Final Report

Assessment of the operational potential of assimilating IASI (Metop) L2 in a regional model

by Bruna Barbosa Silveira[1], Vincent Guidard[1],Nadia Fourrié [1], Philippe Chambon[1], Pierre Brousseau[1], Jean François Mahfouf[1], Patrick Moll[1], Thomas August[2], Tim Hultberg[2]

[1]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[2]EUMETSAT, Darmstadt, Germany

Eumetsat study EUM/CO/17/4600001975/TA

August 29, 2019

Contents

1 Introduction **5**

2 Framework - Model and L2 product **7**

2.0.1 AROME-France and data assimilation 7

2.0.2 Metop combined retrieval L2 product 8

3 Task 1.1 - Evaluation of IASI products **9**

3.1 Task 1.i Evaluation of IASI products, uncertainties and acceptance criteria 9

3.1.1 Using Quality Control Indicator (QCI) to filter the data 9

3.1.2 Evaluation of L2 products (v6.3) against AROME forecasts 11

3.1.3 Evaluation of L2 products (v6.4) against AROME forecasts 14

4 Task 1.2 and 1.3 - Definition of experiments setup and the forecast skills assessment **17**

4.1 Task 1.ii Definition of the assimilation experiments 17

4.1.1 Experiments 17

4.2 Task 1.iii Definition of the forecast skills assessment 21

5 Task 2 - Assimilation experiments and evaluation **23**

5.1 Task 2.i - Impact on the statistics of the other assimilated observations 23

5.2 Task 2.ii - Forecast verification 31

5.2.1 Upper-air verification and surface verification 31

5.2.2 Precipitation Verification 36

6 Task 3 - Feed-back on the EARS-IASI L2 products **39**

6.1 Conclusion 39

6.2 Way Forward 40

Appendices **43**

A Statement of work **45**

B Additional L2 product Statistics **49**

B.1 Monthly variation of L2 product 49

Chapter 1

Introduction

The satellite data have been one of the main sources of observations used in the Numerical Weather Prediction (NWP) models, these data have been used in the NWP together with other observations such as radiosoundings, synop data, radar and others. In the last decades, the satellite data have been assimilated directly in the NWP models (radiance/brightness temperature), however these data can also be assimilated as a retrieval profile. The Advanced Microwave Sounding Unit (AMSU)-A and B, Microwave Humidity Sounder (MHS), High Resolution Infrared Radiation Sounder (HIRS), Advanced Techonology Microwave Sounder (ATMS), Atmospheric Infrared Sounder (AIRS), Cross-track Infrared Sounder (CRIS) and Infrared Atmospheric Sounding Interferometer (IASI) are the main sensors that have been used in the operational NWP systems.

IASI is a hyperspectral sensor with 8461 channels, it is capable to provide humidity and temperature profiles from the atmospheric emission spectra, these profiles are retrieved with a high accuracy and vertical resolution (EUMETSAT, 2017). IASI sensor is onboard Metop-A, B and C satellites. For a NWP purpose, a subset of IASI channels are disseminated (Collard, 2007) on the Global Telecommunication System (GTS), which reduces the redundancy in information and the computational cost in use of all IASI channels in the NWP.

The former version of Applications of Research to Operations at MEsoscale (AROME) model was able to assimilate IASI channels. However, the current AROME version has had the top model change from 1 to 10 hPa (Brousseau et al., 2016). As a result, the quality of the simulation of channels having a strong contribution from the atmosphere above 10 hPa decreased. In this way, these channels were removed from the assimilation process.

IASI L2 products from the EUMETSAT Advanced Retransmission Service (EARS) are statistical retrievals (Piece-Wise Linear Regression - $PWLR^3$) which combine information from IASI and microwave sensors, on board Metop satellites. The products are generated by EUMETSAT from direct broadcast at local acquisition centres, such as Lannion (Satellite Meteorological Centre of Météo-France) in Brittany, France. The products are available to regional users within a maximum of 30 minutes from sensing time. The L2 provide temperature and humidity information on the whole atmosphere, with a high vertical resolution (109 levels below 10hPa).

The goal of this study is to assess the potential benefit and study the practicalities of assimilating L2 temperature and humidity profiles in a regional model in replacement of IASI, AMSU-A and MHS radiances. To target the objectives some tasks were executed. The evaluation of the level 2 products was performed with AROME-France short-range forecasts, which are used as background state in the assimilation process. The configuration of the assimilation experiments using the Level 2 products in AROME-France in replacement of level 1 radiances were then defined. Finally the assessment of the assimilation experiments was performed by comparing Baseline, without Metop satellite sounders, with L1 and L2 assimilation experiments.

The document is divided as follows: the second chapter describes the framework (AROME model and L2 product), the third chapter presents the first task of the project which consisted in evaluating the L2 quality against the AROME model, this evaluation will help for the constructing of the data assimilation experiments. The fourth chapter presents the defenition of experiments setup and the metrics used to evaluated these experiments. The fifth chapter shows the main results. Finally, the sixth chapter presents the conclusions and forward way of this study. The appendix presents an additional set of figures with the monthly evaluation of the L2 product.

Chapter 2

Framework - Model and L2 product

2.0.1 AROME-France and data assimilation

AROME is the operational convective-scale model at Météo-France since 2008 (Seity et al., 2011). In the current AROME version, the horizontal and vertical resolutions are 1.3 km and 90 levels, ranging from 5 m to 10 hPa. Figure 2.1 shows the AROME orography and the domain, which contains 1440 x 1536 points on a Lambert projection centered at 47.5 °N and 2 °E over France (Brousseau et al., 2016).

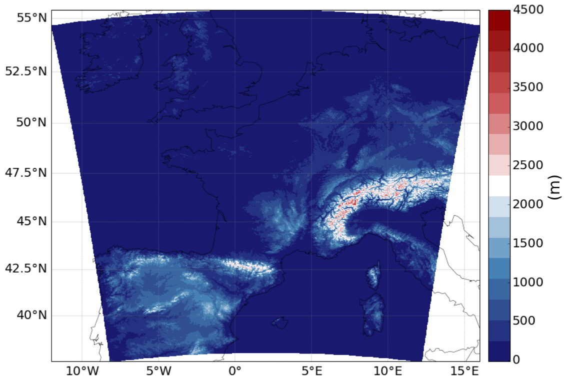


Figure 2.1: AROME domain and orography in shaded.

The initial conditions of the AROME model are provided by a 3D-Var assimilation scheme, which has one-hour assimilation cycle and one-hour assimilation time window (± 30 minutes). Figure 2.2 presents the AROME assimilation cycle scheme, where it is possible to notice that a long forecast is run at 00 UTC, the same configuration is found at 06, 12 and 18 UTC. The forecast range at 00 and 12 UTC is 48 hours and at 06 and 18 UTC is 42 hours. The boundary conditions are given by the forecast fields from the French global model, *Action de Recherche Petite Echelle Grande Echelle* (ARPEGE).

In the system, observations from different sources are assimilated, such as radar measurements (Doppler wind and relative humidity retrieved from reflectivity), surface stations, buoys, ship, aircraft, wind profilers, radiosoundings and satellite observations. The satellite observations include data from infrared and microwave sensors on board geostationary and polar-orbiting satellites. The sensors operationally assimilated in AROME are AMSU-A on board Metop-A and B, NOAA-15, 18 and 19 and AQUA; AMSU-B (MHS) on board NOAA-18 and 19 and Metop-A and B, ATMS on board NPP, SSMIS

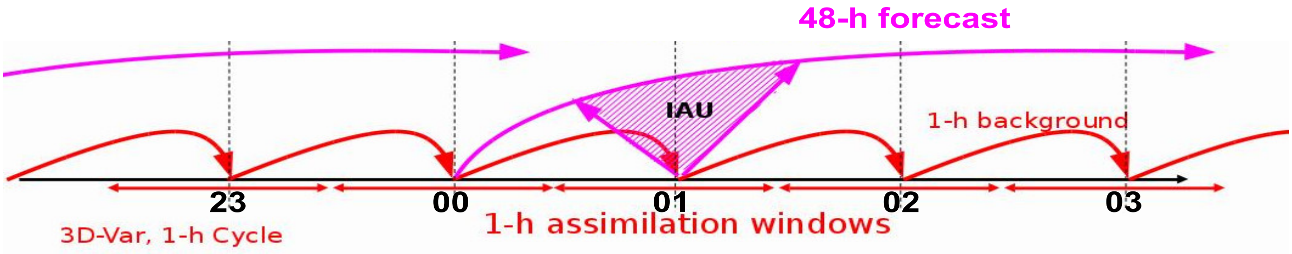


Figure 2.2: AROME assimilation cycle scheme. Adapted from Andrey-Andrés (2017).

(DMSP-17 and 18), GMI (GPM), IASI on board Metop-A and B SEVIRI from Meteosat 11 and scatterometer (Metop-A and B). The GNSS data from ground-based stations are also assimilated.

Some observations are bias corrected and thinned in the AROME system due to misrepresentation of their error and information redundancy. The bias correction coefficients applied to the radiance data come from ARPEGE, except the ones applied to SEVIRI, which are calculated in AROME, it is not available in ARPEGE. Relative humidity profiles from radiosoundings are not assimilated above 300 hPa.

In this study the experiments have been carried out with the same AROME version introduced previously, however some modifications have been applied in order to configure appropriately the system. These differences are shown in detail in the next chapter.

2.0.2 Metop combined retrieval L2 product

The L2 product comes from a statistical retrieval, which utilises IASI toghether with microwave sensors (AMSU-A and MHS), on board Metop satellites, in a synergistic single retrieval. The L2 operational processor and its components were presented in EUMETSAT (2017). This product contains atmospheric profiles of pressure, temperature, water vapour mixing ratio and ozone, some surface parameters (surface temperature, surface emissivity at 10 wavenumbers, surface mean elevation in the pixel and standard deviation of surface elevation in the pixel) and information about the profile quality.

In this study, the profiles of temperature, pressure and water vapour mixing ratio, surface mean elevation in the pixel and the quality control indicator (QCI) for temperature and humidity were used for the assessment and for assimilating these data. The evaluation performed in this work includes the period from August 2017 to February 2018 and May, 15th 2018 to September, 2018. Only the L2 data from observations locally received in Lannion in real time. The L2 data from Metop-A and Metop-B are available for AROME data assimilation from 08 UTC to 12 UTC (AM), and from 19 UTC to 23 UTC (PM) only for Metop-B.

Chapter 3

Task 1.1 - Evaluation of IASI products

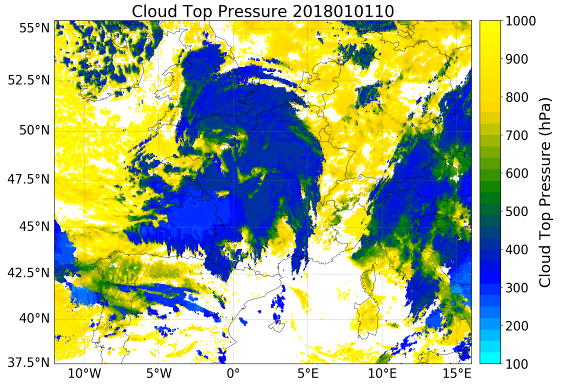
3.1 Task 1.i Evaluation of IASI products, uncertainties and acceptance criteria

The L2 temperature and humidity profiles were compared with the 1 hour forecast from operational AROME model (L2 minus AROME prior state), which is the background for data assimilation. After this, the mean difference and standard deviation of difference were calculated using the following pressure (hPa) bins limits: 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025 and 1050.

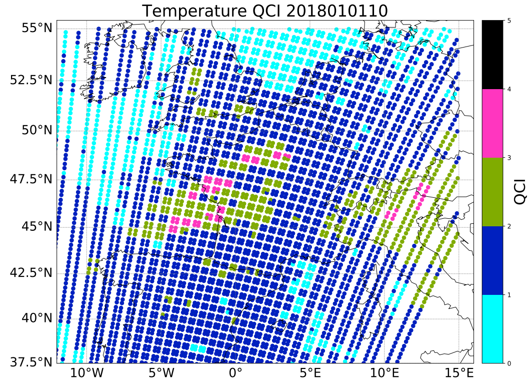
3.1.1 Using Quality Control Indicator (QCI) to filter the data

Correlation between QCI and cloud coverage

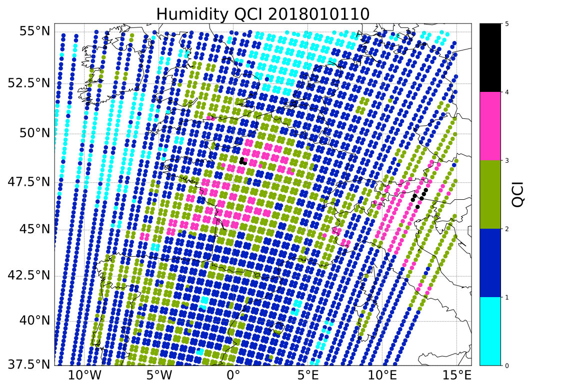
One important information provided with the L2 data is the temperature, humidity and cloud Quality Control Indicator (QCI). The humidity (water vapour) QCI is an uncertainty estimate of surface air water-vapour (it is give in dew point temperature) and for temperature profiles the QCI is an uncertainty estimate in the lower troposphere (K), (EUMETSAT, 2017). The regional IASI L2 product also contains the observation (OBS) minus $CALC_{clear}$ in a IASI window channel, namely OmC, which was used here as cloud QCI. These QCI were compared to the Cloud Top Pressure (CTP) data from SEVIRI. Figure 3.1 shows the CTP, the temperature, humidity and cloud QCI on January 1st, 2018 at 10 UTC over AROME domain. Figure 3.1b shows that the highest QCIs values (magenta) for temperature are located over regions with the CTP between 300 and 400 hPa (figure 3.1a), the green points (temperature QCI greater than 2 K) are also over these regions. The regions with most L2 data with temperature QCI values below 1 K (cyan) are preferably over North and Northwest part of the domain, these regions do not present clouds or the clouds tops are below 800 hPa. The QCIs values over regions without cloud (i.e, over Mediterranean Sea) are not the best QCI marks (below 1 for both, temperature and humidity), the QCIs are the second in the QCI scale, below 2 and greater than 1. The humidity and temperature QCI have a very similar pattern, however, the humidity QCIs present a wider area with QCI values less than 3 (green), and these values seem more related to high cloud tops. Figure 3.1d shows the QCI values for cloud information, the regions with the higher values are over areas with CTP less than 400 hPa, it is the same behaviour observed in the QCI of temperature and humidity. The QCI of cloud presents the small values over regions without clouds.



