate a larger amount of observations over the whole AROME domain.
- A high spatial density assimilation of IRS channels will also be performed to diagnose horizontal observation error correlations. Based on these results, data assimilation studies of IRS observations will be carried out with different thinning.
- In the coming years, the AROME 3D-Var data assimilation system will be replaced by a 4D-EnVar version. This evolution will allow to take into account temporal aspects and to better constrain temperature, humidity and wind fields. This will be particularly beneficial for IRS observations usage and their high temporal frequencies. In addition, the development of ensembles for assimilation and the use of a 4D-EnVar for AROME will make it easier to define background errors, thus including hydrometeors (clouds) more easily. Finally, the high spatial and temporal frequency of IRS will enable progress to be made on the assimilation of radiances for the characterisation of aerosols in AROME, for example. In the years to come, the objective is to include aerosols and atmospheric composition online in the framework of an Earth system model, thus making it possible to develop coupled meteorological and chemical assimilation.

## Acknowledgements

Olivier Coopmann is funded by the EUMETSAT Research Fellowship programme. The authors thank Bertrand Théodore of EUMETSAT for his support and guidance and Vincent Guidard for his expertise on infrared observations.

## References

- Atkinson, N. (2022). *IRSP Top Level Design*. Met Office, Exeter, UK.
- Atlas, R., Hoffman, R. N., Ma, Z., Emmitt, G. D., Wood Jr, S. A., Greco, S., Tucker, S., Bucci, L., Annane, B., Hardesty, R. M., et al. (2015). Observing system simulation experiments (OSSEs) to evaluate the potential impact of an optical autocovariance wind lidar (OAWL) on numerical weather prediction. *Journal of Atmospheric and Oceanic Technology*, 32(9):1593–1613.
- Bessho, K., Owada, H., Okamoto, K., and Fujita, T. (2021). Himawari-8/9 follow-on satellite program and impacts of potential usage of hyperspectral ir sounder. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 1507–1510. IEEE.
- Boukabara, S.-A., Moradi, I., Atlas, R., Casey, S. P., Cucurull, L., Hoffman, R. N., Ide, K., Krishna Kumar, V., Li, R., Li, Z., et al. (2016). Community global observing system simulation experiment (OSSE) package (cgop): description and usage. *Journal of Atmospheric and Oceanic Technology*, 33(8):1759–1777.
- Bouyssel, F., Berre, L., Bénichou, H., Chambon, P., Girardot, N., Guidard, V., Loo, C., Mahfouf, J.-F., Moll, P., Payan, C., et al. (2022). The 2020 global operational nwp data assimilation system at Météo-France. In *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. IV)*, pages 645–664. Springer.
- Brousseau, P., Seity, Y., Ricard, D., and Léger, J. (2016). Improvement of the forecast of convective activity from the AROME-France system. *Quarterly Journal of the Royal Meteorological Society*, 142(699):2231–2243.
- Burrows, C. (2019). Assessment and preliminary assimilation of radiances from GIIRS on FY-4A. In *2019 Joint Satellite Conference*. AMS.
- Chambon, P., Mahfouf, J.-F., Audouin, O., Birman, C., Fourrié, N., Loo, C., Martet, M., Moll, P., Payan, C., Pourret, V., et al. (2022). Global observing system experiments within the météo-france 4d-var data assimilation system. *Monthly Weather Review*.
- Coopmann, O., Fourrié, N., and Guidard, V. (2022). Analysis of MTG-IRS observations and general channel selection for numerical weather prediction models. *Quarterly Journal of the Royal Meteorological Society*.