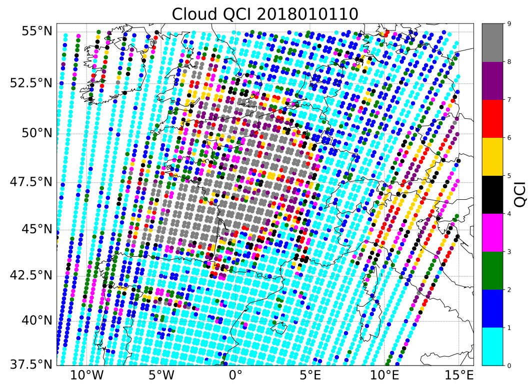
(a) Cloud Top Pressure from SEVIRI



(b) Temperature QCI



(c) Humidity QCI

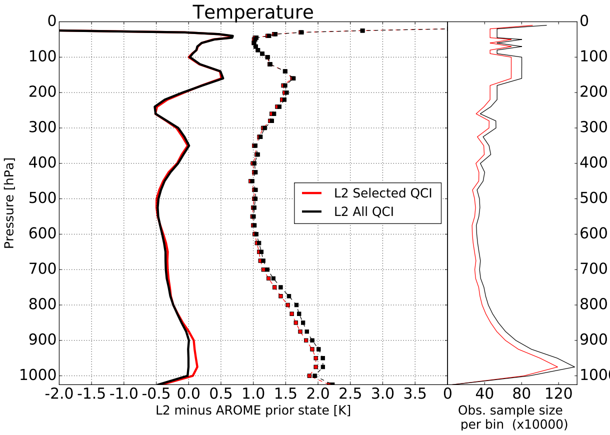


(d) Cloud QCI

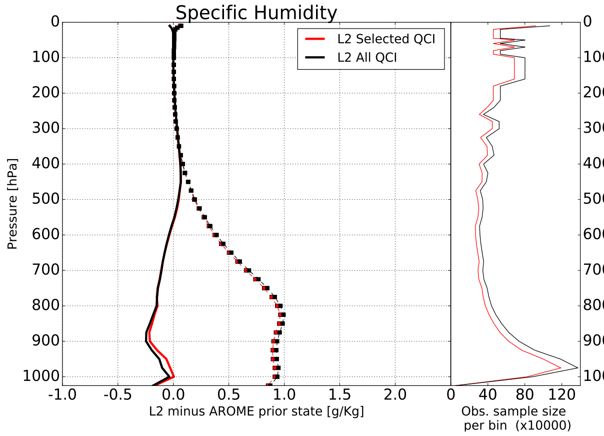
Figure 3.1: a) Cloud Top Pressure Product from SEVIRI in hPa, b)L2 temperature profile Quality Control Indicator (QCI), c) L2 humidity profile QCI and d) L2 cloud profile QCI, all data correspond to January 1st, 2018 at 10 UTC.

Selection of QCI to be selected

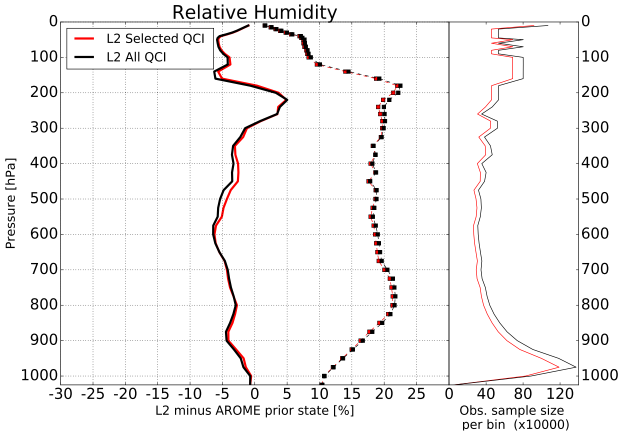
We intend here to perform a pre-selection for the data assimilation based on the QCI. The first evaluation was made using different L2 temperature and humidity QCI combinations. The figure 3.2 shows the temperature and humidity mean differences and standard deviation of the differences for January 2018. The differences between all the L2 QCI values (black lines) and the selected ones (red lines, QCI values less than 2 K for temperature and less than 3 K for temperature of dew point) are not large. The mean differences of all L2 profiles for temperature (figure 3.2a) are slightly smaller than the ones for L2 selected QCI. Figure 3.2 also shows that there is no decrease in the profile quality, e.g., the mean differences and standard deviation of differences values did not increase when the QCI value increase. The number of observations per bin decreased when the QCI values are more restrictive, but it was expected. Therefore, in this study, the L2 data used have QCI values less than 2 K for temperature and the humidity QCI used were less than 3 K for temperature of dew point.



(a) Temperature



(b) Specific Humidity



(c) Relative Humidity

Figure 3.2: L2 minus AROME-F first-guess statistics for all QCI values (black lines) and the QCI selected (red lines). Mean differences are in solid lines, standard deviations of differences are in dashed lines with squares.

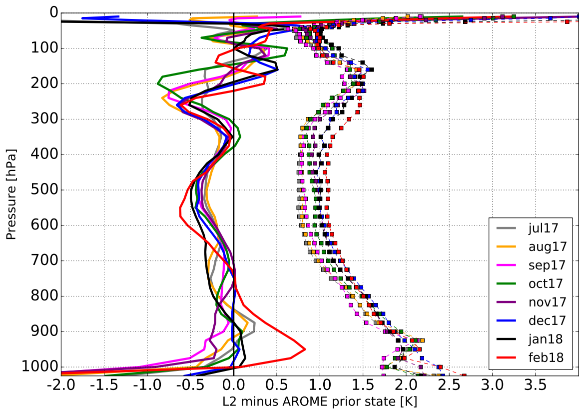
3.1.2 Evaluation of L2 products (v6.3) against AROME forecasts

The assessment was performed for all profiles, independently of the geographic position, and the evaluation was also performed over land, sea and high altitudes separately, in this case the profiles on mixed regions (where the surface mask presents water and continent) were discarded, the geographic classes utilised are the same as those considered in the EUMETSAT (2017).

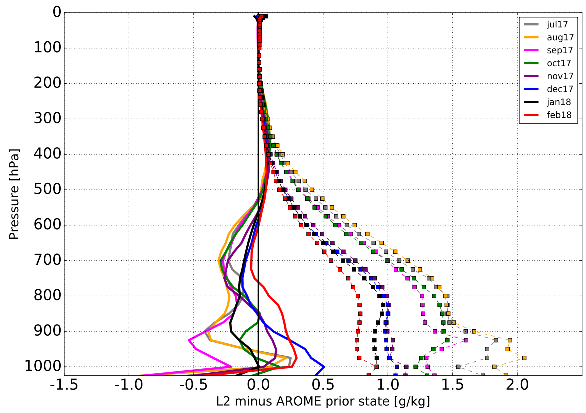
There are differences between the elevation provided with the L2 (surface mean elevation in the pixel) and the orography used in the AROME model, these differences can be greater than 500 meters over elevated terrain. The orography database used in the L2 product and AROME model are different. The profiles located at points where the altitude differences are large should have an atmosphere with different characteristics. In order to exclude these profiles, a filter was applied to eliminate profiles that had an absolute difference between altitudes greater than 25 meters.

Figure 3.3 shows the monthly variation of mean differences and differences of standard deviation from July 2017 to February 2018 for temperature, specific and relative humidity. Mean differences for temperature have a large variation near surface and between 300 and 200 hPa (figure 3.3a), there is also a monthly variation, with the highest value on February, August and September and the lowest values on January and December. February is the most different month, for which the mean difference has a positive peak between 1000 and 900 hPa (more than 0.5 K). All months present the same behaviour for the standard deviation of differences, except near surface where it is possible to notice that there is a monthly variation (1.73 to 2.6 K). The standard deviation of differences increased between 300 and 200 hPa (1 to 1.5 K), in this layer, as shown before, the mean differences also increase.

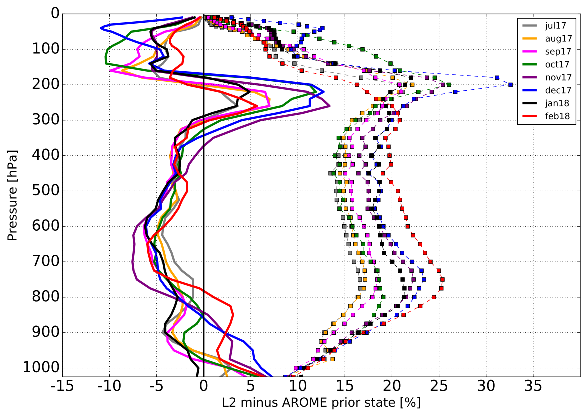
The specific humidity also has a monthly variation (figure 3.3b). The mean difference is negative in most of the cases. In December, below 950 hPa, mean difference is positive, the other months present a negative mean difference near the surface (-0.9 g/kg on August to -0.04 g/kg on October). The standard deviation of differences has strong variation related with the monthly variation, e. g., in August the value is 1.9 g/kg at 975 hPa, which is greater than the ones in February for the same level (0.77 g/kg), for example.



(a) Temperature



(b) Specific Humidity



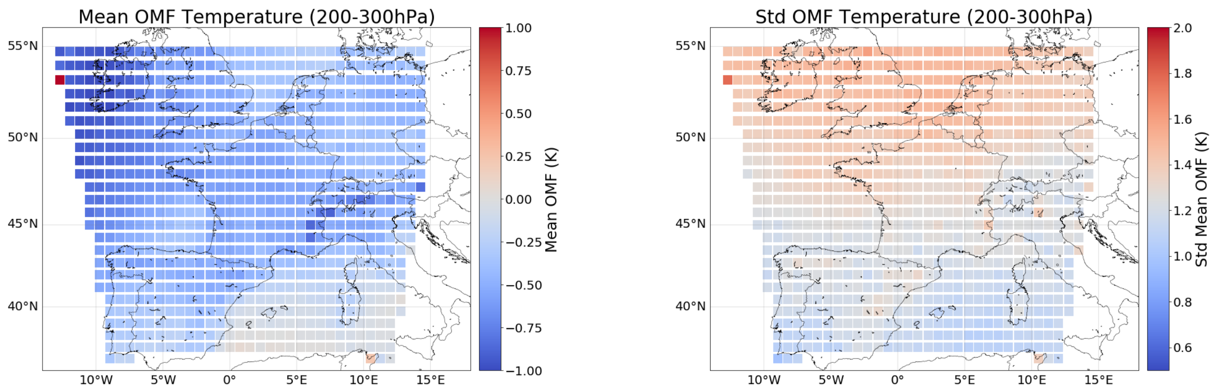
(c) Relative Humidity

Figure 3.3: L2 minus AROME-F first-guess statistics for seven months. Mean differences are in solid lines, standard deviations of differences are in dashed lines with squares. Gray: July, yellow: August/2017, magenta: September/2017, green: October/2017, purple: November/2017, blue: December/2017, black: January/2018 and red: February/2018.

The relative humidity mean difference variation is positive near the surface (except in January), these values vary between 2 and 7 %, above 1000 hPa for some months and above 950 hPa for other months

(November, December and February) the mean differences values change of sign (figure 3.3c). Although, the relative humidity standard deviation of difference variation has an opposite behaviour when compared with specific humidity, e. g., the standard deviation values are large on February (25 % at 750 hPa) and small on July (16 %). The standard deviation between 300 hPa and 200 hPa has a behaviour similar to temperature (except in February, where the standard deviation decreases), values are greater than 30 % on December. The number of observations per bin is almost the same over all periods analysed (approximately 1200000 at 975 hPa and approximately 280000 at 575 hPa). The appendices present monthly variation of the L2 data over land and sea, separately, and for the Metop-A AM and Metop-B AM and PM.

The spatial distribution of mean difference of temperature between levels 200 and 300 hPa were analysed in order to evaluate the spatial characteristics of these values. Figure 3.4a shows that the mean differences change of sign (negative to positive) from Northwest to Southeast. The large differences are located in the Northwest of the domain. It is not possible to confirm which is responsible for this behaviour, the AROME model or the L2 data. Over the Alps region (elevated terrain), the mean difference is large, which could be related with the difference in the orography. A speculation is such characteristic can come from AROME model because of the most part of the flux in this layer comes from the west to east, with boundary conditions coming from ARPEGE. Figure 3.4b shows the spatial distribution of standard deviation of mean difference of temperature between levels 200 and 300 hPa, these values increase from Southeast to Northwest.



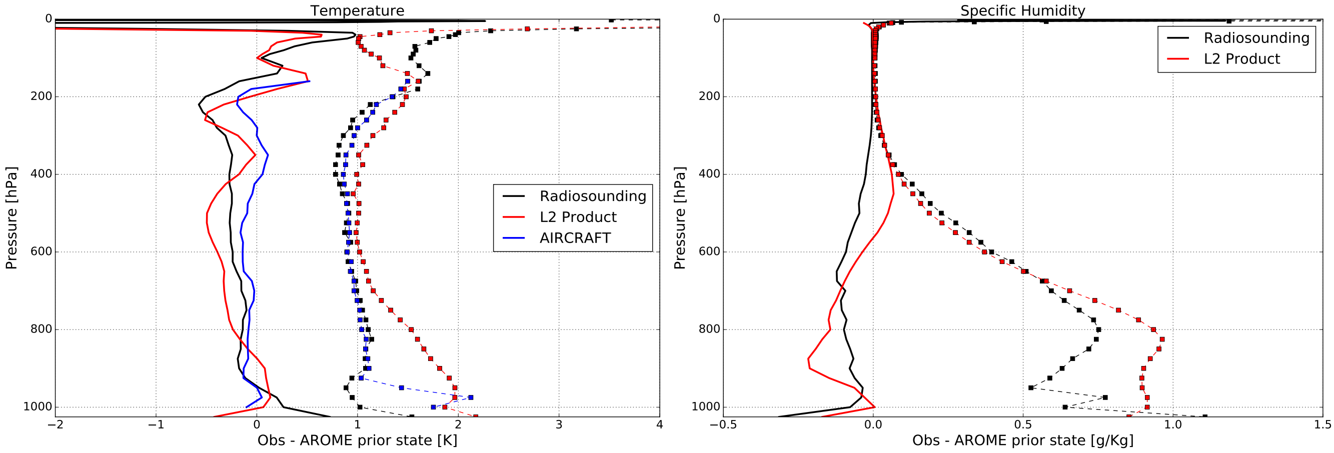
(a) Mean OMF

(b) Standard Deviation OMF

Figure 3.4: a) Mean L2 minus AROME-F first-guess values for January/2018 between 200 and 300 hPa and b) standard deviation of mean L2 minus AROME-F first guess for the same period and layer.

The statistics of L2 departures from January 2018 were compared against the statistics of radiosoundings and aircraft available for assimilation in the baseline experiment (explained in the next section). The humidity data from radiosoundings above 300 hPa are not assimilated into operational AROME, but for a comparison between radiosoundings and L2 these data were also considered. The radiosoundings first guess departure (observation minus one-hour forecast) are used to calculated the mean differences and standard deviation of differences.

Figure 3.6a shows the L2 data (red lines) and radiosoundings (black lines) mean differences for temperature have opposite sign near the surface (Figure 3.6 on left column), the mean difference for L2 data is -0.5 K and the one for radiosoundings is 0.5 K. The three observations types are in good agreement for the mean differences amplitudes. The standard deviation of differences get closer when going up in the atmosphere (above 700 hPa), near the surface, the L2 standard deviation is greater than 2 K and for radiosoundings it is 1.5 K. The AIRCRAFT standard deviation (dashed blue line) has a peak of 2 K near 925 hPa. The figure 3.6b shows the specific humidity mean differences, and it is possible to notice that there is an agreement between the two data over the whole atmosphere. The standard deviation are different, but above 700 hPa the lines start to get closer.



(a) Temperature **(b)** Specific Humidity

Figure 3.5: L2 (red lines), Radiosoundings (black lines) and aircraft (blue line) minus AROME-F first-guess statistics for January/2018. Mean differences are in solid lines, standard deviations of differences are in dashed lines with squares. a) Temperature and b) specific humidity

To assess the diurnal variation, the L2 data were evaluated against the aircraft data (AIREP) and radiosoundings over land (both available for assimilation in the baseline experiment). The L2 data were separated in Metop-A AM and Metop-B AM and PM, the AIREP data were selected for the same time at which the L2 data are available. The radiosoundings was divided in AM (07 UTC to 18 UTC) and PM (19 UTC to 06 UTC). Figure 3.6 shows that there are differences between L2 data AM (green and purple lines) and PM (blue lines) mean difference profiles, for the other two observation types, the differences between the two curves (AM and PM) are not large. The standard deviation profiles are similar for the different L2 data time, the same is noticed for the radiosoundings. The AIREP PM standard deviation profile (red line) exhibit a peak near 950 hPa (almost 2 K).

The L2 agree well with the radiosoundings and aircraft data (AIREP). For the use of the L2 data in the data assimilation more evaluation was performed. The next section will show other assessments realized. The L2 profiles seem suitable for assimilation into AROME system.

3.1.3 Evaluation of L2 products (v6.4) against AROME forecasts

Since March 2018, the L2 product has a new version (v6.4). For the purpose to comparing the two L2 versions this section presents the bias and standard deviation profiles for the L2 product from May, 15th to 30th September 2018.

The temperature mean bias profiles (figure 3.7a) of the version 6.4 present an improvement in the bias in the layer between 400 and 800 hPa. The standard deviation of differences is smaller near surface, the values are closer to 2 K in the v6.4 when compared to the previous version. Figure 3.7b shows the specific humidity profiles, the mean bias profile also has an improvement between 600 and 800 hPa (closer to zero or change the sign), however below in the atmosphere there is a degradation, near 900 hPa the bias has a negative peak in all evaluated months, close to -0.5 g/kg. The standard

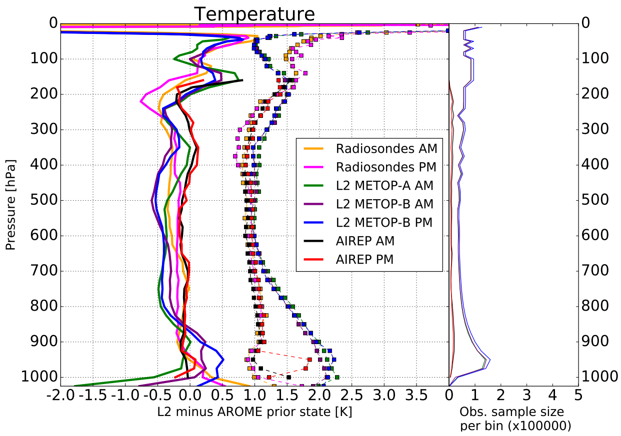
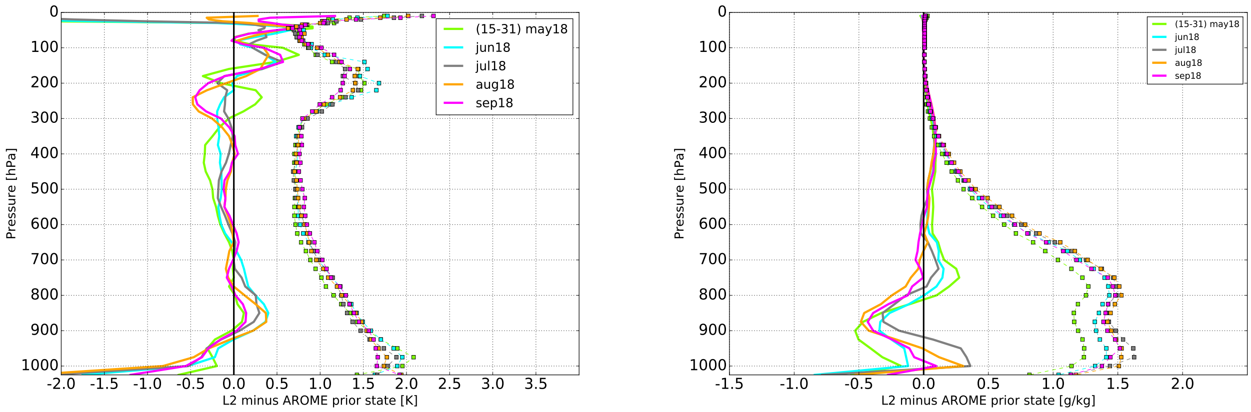


Figure 3.6: L2, radiosoundings and AIREP minus AROME-F first-guess statistics for January/2018 over land. Mean differences are in solid lines, standard deviations of differences are in dashed lines with squares. Yellow: radiosoundings AM, magenta: radiosoundings PM, green: L2 Metop-A AM, purple: L2 Metop-B AM, blue: L2 Metop-B PM, black: AIREP AM red: AIREP PM.

deviation has a small spread in the v6.4 when compared against v6.3. It is really evident near surface.



(a) Temperature

(b) Specific Humidity

Figure 3.7: L2 minus AROME-F first-guess statistics for seven months. Mean differences are in solid lines, standard deviations of differences are in dashed lines with squares.Green: May/2018, cyan: June/2018, gray: Julliet/2018, orange: August/2018 and magenta: September/2018.

The L2 product version 6.4 statistics against AROME model have behaviour similar to L2 product evaluated against radiosoundings in the IASI L2 PPF v6.4 EUMETSAT (2018).

The evaluation of L2 products mean differences and standard deviation of mean differences against the same information from radiosounding and aircraft showed that the L2 products profiles present a good quality in terms of first guess departure (OMF). It means that it is possible to use this data into data assimilation.

Chapter 4

Task 1.2 and 1.3 - Definition of experiments setup and the forecast skills assessment

4.1 Task 1.ii Definition of the assimilation experiments

4.1.1 Experiments

The L2 data evaluation against aircraft and radiosoundings helped to build the data assimilation experiment setting using the L2 data. Table 4.1 presents the experiments and their configuration. The baseline experiment utilises the same observations as the operational AROME-France, except the IASI, AMSU-A and MHS radiances. The control experiment utilises the data used in baseline experiment and IASI radiances from Lannion, AMSU-A and MHS radiances from EUMETSAT, the Metop-A data in the evening is not assimilated. The L2 experiment assimilates the observations used in the baseline experiment and the L2 data. The experiment were performed over three different periods. The first period is the summer, from July, 15th 2017 to September, 15th, 2017 (63 days), the second is the winter, from January, 1st 2018 to February, 28th 2018 (59 days) and the third period is the spring, from May, 15th 2018 to June, 26th 2018 (43 days). The spring period has less days because after June, 26th there was a decontamination in the Metop-B satellite leading to data unavailability.

Experiment	**Configuration**
Baseline	No IASI, AMSU-A and MHS data
Control	Baseline + With IASI from Lannion (only), AMSU-A and MHS data from EUMETSAT. No Metop-A in the evening
L2 Experiment	Baseline + L2 product from Lannion

Table 4.1: Experiments configuration and period.

The L1 and the L2 observations are not assimilated in the same way. It is important to show the main characteristics of assimilating these two observations types.

L1 assimilation details

The IASI, AMSU-A and MHS, together, represent around 5% of the data assimilated in the AROME, this amount depends on the weather conditions (clear or cloudy sky). The IASI, AMSU-A and MHS are assimilated as brightness temperatures. The observation operator used into AROME system is the RTTOV. The thinning applied to the IASI (only pixels 1 ond 3 for IASI) and MHS data is 80 km and to the AMSU-A is 100 km. The figure 4.1a shows the observations position (at least one channel) of IASI (black squares), AMSU-A (red diamonds) and MHS (yellow circles) assimilated in one-hour assimilation cycle (January, 1st 2018 at 10UTC).

In AROME, the IASI channels peaking above 200 hPa and below 600 hPa over land are not assimilated. In the ARPEGE model, 129 IASI channels are assimilated, however because of the AROME top level (10 hPa) some channels are discarded, 44 channels remaining (20 temperature, 20 water vapour and 4 surface channels) over sea and 8 water vapour channels over land peaking in the mid and upper troposphere. Figures 4.1b and 4.1c show the jacobians of IASI channels over land and over sea, respectively. AROME assimilates four AMSU-A channels (5, 6, 7 and 8) and three MHS channels (3, 4 and 5), the jacobians of these channels are shown in figures 4.1d and 4.1e, respectively. The jacobians give the vertical distribution of the channels from these sensors used into AROME data assimilation.

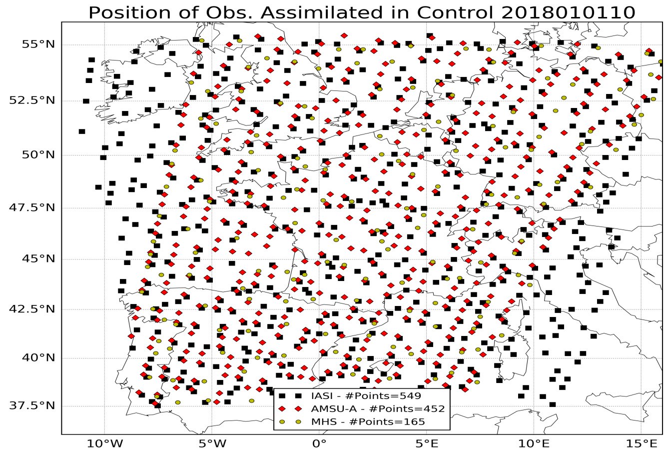
L2 assimilation details

The L2 products, as for the IASI, AMSU-A and MHS L1 data, do not cover the whole domain at each assimilation time, some times there is no L2 product. Figure 4.2 represents one day (January, 1st 2018) of L2 products (without data selection). The figures with black circle are the Metop-B and those with blue circle are the Metop-A. It is possible to notice that the horizontal distribution of the data are not homogeneous in space and time.

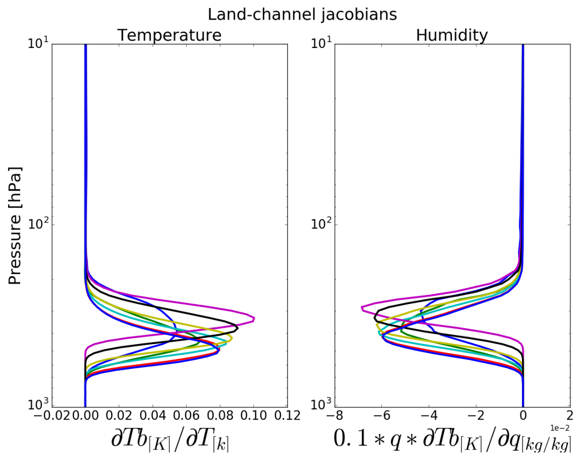
However, a thinning procedure is necessary in L2 data to avoid things as correlation between the observation which are not currently taken into account and representativity issues, the value used for the L2 observations was 160 km. The L2 profiles with the best QCI for temperature, humidity and cloud were select inside the 160 x 160 km box. Figure 4.3a shows the horizontal distribution of L2 product (black squares) assimilated on January, 1st 2018 at 10 UTC, at least one level per profile was assimilated in each square. The vertical resolution of L2 product is 109 levels below 10 hPa (model top), although a vertical thinning was also necessary, the vertical thinning was one level every 3 levels. Figure 4.3b shows the L2 vertical distribution of temperature (on the left) and specific humidity on the right. The blue line represents an observation profile over sea, blue star in the figure 4.3a, and the green line represents a L2 profile over land, green star in the figure 4.3a.

The profiles over sea have information up to 1000 hPa and it is not observed in the profiles over land (figure 4.3b). These differences in the profiles over different regions occur because some decisions were taken based in the L2 evaluation shown before. The L2 product was also evaluated separately over different surfaces (over land, sea and high altitudes). Based on this, three filters were applied in the L2 experiment to avoid some discrepancies found in this L2 data assessment. These filters are shown in table 4.2. Due to inaccuracy in surface properties we decide to discard the lowest part of the profile. When orography is below 1 km, the data above 900 hPa are assimilated, in presence of relief (over 1 km) data from 700 hPa to 10 hPa were assimilated.

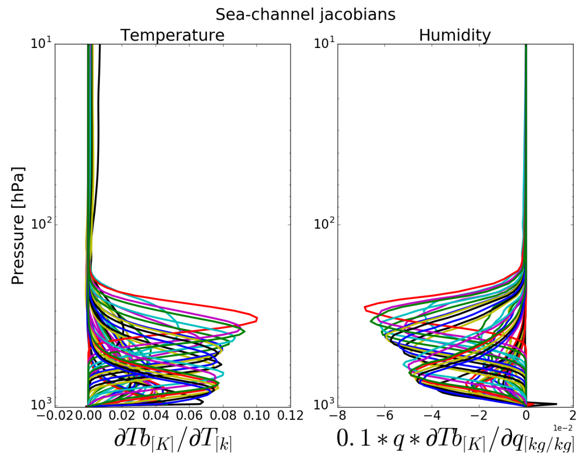
The L2 data will be assimilated as pseudo-radiosoundings and the L2 observation error (σo) must be determined. The radiosounding and aircraft data have the same σo. Their σo was used as a guide for estimating the σo for the L2 data. The figure 4.4a shows the vertical variation of the mean temperature σo (green line) and the standard deviation from radiosoundings (black line), L2 data for all period (blue



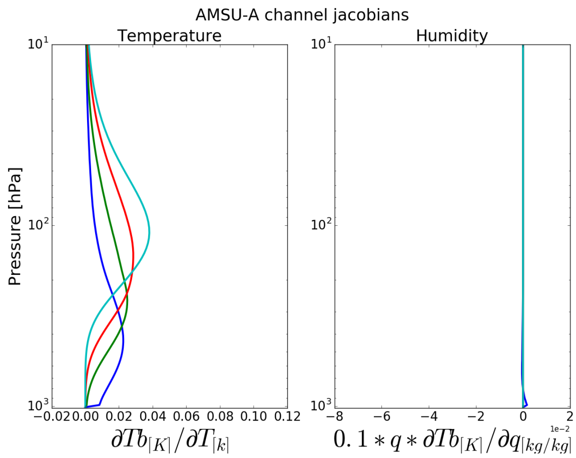
(a) Horizontal distribution of assimilated L1 observation from IASI (blue squares), AMSU-A (magenta diamonds) and MHS (yellow disks).