- Courtier, P., Freyrier, C., Geleyn, J.-F., Rabier, F., and Rochas, M. (1991). The ARPEGE project at Météo—France. *ECMWF Seminar Proceedings*, (7):193–231.
- Desroziers, G., Berre, L., Chapnik, B., and Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3385–3396.
- Duruisseau, F., Chambon, P., Guedj, S., Guidard, V., Fourrié, N., Taillefer, F., Brousseau, P., Mahfouf, J.-F., and Roca, R. (2017). Investigating the potential benefit to a mesoscale NWP model of a microwave sounder on board a geostationary satellite. *Quarterly Journal of the Royal Meteorological Society*, 143(706):2104–2115.
- Eresmaa, R. (2020). ECMWF aerosol and cloud detection software user guide. volume NWPSAF-EC-UD-015. Numerical Weather Prediction Satellite Application Facilities.
- Erico, R. M., Yang, R., Masutani, M., and Woollen, J. (2007). The use of an OSSE to estimate characteristics of analysis error. *Meteorologische Zeitschrift*, 16:695–708.
- Erico, R. M., Yang, R., Privé, N. C., Tai, K.-S., Todling, R., Sienkiewicz, M. E., and Guo, J. (2013). Development and validation of observing-system simulation experiments at NASA's global modeling and assimilation office. *Quarterly Journal of the Royal Meteorological Society*, 139(674):1162–1178.
- Fourrié, N., Bresson, É., Nuret, M., Jany, C., Brousseau, P., Doerenbecher, A., Kreitz, M., Nuissier, O., Sevault, E., Bénichou, H., et al. (2015). AROME-WMED, a real-time mesoscale model designed for the HYMEX special observation periods. *Geoscientific Model Development*, 8(7):1919–1941.
- Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., ... & Fujita, T. (2018). Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, 144(713), 1218–1256.
- Guedj, S., Guidard, V., Ménétrier, B., Mahfouf, J.-F., and Rabier, F. (2014). *Future benefits of high-density radiance data from MTG-IRS in the AROME fine-scale forecast model Final Report*.
- Hoffman, R. N. and Atlas, R. (2016). Future observing system simulation experiments. *Bulletin of the American Meteorological Society*, 97(9):1601–1616.
- Holmlund, K., Grandell, J., Schmetz, J., Stuhlmann, R., Bojkov, B., Munro, R., Lekouara, M., Coppens, D., Viticchie, B., August, T., et al. (2021). Meteosat third generation (MTG): Continuation and innovation of observations from geostationary orbit. *Bulletin of the American Meteorological Society*, 102(5):E990–E1015.
- Kelly, G. and Thépaut, J.-N. (2007). Evaluation of the impact of the space component of the global observing system through observing system experiments. *ECMWF Newsletter*, 113:16–28.
- Lee, J.-W. and Min, K.-H. (2021). The effect of various observations on the forecast accuracy of a numerical model based on ose. In *AGU Fall Meeting Abstracts*, volume 2021, pages A45J–1986.
- Lupu, C., August, T., Coppens, D., Hultberg, T., and McNally, T. (2021). The assimilation of EUMETSAT reconstructed radiances for iasi data compression.
- Ma, Z., Riishøjgaard, L. P., Masutani, M., Woollen, J. S., and Emmitt, G. D. (2015). Impact of different satellite wind lidar telescope configurations on ncep gfs forecast skill in observing system simulation experiments. *Journal of Atmospheric and Oceanic Technology*, 32(3):478–495.
- Martet, M., Brousseau, P., Wattrelot, E., Guillaume, F., and Mahfouf, J.-F. (2022). Operational assimilation of radar data from the european eumetnet programme opera in the Météo-France convective-scale model AROME. In *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. IV)*, pages 629–644. Springer.
- Masutani, M., Campana, K. A., Lord, S. J., and Yang, S.-K. (1999). Note on cloud cover of the ecmwf nature run used for OSSE/NPOESS project.