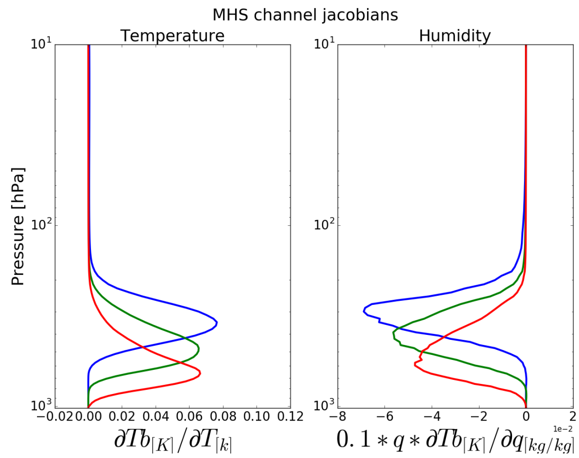
(b) IASI Jacobians over land



(c) IASI Jacobians over sea



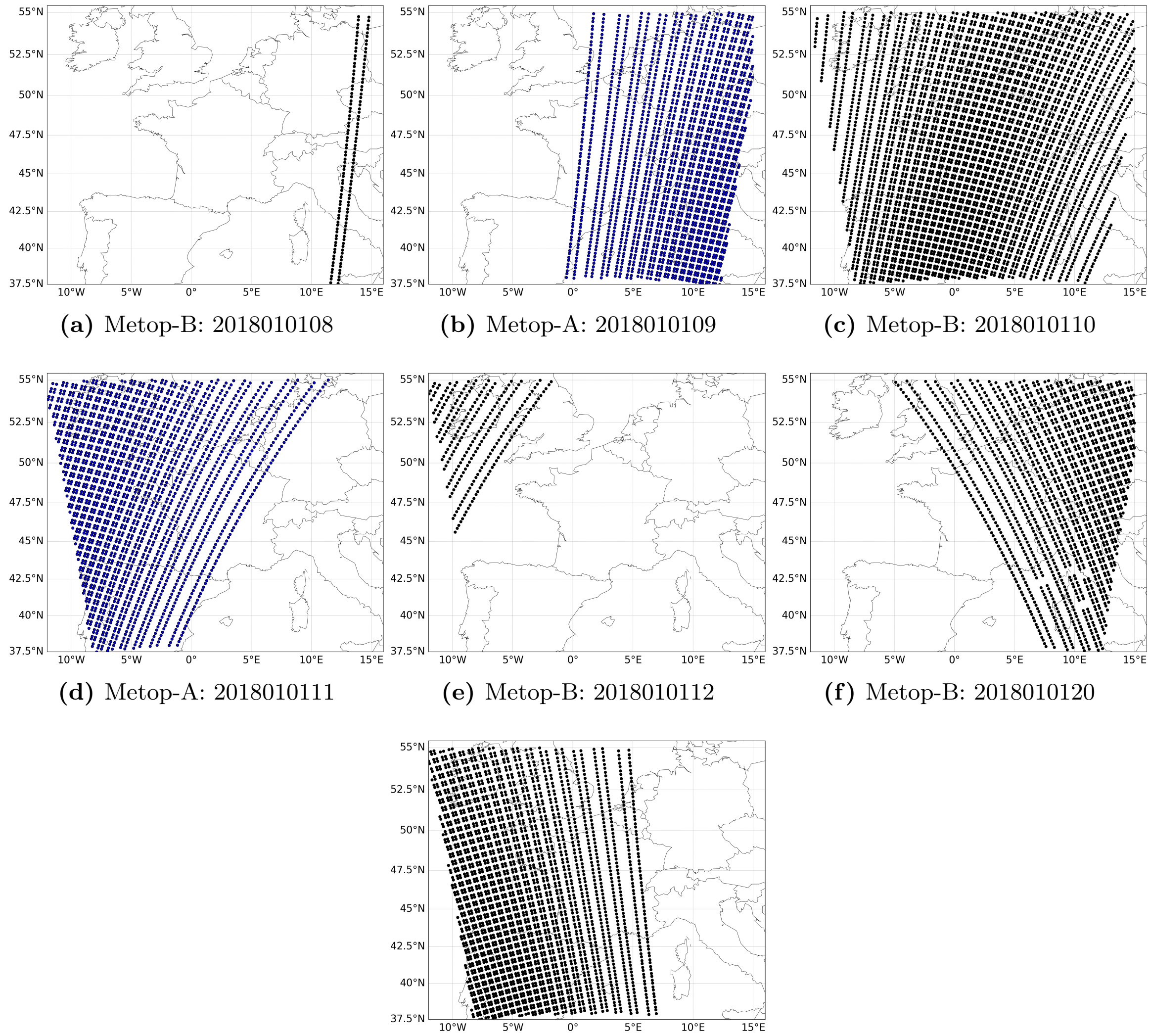
(d) AMSU-A Jacobians



(e) MHS Jacobians

Figure 4.1: a) L1 position (AMSU-A, IASI and MHS) on January, 1st 2018 at 10UTC, b) IASI channels jacobians over land and c) sea, d) AMSU-A jacobians and e) MHS jacobians.

line) and thinned L2 data (red line), noteworthy that the σo for temperature is a fixed profile. The radiosoundings relative humidity mean σo is fixed, to 12 % (figure 4.4c). The specific humidity mean



(a) Metop-B: 2018010108 **(b)** Metop-A: 2018010109 **(c)** Metop-B: 2018010110

(d) Metop-A: 2018010111 **(e)** Metop-B: 2018010112 **(f)** Metop-B: 2018010120

(g) Metop-B: 2018010121

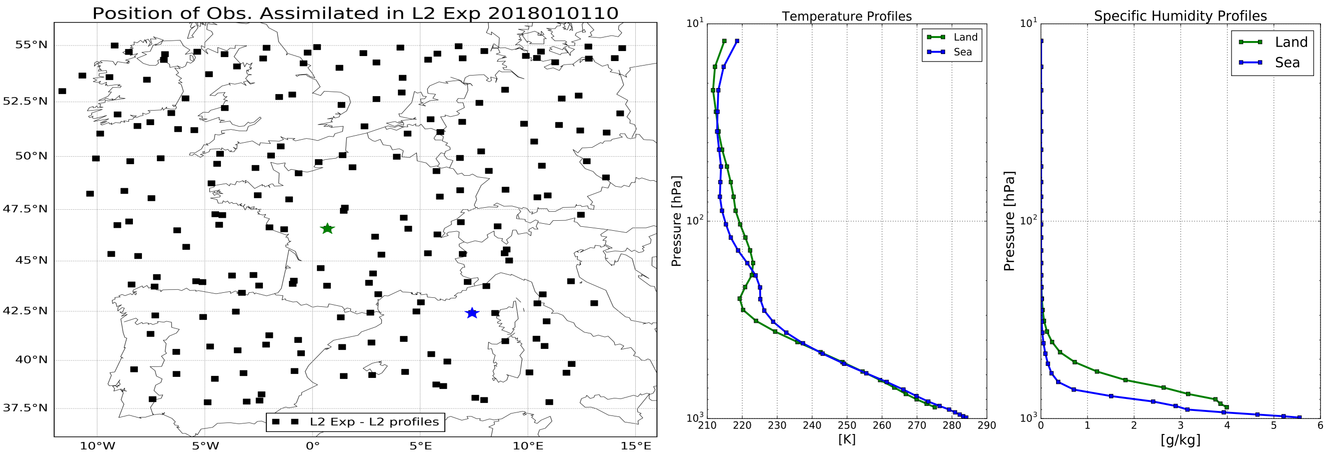
Figure 4.2: Spatial distribution L2 product on January,1st 2018. Blue circles are data from Metop-A and the black ones from Metop-B.

Region	**Filter applied**
Sea	Use data only above level 1000 hPa
Land, orography below 1 km	use data only above level 900 hPa
Land, orography above 1 km	use data only above level 700 hPa

Table 4.2: Filters applied to L2 data.

σo has a standard deviation (figure 4.4b) and its profile is based on the relative humidity observation error σo, the standard deviation of radiosoundings and L2 thinned show an agreement, the both data are from January 2018.

After this evaluation, the estimated error is 1.2 times of the radiosounding error for L2 temperature profiles. Specific Humidity errors assigned to L2 product are 15 % of the relative humidity, in comparison, 12 % relative humidity is usually specified to assimilated radiosoundings measurements.



(a) Spatial distribution **(b)** Vertical distribution

Figure 4.3: a)Spatial distribution of L2 product on January, 1st 2018 at 10 UTC. b)Vertical distribution L2 profiles over land (green lines) and over sea (blue lines).

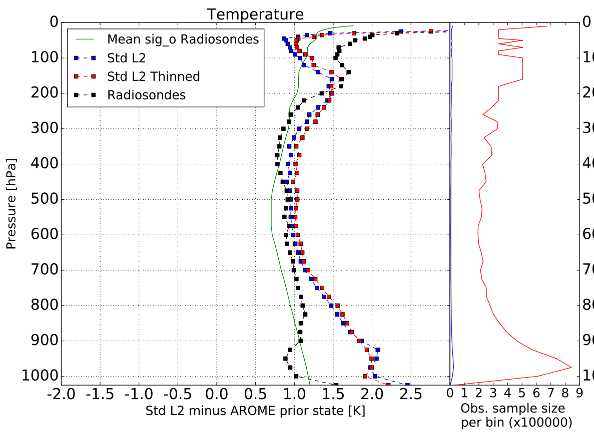
Some evaluations were carried out to investigate the impact of using L2 data instead of L1 into data assimilation and forecasting. The assessment was divided into two parts: first one related to the impact of the use of the L2 product on the assimilation of the other observations and the second related to the forecast verification.

4.2 Task 1.iii Definition of the forecast skills assessment

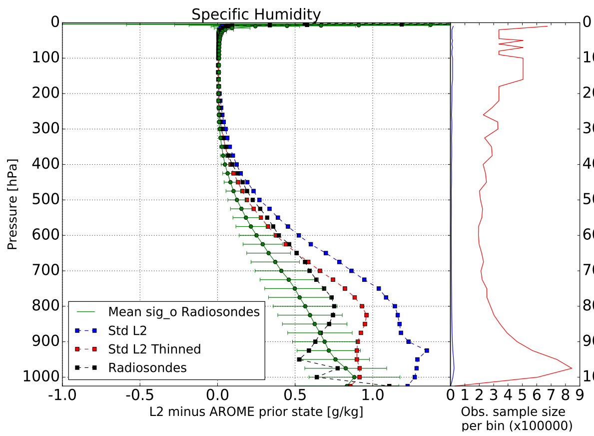
In order to evaluate experiments some metrics will be applied. The observation minus first guess (OMF) and observation minus analysis (OMA) will be used to assess the data assimilation statistics, as well as the standard deviation of OMF and OMA. The analysis increment will be also evaluated.

The forecast skills for the experiments will be assessed using statistics such as root mean square error (RMSE), bias and standard deviation of differences between experiment forecast and reference. The experiments are compared with independent data (radiosoundings and ECMWF analyses). The forecast skills are calculated every 6 forecast hours (0, 6, 12, 18, 24, 30, 36, 42 and 48 hours) and different vertical levels (surface, 1000 hPa, 850 hPa, 700 hPa, 500 hPa, 400 hPa, 300 hPa, 250 hPa and 200 hPa). The skills are computed for geopotencial, temperature, relative humidity and wind (magnitude and direction). The forecast skills are also compared against synop data (sea level pressure, temperature, precipitation, nebulosity, wind (magnitude and direction) and relative humidity) and surface observation from stations over France.

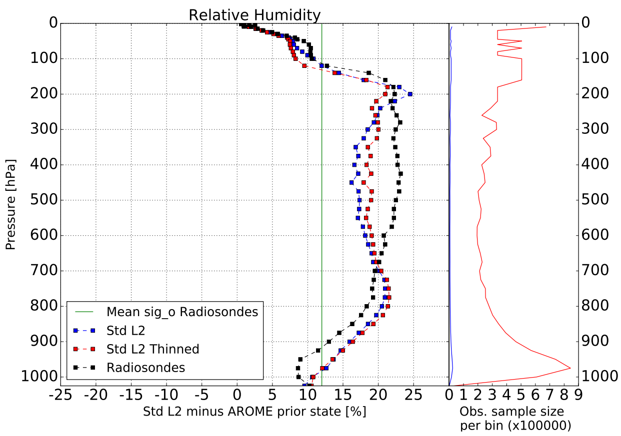
The precipitation accumulation (6 and 24 hours) forecast are evaluated using the frequency bias, detection rate, false alarm rate, Heidke Skill Score (HSS), Brier skill score with neighborhood observation (BSS_NO) and Brier skill (BS), Amodei and Stein (2008). These evaluations were done using different precipitation thresholds (0.5, 2, 5 and 10 mm) and different neighbourhoods (1.3, 20.6, 52.8 and 120.2 km). The forecast range for the precipitation skills is the same used as in the others forecast skills assessments.



(a) Temperature



(b) Specific Humidity



(c) Relative Humidity

Figure 4.4: L2 and radiosondes minus AROME-F first-guess statistics for January 2018 and mean σo profile of radiosondes (green line) for the same period. Standard deviations of differences are in dashed lines with squares.

Chapter 5

Task 2 - Assimilation experiments and evaluation

5.1 Task 2.i - Impact on the statistics of the other assimilated observations

The first evaluation performed was related with the number of observations assimilated in each experiment and period. Table 5.1 shows the amount of observations assimilated separated by observation types (ObsType). The summer and winter experiments assimilate a similar number of observations (when the L2 data and the L1 data of the IASI, AMSU-A and MHS sensors are not considered), except for radar and radiosounding data. Considering the two periods, the control experiment assimilates more radiosounding data and rejects more radar data when compared with baseline experiment, but the relative differences between the baseline and control experiments do not reach 2 ‰ and 3.5 ‰ (column 7th on Table 5.1), respectively. When the L2 experiment is compared with the baseline, a greater difference is observed between the number of radar observations assimilated. IASI, AMSU-A and MHS observations represent 4.85 % of the total data assimilated in the control experiment (considering both periods). The L2 data represents 7.86 % of the observations assimilated in the L2 experiment in the two period verified. The spring period has a similar behaviour to the others two periods.

The assimilation of L2 product also can produce some impact in the mean first guess departure (OMF) and standard deviation of OMF of other observations. The observations more impacted by L2 assimilation were temperature profiles from aircraft and radiosoundings, specific humidity profiles from radiosoundings and relative humidity derived from radar observation. A significance test (t-student with 95% of confidence) was applied in all results, and the reference was always the baseline experiment.

Figure 5.1 and 5.2 show the vertical profiles of mean first guess departure (OMF) and standard deviation of OMF from temperature (aircraft and radiosounding observations), respectively. Generally, the mean OMF of temperature from aircraft vary between -0.14 (winter period) and 0.3 K (spring). Figure 5.1a, b and c show the OMF of temperature from aircraft over summer, winter and spring, respectively. In these figures, the blue lines represent the L2 experiment and the red ones the baseline. In the middle atmosphere, it is possible to notice an improvement in the mean OMF values (it means a reduction in the mean OMF), and this improvement is statistically significant in the levels where there are a blue filled up triangle. The control experiment (black line in the figures) also improve the mean OMF of temperature from aircraft and it is more evident between 800 and 600 hPa (black filled up triangle), figures 5.1a, b and c. Above 400 hPa is not possible to identify a similar behaviour on the three periods. Bellow 925 hPa there are a degradation in the L2 experiment and these are statistically significant, it

Summer (59 days)							
ObsType	Baseline	Control	L2 Exp	(Control-Baseline)	(L2 Exp-Baseline)	Diffcontrol[‰]	DiffL2[‰]
Aircraft	5828210	5828393	5827505	183	-705	0.031	-0.12
SatOb	84368	84304	84292	-64	-76	-0.75	-0.90
Drift Buoys	61500	61503	61500	3	0	0.048	0.0
Radiosoundings	5795847	5801499	5796501	5652	654	0.975	0.112
Profiler	936600	936472	936542	-128	-58	-0.13	-0.06
Scatterometers	75088	75210	75210	122	122	1.624	1.624
Radar	2508420	2505212	2516675	-3208	8255	-1.27	3.290
Synoptic	10923624	10922966	10923776	-658	152	-0.06	0.013
GPS Surface	671056	671089	671187	33	131	0.049	0.195
L2 Profiles	-	-	2503509	-	-	-	
Radiances∗	3904165	3903479	3903876	-686	-289	-0.176	-0.07
IASI, AMSU-A and MHS	-	1768708	-	-	-	-	-
Winter (63 days)							
Aircraft	4569141	4569231	4569384	90	243	0.019	0.053
SatOb	61180	61246	61234	66	54	1.078	0.882
Drift Buoys	48801	48805	48803	4	2	0.081	0.040
Radiosoundings	5349488	5353309	5349056	3821	-432	0.714	-0.08
Profiler	701110	701082	701064	-28	-46	-0.03	-0.06
Scatterometers	68694	68694	68694	0	0	0.0	0.0
Radar	5330942	5320481	5330427	-10461	-515	-1.96	-0.09
Synoptic	10220731	10220727	10221092	-4	361	-0.00	0.035
GPS Surface	605808	605812	605813	4	5	0.006	0.008
L2 Profiles	-	-	2275229	-	-	-	
Radiances∗	3023293	3023789	3023603	496	310	0.165	0.102
IASI, AMSU-A and MHS	-	1179720	-	-	-	-	-
Spring (43 days)							
Aircraft	3664321	3664399	3664547	78	226	0.029	0.085
SatOb	50134	50122	50112	-12	-12	-0.024	-0.024
Drift Buoys	42263	42263	42260	0	-3	0.0	-0.007
Radiosounding	3988110	3989883	3987024	1773	-1086	0.044	-0.027
Profiler	635299	635221	635238	-78	-61	-0.12	-0.096
Scatterometers	51538	51538	51538	0	0	0.0	0.0
Radar	1699764	1703340	1703560	-3576	-3796	-2.1	-2.2
Synoptic	7660253	7660167	7660296	86	-43	0.011	-0.0056
GPS Surface	432098	432103	432107	-5	-9	-0.011	-0.021
L2 Profiles	-	-	1633759	-	-	-	-
Radiances∗	2610160	2610070	2609840	90	320	0.034	0.12
IASI, AMSU-A and MHS	-	1140697	-	-	-	-	-

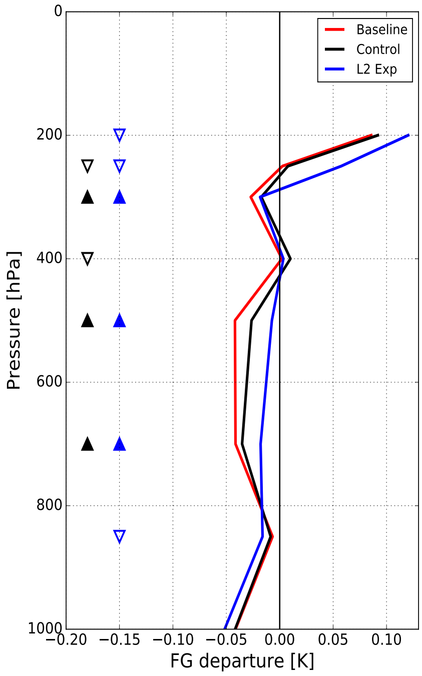
∗ATMS, SSMIS, SEVIRI, GMI and MWHS2

Table 5.1: Observations assimilated per observation type.

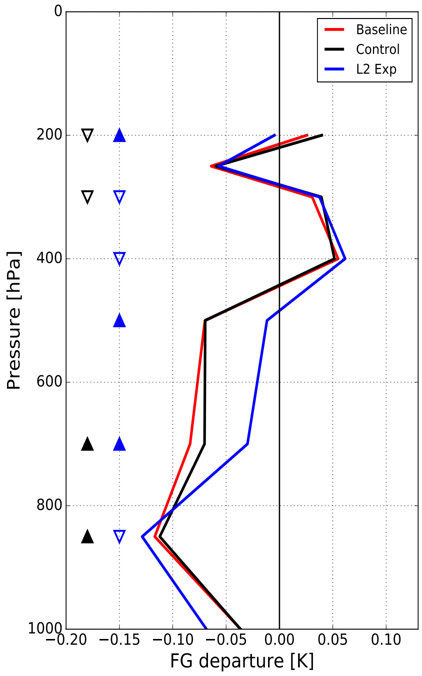
is more evident in the summer and winter.

The OMF of temperature observation from radiosoundings show an improvement in the mean OMF of L2 and control experiment between 800 and 300 hPa over all periods (figure 5.1d, e and f), except on summer where the L2 experiment present a degradation statistically significant (blue empty down triangle) bellow 600 hPa. Above 400 hPa the OMF of temperature from radiosoundings as the ones from aircraft do not present a clear behaviour, it varies between the levels and periods.

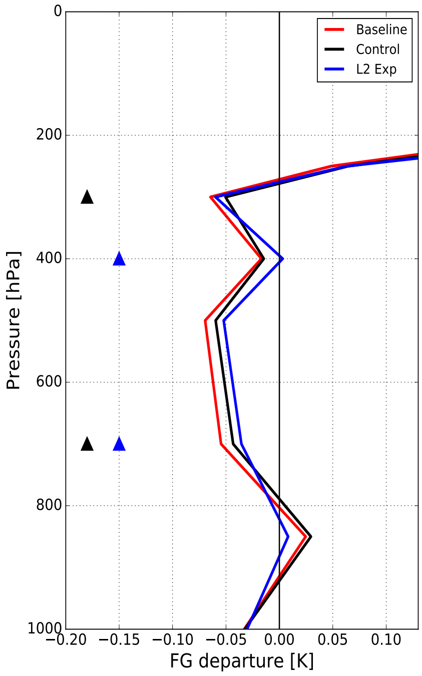
The standard deviation of OMF from aircraft (figures 5.2a, b and c) and radiosoundings (figures 5.2d, e and e) were also evaluated. The standard deviation of OMF of temperature from aircraft of L2 (blue lines), control (black lines) and baseline (red lines) experiments are very similar. However, the differences between the experiments and baseline are statistically significant and there is a degradation (down triangles) in the standard deviation of OMF, more evident above 600 hPa. The L2 experiment standard deviation of OMF of temperature from radiosoundigs show a degradation in almost all profile, it is more evident above 300 hPa where the differences between L2 experiment and baseline are statistically significant (blue empty down triangles in the figures 5.2d, e and f).



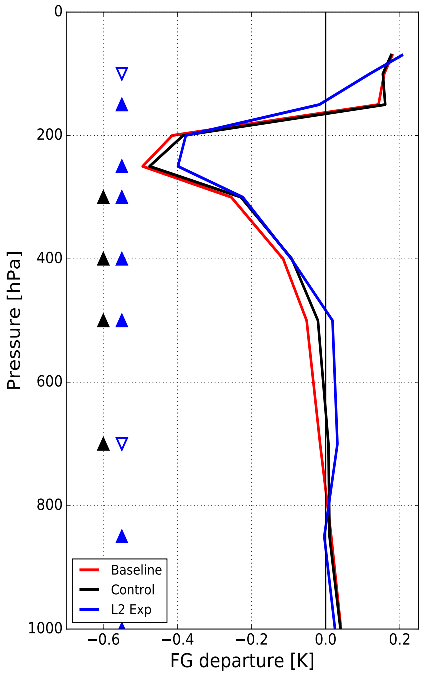
(a) Aircraft - Summer



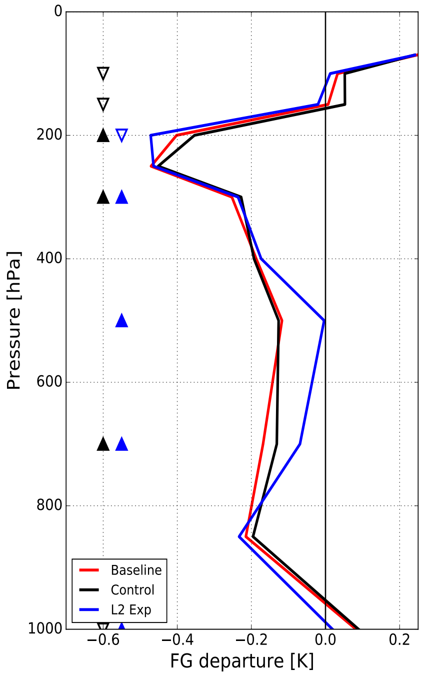
(b) Aircraft - Winter



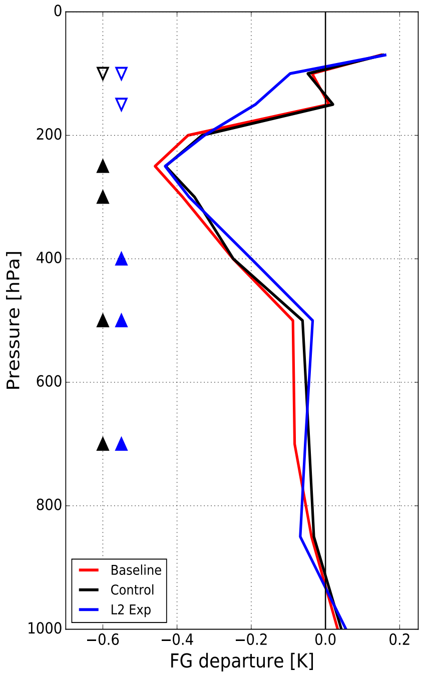
(c) Aircraft - Spring



(d) Radiosounding - Summer

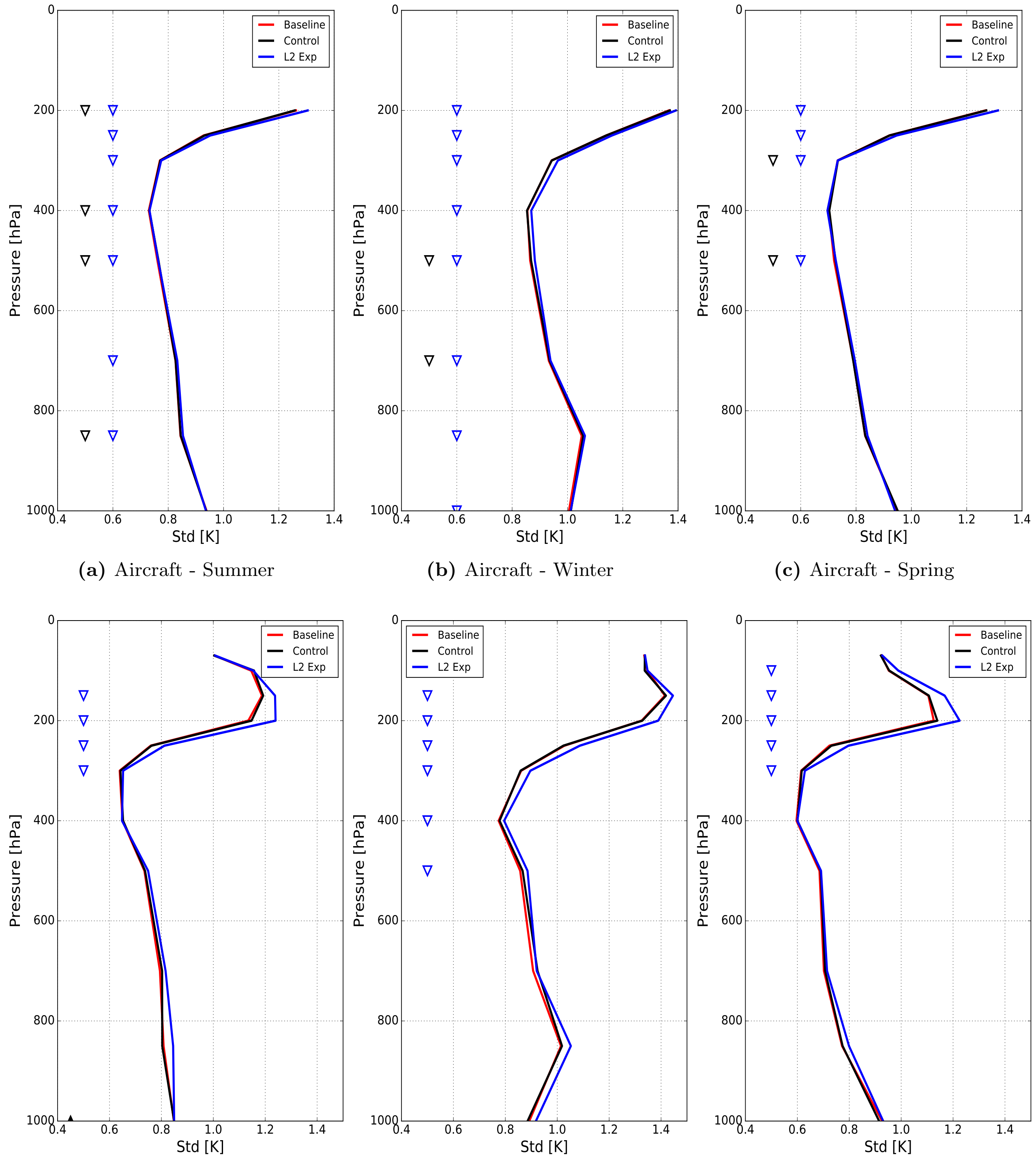


(e) Radiosounding - Winter



(f) Radiosounding - Spring

Figure 5.1: First guess departure (OMF) of temperature from aircraft, top line, and temperature from radiosoundings, bottom line. The blue lines represent L2 experiment, red lines the baseline and the black ones the control experiment. The black (blue) filled up triangle means that in the level the control experiment (L2 experiment) is better than the baseline with 95% of confidence (t-student test). The black (blue) empty down triangle means that in the level the baseline is better than the control experiment (L2 experiment) with 95% of confidence (t-student test). Figure a) aircraft over summer period, b) aircraft over winter, c) aircraft over spring, d)radiosounding over summer period, e) radiosounding over winter and f) radiosounding over spring.



(a) Aircraft - Summer **(b)** Aircraft - Winter **(c)** Aircraft - Spring

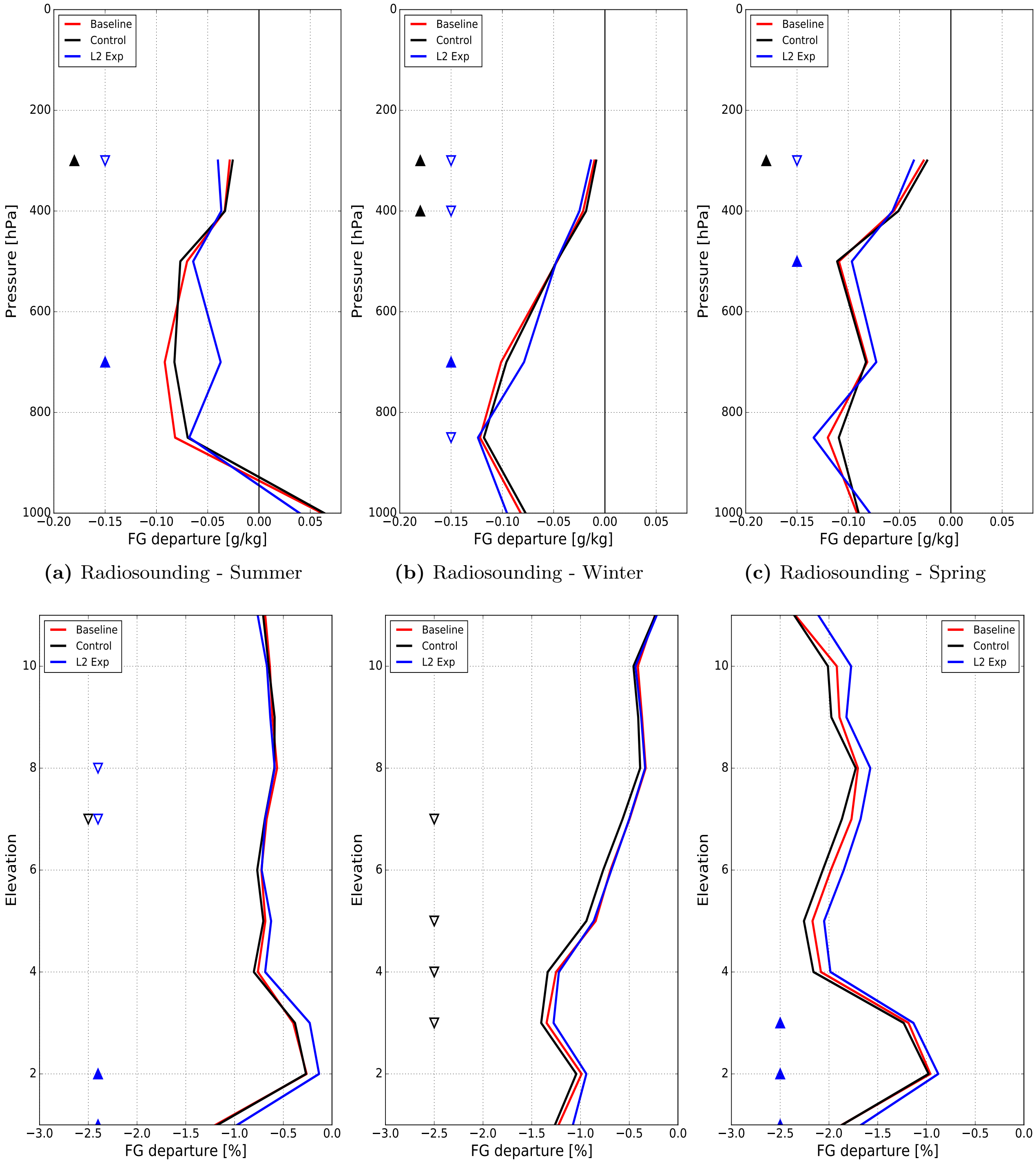
(d) Radiosounding - Summer **(e)** Radiosounding - Winter **(f)** Radiosounding - Spring

Figure 5.2: Same as figure 5.1 for standard deviation of OMF.