- Masutani, M., Schlatter, T., Errico, R., Stoffelen, A., Andersson, E., Lahoz, W., Woollen, J., Emmitt, G., Riishøjgaard, L. P., and Lord, S. (2010a). Observing system simulation experiments. *Data Assimilation: Making Sense of Observations*, pages 647–679.
- Masutani, M., Schlatter, T. W., Errico, R. M., Stoffelen, A., Andersson, E., Lahoz, W., Woollen, J. S., Emmitt, G. D., Riishøjgaard, L.-P., and Lord, S. J. (2010b). Observing system simulation experiments. In *Data Assimilation*, pages 647–679. Springer.
- Matricardi, M. (2010). A principal component based version of the rtov fast radiative transfer model. *Quarterly Journal of the Royal Meteorological Society*, 136(652):1823–1835.
- McCarty, W., Carvalho, D., Moradi, I., and Privé, N. C. (2021). Observing system simulation experiments investigating atmospheric motion vectors and radiances from a constellation of 4–5- $\mu$  m infrared sounders. *Journal of Atmospheric and Oceanic Technology*, 38(2):331–347.
- McCarty, W., Errico, R. M., and Gelaro, R. (2012). Cloud coverage in the joint OSSE nature run. *Monthly weather review*, 140(6):1863–1871.
- McGrath-Spangler, E. L., McCarty, W., Privé, N., Moradi, I., Karpowicz, B. M., and McCorkel, J. (2022). Using osse to evaluate the impacts of geostationary infrared sounders. *Journal of Atmospheric and Oceanic Technology*.
- McNally, A. and Watts, P. (2003). A cloud detection algorithm for high-spectral-resolution infrared sounders. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 129(595):3411–3423.
- Okamoto, K., Owada, H., Fujita, T., Kazumori, M., Otsuka, M., Seko, H., Ota, Y., Uekiyo, N., Ishimoto, H., Hayashi, M., et al. (2020). Assessment of the potential impact of a hyperspectral infrared sounder on the himawari follow-on geostationary satellite. *SOLA*.
- Privé, N., Errico, R., and Tai, K.-S. (2013a). The influence of observation errors on analysis error and forecast skill investigated with an observing system simulation experiment. *Journal of Geophysical Research: Atmospheres*, 118(11):5332–5346.
- Privé, N., Errico, R., and Tai, K.-S. (2013b). Validation of the forecast skill of the global modeling and assimilation office observing system simulation experiment. *Quarterly Journal of the Royal Meteorological Society*, 139(674):1354–1363.
- Privé, N., Xie, Y., Koch, S., Atlas, R., Majumdar, S. J., and Hoffman, R. N. (2014). An observing system simulation experiment for the unmanned aircraft system data impact on tropical cyclone track forecasts. *Monthly Weather Review*, 142(11):4357–4363.
- Privé, N. C., Xie, Y., Woollen, J. S., Koch, S. E., Atlas, R., and Hood, R. E. (2013c). Evaluation of the earth systems research laboratory’s global observing system simulation experiment system. *Tellus A: Dynamic Meteorology and Oceanography*, 65(1):19011.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V. (2011). The AROME-France convective-scale operational model. *Monthly Weather Review*, 139(3):976–991.
- Stark, J. D., Donlon, C. J., Martin, M. J., and McCulloch, M. E. (2007). Ostia: An operational, high resolution, real time, global sea surface temperature analysis system. In *Oceans 2007-europe*, pages 1–4. IEEE.
- Stewart, LM and Dance, Sarah L and Nichols, Nancy K and Eyre, JR and Cameron, J. (2014). Estimating interchannel observation-error correlations for IASI radiance data in the Met Office system *Quarterly Journal of the Royal Meteorological Society*, 140(681):1236–1244.
- Vittorioso, F., Guidard, V., and Fourrié, N. (2021). An infrared atmospheric sounding interferometer–new generation (IASI-NG) channel selection for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 147(739):3297–3317.
- Wang, H., Huang, X.-Y., and Chen, Y. (2013). An observing system simulation experiment for the impact of MTG candidate infrared sounding mission on regional forecasts: system development and preliminary results. *International Scholarly Research Notices*, 2013.
- Wattrelot, E., Caumont, O., and Mahfouf, J.-F. (2014). Operational implementation of the 1D+ 3D-var assimilation method of radar reflectivity data in the arome model. *Monthly Weather Review*, 142(5):1852–1873.
- Yang, J., Zhang, Z., Wei, C., Lu, F., and Guo, Q. (2017). Introducing the new generation of chinese geostationary weather satellites, fengyun-4. *Bulletin of the American Meteorological Society*, 98(8):1637–1658.
- Yin, R., Han, W., Gao, Z., and Di, D. (2020). The evaluation of FY4A’s geostationary interferometric infrared sounder (GIIRS) long-wave temperature sounding channels using the grapes global 4D-var. *Quarterly Journal of the Royal Meteorological Society*, 146(728):1459–1476.