The mean OMF of specific humidity (radiosounding observations) and relative humidity (radar observations) are presented in the Figure 5.3. Generally, in the L2 experiment (blue lines) the OMF is reduced in the lower troposphere for both observation types. The OMF of specific humidity from radiosoundings (figure 5.3a, b and c) present a degradation (improvement) in the L2 (control) experiments above 400 hPa (this observation type is assimilated up to 300 hPa) and the difference between the experiments and the baseline are statistically significant for levels with blue empty down triangle (it represents a degradation in L2 experiment) and black filled up triangle (it represents an improvement in the control experiment). The L2 experiment improves the first guess departure between 800 and 600 hPa during the summer and winter periods (figures 5.3a and b), this improvement is statistically significant. However, over the spring period (5.3c) the improvement occurs between 700 and 500 hPa, which is also statistically significant. Close to surface the L2 has an improvement. The control experiment shows an improvement between 850 and 500 hPa during the summer and winter, nevertheless the differences between the control and baseline experiment are not statistically significant.

Figures 5.3d, e and f show the first guess departure of relative humidity observation retrieved from radar data (reflectivity). The vertical information of these observation is in terms of elevation. L2 experiment (blue lines) presents an improvement up to elevation 6 during summer (figure 5.3d), up to elevation 4 during the winter (figure 5.3e) and over all elevations during spring (figure 5.3f). The control experiment (black lines) presents a degradation in the OMF of relative humidity up to elevation 6 during summer and over all profile during the other two periods. There are levels with the differences are statistically significant during the three periods, but they are not similar in terms of elevation where it occurs.

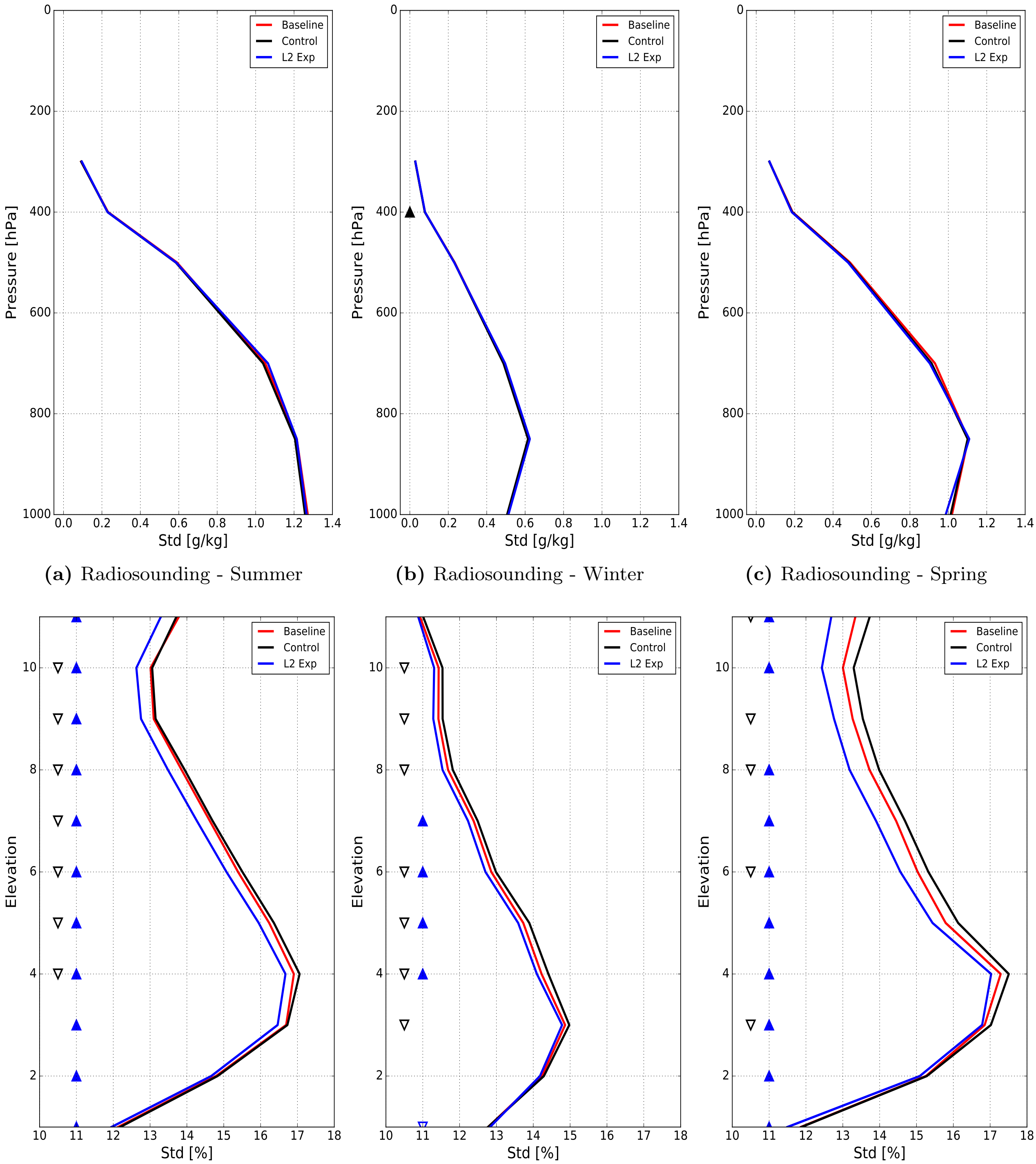
The standard deviation of OMF of specific humidity from radiosoundings were also analysed. Figures 5.4a, b and c show the standard deviation profiles. The experiment presents very similar curves, there is a small improvement in the standard deviation of L2 experiment close to surface during spring period (figure 5.4c). The difference between the control (black lines) and baseline (red lines) is statistically significant at level 400 hPa over winter period (figure 5.4b), however is not possible to see any difference between the two lines. The L2 experiment (blue line) show an improvement in terms of standard deviation of OMF of relative humidity from radar (figures 5.4d, e and f) during the three periods analysed, and these improvement is statistically significant in almost all elevations (blue filled up triangles). An opposite behaviour is observed in the control experiment (black lines) which has a degradation in all periods. This experiment present some levels with a degradation which is statistically significant (black empty down triangles).



(a) Radiosounding - Summer **(b)** Radiosounding - Winter **(c)** Radiosounding - Spring

(d) Radar - Summer **(e)** Radar - Winter **(f)** Radar - Spring

Figure 5.3: Same as figure 5.1 for specific humidity from radiosoundings, first line, and for relative humidity retrieved from radar data.



(a) Radiosounding - Summer **(b)** Radiosounding - Winter **(c)** Radiosounding - Spring

(d) Radar - Summer **(e)** Radar - Winter **(f)** Radar - Spring

Figure 5.4: Same as figure 5.3 for the standard deviation of OMF.

Two additional L2 experiments were performed. In the first one the σo of temperature and humidity of L2 experiment were modified, they were increased close to surface, and the σo of temperature were decreased between 700 hPa and 200 hPa, this experiment was performed from January, 1st to 31st 2018. Another additional L2 experiment used the original σo and relaxed the vertical filter applied to the L2 profiles, which means that the observation close to surface (below 1000 hPa over sea and below

900 hPa over land orography bellow 1 km) were assimilated. This second experiment was carried out from January, 1st to 5th 2018. The experiments did not present any improvement in the quality of OMF and OMA close to surface. Therefore, it was decided to maintain the configurations presented in the previous chapter.

Two specific cases were also studied. The first one was an hail storm (on May, 26th 2018), it occurred in Bordeaux and Cognac, France. The experiments were run during 5 days, from May, 23rd to May, 28th 2018. The second one was a storm in Dordogne (on July, 4th 2018), France, the experiments were also performed during 5 days, from July, 1st to July, 6th 2018. Both events represented an extreme weather and they represent some damage in the cities affected. It was chosen not to present the OMF figures of the study cases because in general they have the same behaviour presented in the others three periods analysed.

In summary, L2 experiment helped to decrease the OMF of temperature observations in the middle atmosphere. These behaviour are not observed in the standard deviation profiles where there is a degradation (top of profiles). Control experiment also helped to decrease the first guess departure of temperature observations in the middle atmosphere and up to 300 hPa (almost all periods). As in L2 experiment this characteristic are not observed in the standard deviation profiles. In terms of humidity observations, the L2 Experiment helped to decrease the first guess departure (OMF) between 800 and 700 hPa (radiosounding), however at 300 hPa there is a degradation. In the relative humidity, retrieved from radar observations, the OMF present an improvement up to elevation 6. The standard deviation of OMF of radar observations shows an improvement in almost all elevations. Control Experiment helped to decrease the OMF of radiosoundings between 800 and 600 hPa (except spring period) and up to 300 hPa. However, this experiment contributed to increase the OMF of radar (RH), more evident during the winter and spring period. The standard deviation of radar observations is increased in this experiment and it is more evident during winter and spring periods.

5.2 Task 2.ii - Forecast verification

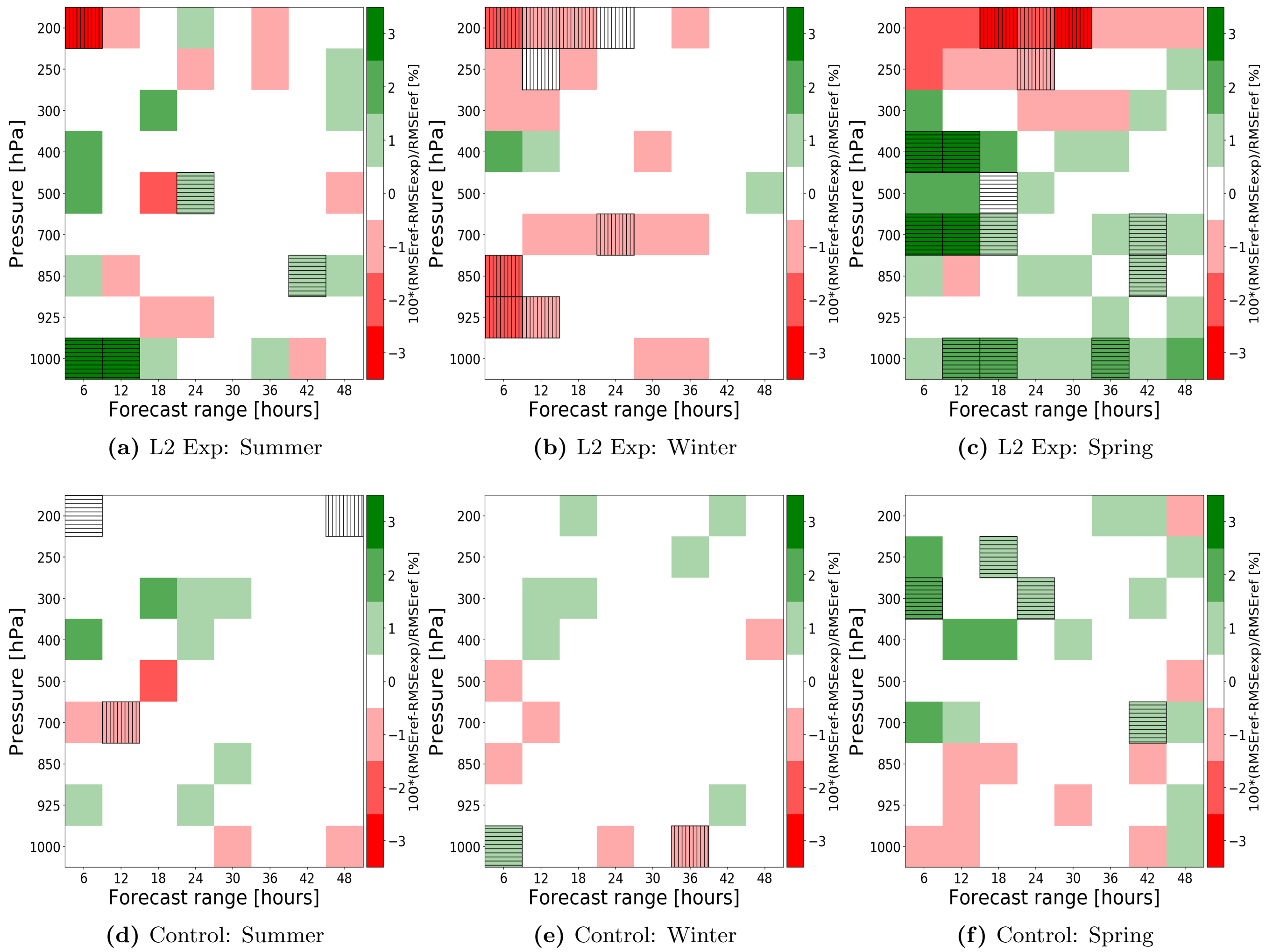
5.2.1 Upper-air verification and surface verification

Upper-air verification

The forecast skill verification of upper-air forecast of control, L2 and baseline experiments were performed against ECMWF analyses as independent analyses. Figure 5.5 presents the relative differences of root mean square error (RMSE) in % of temperature, the figures 5.5a, b (top figures 5.5) represent the differences between L2 and baseline experiment during summer, winter and spring, respectively. The bottom figures (5.5d, e and f) represent the differences between control and baseline experiment (summer, winter and spring, respectively). The positive values (green color) in the figures represent an improvement in the experiment when compared against baseline, the negative values (red colors) represent a degradation in the experiments. The levels and forecast hours where there is vertical (degradation) or horizontal (improvement) lines represent that the differences between the RMSE of experiments and baseline are statistical significant with 95% of significance (t-student test).

L2 experiment present a degradation during winter and spring above 250 hPa, however the differences of RMSE are statistical significant in the winter at 200 hPa from 6 to 24 hours of forecast, in spring period the differences are significant between 18 and 30 hours. In spring period, 5.5c there is some improvement between 700 and 400 hPa and at 1000 hPa, it is significant in the first 12 hours of forecast.

In the summer period, 5.5a, at 1000 hPa present improvement at 6 and 12 hours forecast, but it seems to be neutral in almost all levels and forecast ranges. The control experiment has neutral impact during the winter and summer periods, figures 5.5d and e, respectively. In the spring period, it is possible to notice some improvement between 700 and 250 hPa (figure 5.5f), however, the signal is almost neutral as in the others periods.



(a) L2 Exp: Summer

(b) L2 Exp: Winter

(c) L2 Exp: Spring

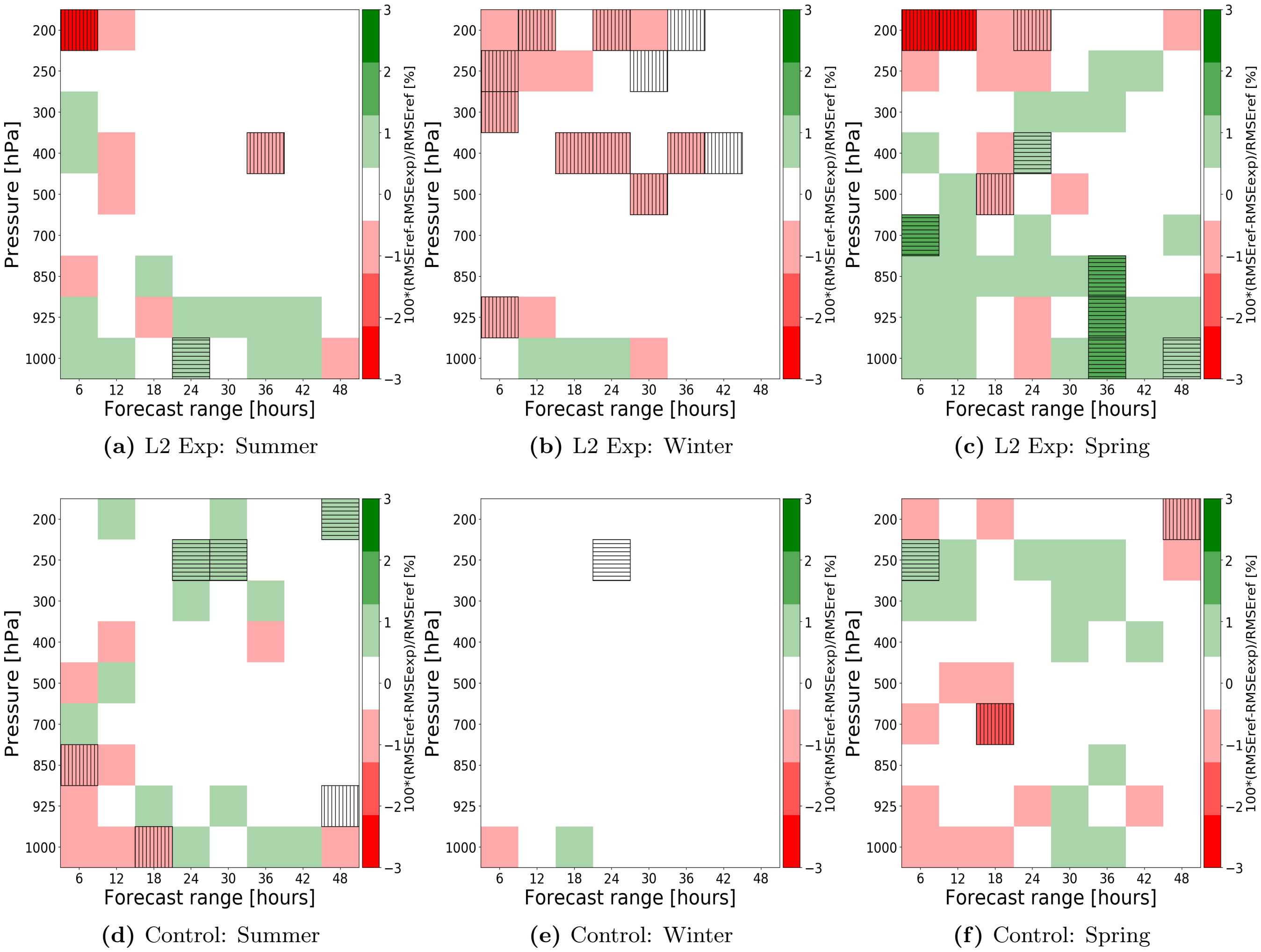
(d) Control: Summer

(e) Control: Winter

(f) Control: Spring

Figure 5.5: Relative difference of root mean square error (RMSE) in % of temperature, 100*((RMSEBaseline-RMSEExp)/RMSEbaseline). The RMSE of baseline and experiments were calculated against ECMWF analyses. Horizontal axis represents the forecast range (from 6 hours to 48 hours) and the vertical axis represents the pressure levels (1000 hPa to 200 hPa). The vertical lines and color red represent that the baseline is better than the experiments with 95% of confidence (t-student test). Horizontal lines and color green represent that the experiment is better than baseline experiment with 95% of confidence. a) L2 Experiment over summer period, b) L2 Experiment over winter, c) L2 Experiment over spring, d) control Experiment over summer period, e) control Experiment over winter and f) L2 Experiment over spring.

The forecast skill of wind intensity was also assessed. The figures 5.6 have the same meaning that figures 5.5 but the relative differences are related with the wind intensity fields. As in the temperature evaluation, on the top of figure (200 hPa during summer and between 250 and 200 hPa during the others periods) at forecast ranges 6-12 hours it is possible to notice degradation in the skill of L2 experiments (figures 5.6a, b and c). However, as for temperature there are levels and forecast ranges where the impact of assimilating L2 product is neutral. The control experiment did not show a clear signal in terms of improvement/degradation (figures 5.6d, e and f) when compared to the baseline experiment. Over winter period the neutrality is very well defined.

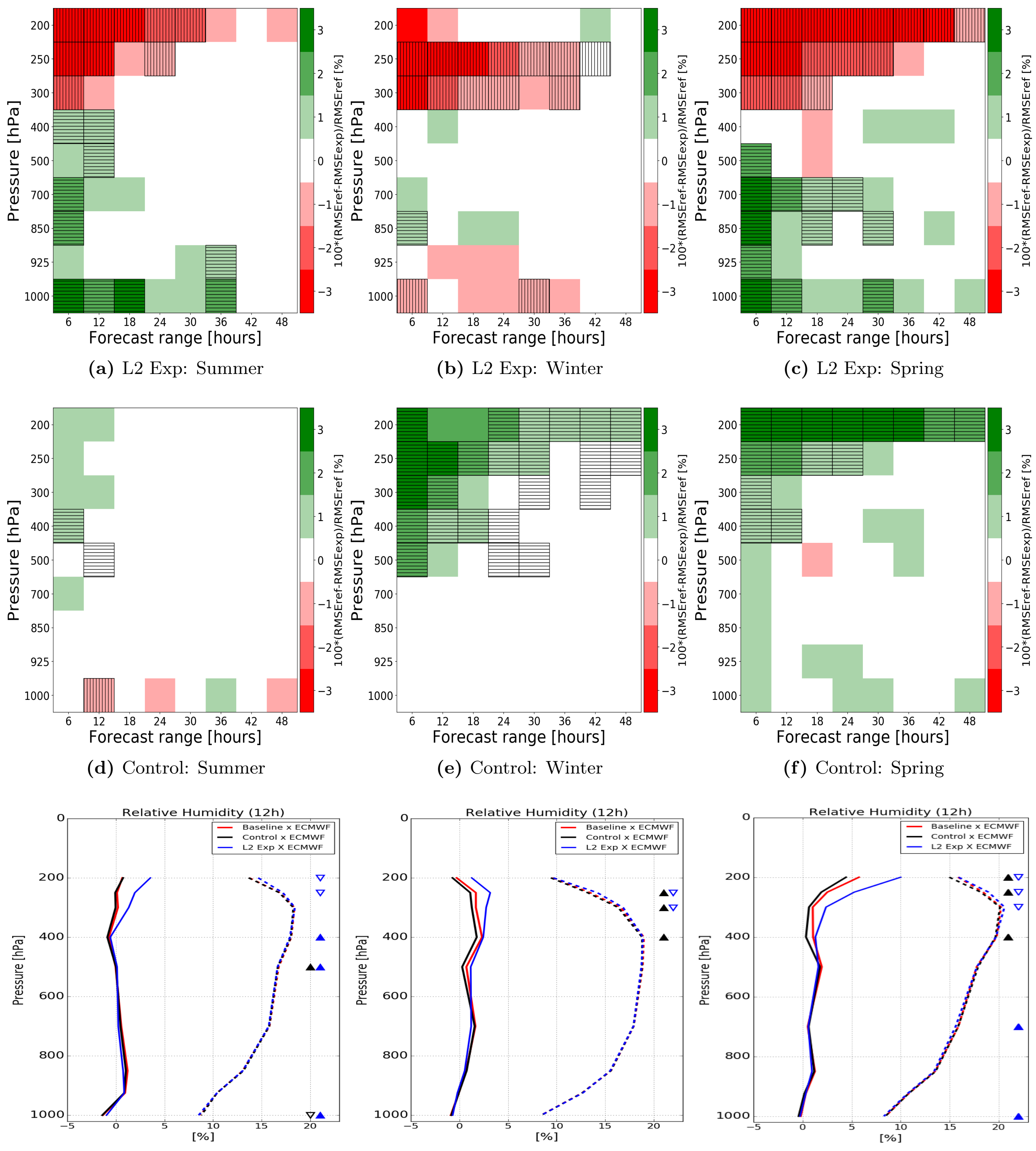


(a) L2 Exp: Summer **(b)** L2 Exp: Winter **(c)** L2 Exp: Spring

(d) Control: Summer **(e)** Control: Winter **(f)** Control: Spring

Figure 5.6: Same as Figure 5.5 for wind intensity.

The third evaluation in terms of upper-air fields was the relative humidity. The figures 5.7a-f represent the same as figures 5.5. The evaluation of L2 experiment shows a degradation (statistic significant) between 300 and 200 hPa present in the 3 periods evaluated which is present in almost all forecast ranges (Figures 5.7a, b and c). Nevertheless, there are levels and forecast ranges (more evident in the first 18 hours of forecast) where some improvement is observed in the L2 experiment, mainly during summer and spring periods. Control experiment is represented in the figures 5.7d, e and f. This experiment present an improvement in the humidity fields well characterized during winter and spring, above 500 hPa. During the summer, the contribution of this data was neutral in almost all levels and forecast ranges. In all variables evaluated here, there is a small (close to zero %) relative differences of RMSE in areas where the differences between the baseline and the experiment present significance statistical. Another way to see the differences between the experiments is plot of the bias and standard deviation of bias. Figures 5.7g-i present the profiles of bias and standard deviation of bias from relative humidity forecast at 12 hours. The big differences are present in the bias and standard deviation profiles in the top of atmosphere (same signal observed in the others figures), where the L2 experiment (blue line in the figures) contributed to degrade the forecast of relative humidity in these regions. And it is also possible to observe that the differences in the RMSE are more related with differences in the bias than in the standard deviation. And during the three periods the control experiment (black lines on the figures 5.7g-i) presents an improvement in the bias above 400 hPa (during summer) and above 500 hPa over the other periods.



(a) L2 Exp: Summer

(b) L2 Exp: Winter

(c) L2 Exp: Spring

(d) Control: Summer

(e) Control: Winter

(f) Control: Spring

(g) 12h Forecast - Summer

(h) 12h Forecast - Winter

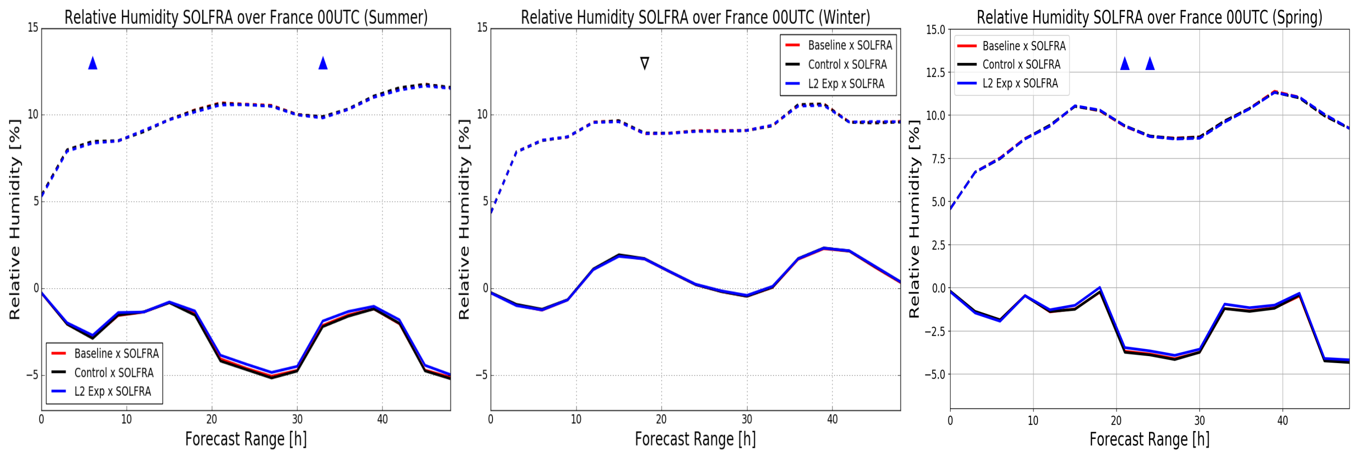
(i) 12h Forecast - Spring

Figure 5.7: Same as figure 5.5 for relative humidity. The figures g, h and i represent the bias (continuous) and standard deviation of bias (dashed lines) of AROME forecast against ECMWF analyses at 12 h over summer, winter and spring, respectively. Red lines are the baseline experiment, black lines represent the control experiment and the blue lines the L2 Experiment. The triangles have the same meaning that in figure 5.1 but the statistical test is related with the differences of RMSE between the baseline and experiments.

Other variables were also evaluated, however the variables presented before are the ones where the impact in the control and L2 experiments are more evident. The same evaluation showed before was also made against radiosoundings as independent information. The results were not shown because the L2 and control experiments presented differences which were not statistically significant, contrary to the ones showed for the evaluation using the ECMWF's analyses as independent analyses.

Surface verification

An evaluation were performed against surface stations which are situated over France domain, representing more than 600 stations. The variables evaluated were surface pressure, temperature, relative humidity, wind, rain and nebulosity. But here, only the results for the relative humidity are presented, because it is the only variable for which it is possible to see some differences between the 3 experiments. As in the evaluation against the ECMWF analyses the differences between the experiments are more evident in the bias than in the standard deviation. Figures 5.8a-c present the bias (continuous lines) and standard deviation of bias (dashed lines) of relative humidity, the test of significance statistical between the experiment and baseline was performed in terms of RMSE. In these figures the triangle meaning is the same as for figures 5.1. During the summer and spring the bias from L2 experiment (blue lines) is smaller (closer to zero) when compared to other 2 experiments. It is more evident between 20 and 40 hours of forecast range. It is necessary to emphasize that the differences between the experiments are small. The surface verification were also made using the synop data as independent information. The results were not shown because the differences between the experiments (same variables used in the evaluation against surface stations) were very small, which did not bring new informations for the assessment.

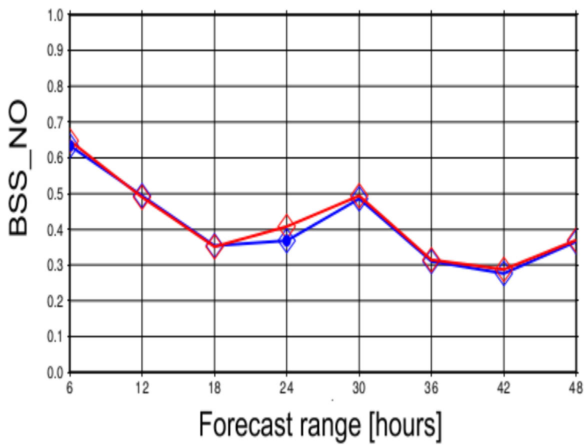


(a) Summer (b) Winter (c) Spring

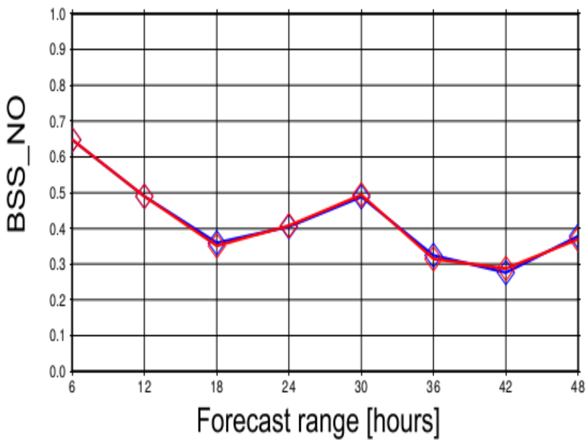
Figure 5.8: Bias (continuous lines) and standard deviation of bias (dashed lines) for relative humidity form surface stations. Figure a) summer period, b) winter period and c) spring period. The lines and triangles have the same mean as in figure 5.7g. The vertical axis is the relative humidity (%) and the horizontal axis is the forecast range.

5.2.2 Precipitation Verification

The precipitation verification were performed for all periods, skill scores, neighbourhood and precipitation thresholds proposed in the chapter 4.2. The results did not show big differences between the experiments and baseline, and if there were differences they were not present in all forecast ranges and precipitation threshold evaluated. For this reason, two skill scores from summer period were selected to represent the evaluation. These skills give the quality of precipitation forecast. The first one, present in the figure 5.9a and b, is the Brier Skill Score with neighbourhood observation (BSS_NO), of L2 and control experiments, respectively. The closer to 1 the score is, better is the forecast. The precipitation threshold chosen was 5 mm in 6 hours with neighbourhood of 50.28km. In the figures the forecast range with a circle inside the square means that the differences between the experiment and baseline are statistically significant with 95%. The BSS_NO of control (blue line) and baseline (red line) experiments, figure 5.9b, did not show differences that are statistically significant. Figure 5.9a presents the BSS_NO of L2 (blue lines) and baseline (red lines), the L2 experiment presents a degradation, when compared with the baseline experiment, at 24 hours forecast, and the differences between the two experiments are statistically significant (circle blue inside square).



(a) L2 Experiment



(b) Control Experiment

Figure 5.9: Brier Skill score (BSS NO) for 5mm of precipitation accumulated in 6 hours, with the neighbourhood of 50.28.km. a) Red line is the baseline and blue ones is the control experiment and b) red line is the baseline and blue ones is the L2 experiment. The squares with a circle inside represent that the differences in terms of BSS NO between the experiment and baseline are statistically significant with 95% of confidence. These figures represent the summer period.

Despite there are some differences between the experiments in the figures 5.9, it is not easy to identify these differences in the quality of precipitation forecast when a time series of other skill score is analysed. Figure 5.10 shows the time series of Brier Score (BS) for 5 mm of precipitation with neighbourhood of 50.28km at 24 hours forecast (the closer to 0 the score is, better is the forecast). The value of BS of the three experiments varies a lot during the period analysed (summer). For example, on August, 7th the L2 (blue line) and control (black line) experiments has a value smaller than the baseline, which means a better forecast, characteristic not observed in figure 5.9b. However, on August, 15th it is possible to see a relationship between the BSS_NO and BS for the L2 experiment, both scores show a degradation of precipitation forecast.

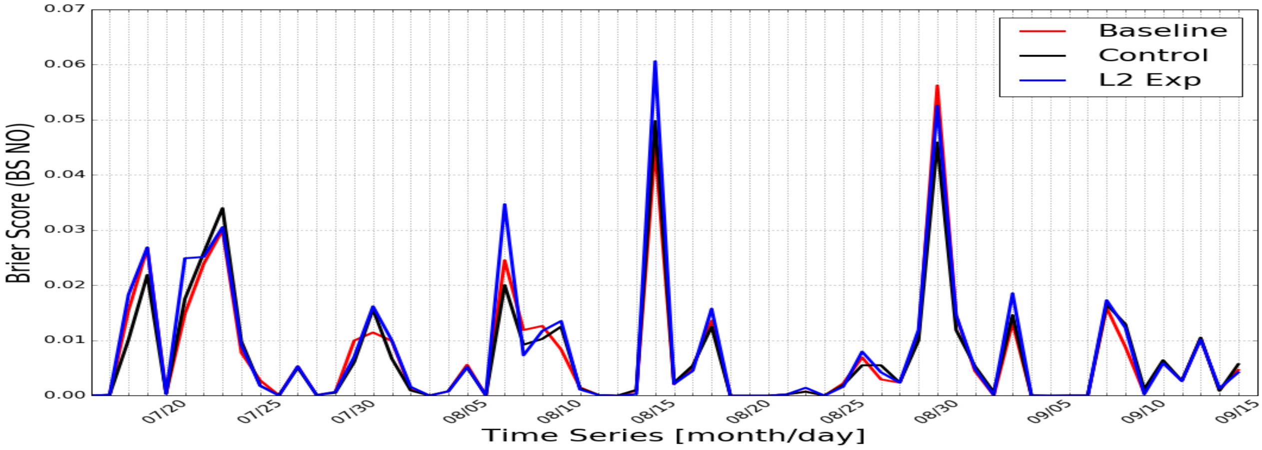


Figure 5.10: Temporal variation of Brier Score (BS) for 5 mm of precipitation accumulated in 6 hours at 24 hours forecast range. Red line represent baseline experiment, black line is the control experiment and blue line is the L2 experiment. The period analysed represent the summer experiment (July, 15th to September 15th 2018).

The study cases experiments were also evaluated in terms of forecast skill, but no conclusion can be drawn from these results. These kind of analyses were not made for the two additional L2 experiment performed (the main objective of these experiment were to verified the impacts in the OMF profiles).

In summary, the upper-air forecast skill evaluation showed that for temperature and wind intensity the scores are almost neutral in the control experiment in the three periods and there is degradation in the top of atmosphere in the L2 experiment. The relative humidity forecast present an improvement in the control experiment and improvement/degradation in the L2 experiment. However, when a specific profile (12 hours forecast) is analysed the standard deviation lines are similar and the differences in the RMSE come from the bias values. The surface skill score verification did not show big differences and when there is some differences they were only evident in the bias.

The precipitation verification scores vary during the periods analysed, as shown in the time series of the summer period (figure 5.10). The control and L2 experiments showed a similar behaviour to baseline experiment, it means that the differences between the experiment and baseline were almost neutral. But, neutral does not mean same precipitation forecast (as presented in the figure 5.10).

Chapter 6

Task 3 - Feed-back on the EARS-IASI L2 products

6.1 Conclusion

This study aimed to assess the potential benefit and also the practicalities of assimilating L2 temperature and humidity profiles in a regional model (AROME-France) in replacement of IASI and AMSU radiances.

The statistics of L2 data showed a good quality when compared with other observations first guess departures (radiosoundings and AIREP). The thinning applied to the L2 data, in order to have a spatial distribution comparable to L1 data, kept the quality of the L2 product profiles, demonstrating that the profile with best QCI is selected.

The radiosounding σo for temperature and humidity were considered to estimate the L2 product σo, 15 % of relative humidity for the specific humidity profile and 1.2 times the radiosounding error profile for temperature.

After the statistics assessment, three experiments were built, the baseline, control (with L1 observation from IASI, AMSU-A and MHS) and L2 experiments to assess the impact of L2 products assimilation in the AROME model. Some configurations of the control experiment were based in the L2 products to have a fair evaluation between the two experiments.

The assimilation results showed that the L2 product is suitable to be assimilated in the Numerical Weather Prediction (NWP) models. The assimilation of L2 temperature and specific humidity profiles helped to decrease the first guess and the analysis departure of the other observations (during the assimilation cycle).

The impact of the two experiments, control and L2, against baseline in the forecast skills were also evaluated. The Control experiment improves the upper-troposphere forecast skill for humidity. The L2 experiment helped to improve the humidity skill scores in the middle-lower troposphere where no information from L1 radiance is assimilated and it also degraded the scores above 300 hPa for the temperature and wind intensity, not all periods, and also humidity. The surface verification showed that if there was some differences between the experiments it was more evident in the bias. The forecast skill scores for precipitation were almost neutral over all periods analysed.

6.2 Way Forward

The L2 profiles are fully suitable for assimilation in a regional NWP model. Nevertheless, several points have to be taken into account and tuned in each NWP system to optimise the exploitation of the information in the L2 products. In particular, both horizontal and vertical data selection have to be carefully studied.

The assimilation experiments in the present study were carried out with a simplified approach, taking advantage of the existing assimilation framework for radiosoundings. In this approach, the IASI L2 profiles were assumed as pseudo-sonde, i.e. without taking vertical error correlation and vertical sensitivity into account. It is expected that a more sophisticated assimilation scheme with an observation operator, e.g. applying averaging kernels from IASI L2 if they can be supplied by EUMETSAT, could be beneficial both in the product evaluation and assessment processes.

In addition, among possible directions to further study:

- evaluation the horizontal error correlation to define the optimal horizontal thinning;
- evaluation the vertical error correlation;
- use of the vertical error correlation in the assimilation process, investigating also cross-variable ;
- adaptive vertical selection, e.g. to have only one L2 level per model level.

On the latter point, as the quality within a profile varies depending on the presence of cloud within the pixel and its elevation, some information e.g. by provision of a cloud top or of uncertainty profiles could be useful to determine the parts of the profiles which are most valuable for assimilation (for instance where in the vertical the hyperspectral IR information can be fully exploited). In the current settings, the quality control is pixel-based and a whole profile can be rejected because the quality indicator is a single uncertainty estimate relating to the quality in the lower troposphere.

In preparation for IRS and in mitigation of the loss of MW sensors, L2 profiles obtained from IR sensor only should be assessed for assimilation. In that case, a detailed cloud characterisation or uncertainty estimates in the vertical would be even more important information.

Bibliography

EUMETSAT. IASI level 2: Product guide, 2017.

A. D. Collard. Selection of iasi channels for use in numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 133(629):1977–1991, 2007. doi: 10.1002/qj.178. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.178.

P. Brousseau, Y. Seity, D. Ricard, and J. Léger. Improvement of the forecast of convective activity from the AROME-France system. *Quarterly Journal of the Royal Meteorological Society*, 142(699): 2231–2243, 2016. doi: 10.1002/qj.2822. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2822.

Y. Seity, P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson. The arome-france convective-scale operational model. *Monthly Weather Review*, 139(3):976–991, 2011. doi: 10.1175/2010MWR3425.1. URL https://doi.org/10.1175/2010MWR3425.1.

J. Andrey-Andrés. Impact of principal components compression on the assimilation of hyperspectral infrared data. eumetsat fellowship report, 2017.

EUMETSAT. IASI L2 PPF v6.3 validation report., 2017. [Available online at http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_DMT_920559&RevisionSelectionMethod=LatestReleased&Rendition=Web].

EUMETSAT. IASI L2 PPF v6.3 validation report, 2017.

EUMETSAT. IASI L2 PPF v6.4 validation report, 2018.

M. Amodei and J. Stein. Deterministic and fuzzy verification methods for a hierarchy of numerical models. *Meteorological Applications*, 16(2):191–203, 2008. doi: 10.1002/met.101. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.101.

Appendices

Appendix A

Statement of work

4 TASKS BREAKDOWN

Task 1 Definition of the experiments setup

This first task consists in the evaluation of the IASI L2 products in the regional model framework and in the preparation and definition of the assimilation experiments. The evaluation shall be performed on a statistically significant and representative period, in view of the subsequent assimilation experiments.

Task 1.i. Evaluation of IASI L2 products, uncertainties and acceptance criteria

The framework of the regional model shall be adapted to the ingestion of IASI L2 temperature and humidity products as pseudo-radiosoundings, using the near-real time EARS-IASI L2 regional service.

The profiles shall be monitored with the routine facilities coupled to the numerical model, including intercomparisons with the model itself and synoptic sondes where possible. The departures shall be analysed as a function of the quality indicators as well as of the cloud signal, which are provided together with the atmospheric parameters. This analysis is intended to determine the selection criteria and thresholds to be applied to the IASI L2 data in the respective assimilation experiments.

In addition to defining acceptance criteria and thresholds, the observation error estimates to be used for the assimilation of individual soundings shall be derived from this analysis. This may result in assigning different uncertainty characteristics to the IASI L2 profiles depending on their quality/cloudiness level, on the surface type (e.g. distinguishing maritime/continental pixels) or any relevant stratification.

Task 1.ii. Definition of the assimilation experiments

These activities are to define the baseline and the various assimilation experiments.

In particular, as an objective is to assess the impact of assimilating L2 profiles relatively to the current operational assimilation of L1 measurements from EPS satellites, an experiment with L2 products shall be designed with similar data selection/sampling. This particular configuration may be restricted to pixels classified as cloud-free in the assimilation system or to pixels where clear channels are assimilated, consistently with the operational assimilation settings in the regional model.

In addition, in order to possibly take advantage of the large yield allowed by the combined microwave + infrared sounding, an assimilation experiment shall be designed including sounding in clear and cloudy pixels. Hence, this may attempt to assimilate soundings in more pixels than what is done in the current operational setup where Level 1 measurements are ingested. The L2 data selection shall take into account the quality control criteria established in Task 1.i and the possible need to subsample the information horizontally, depending on the assumptions made for data assimilation, e.g. in terms of horizontal correlations. Depending on how much redundancy or similarities there may be with between Metop/IASI & AMSU L1 with measurements from other hyperspectral missions, the baseline experiment may deny also data from other sounders. This is in order to isolate more the effect of assimilating Metop L1 and L2 products relatively to each other. The following assimilation experiments are typically expected, but not necessarily limited to:

- Exp.1 Baseline: All but IASI and AMSU (possibly excluding other satellite sounders)
- Exp.2 Control: Baseline + Metop/IASI L1C and Metop/AMSU L1B
- Exp.3 Baseline + EARS-IASI L2 products, with similar pixel sampling as in Exp.2
- Exp.4 Baseline + EARS-IASI L2 products on broader useful yield as defined in Task 1.i

The period of the study shall include seasonal variations and the convective season, from 6 to ideally 12 months. The definition of the assimilation experiments shall also explain and discuss the respective configurations: selection of input data, choice of the vertical resolution at which the EARS-IASI L2 products will be assimilated (e.g. on their original grid sampling or on slab layers).

Task 1.iii. Definition of the forecast skills assessment

These activities are to define the reference data, the methodology and the metrics to quantify the relative merits and limitations of the respective experiments proposed in Task 1.ii.

The assessment shall include 3-way intercomparisons of forecasted atmospheric profiles, satellite profiles and suitable independent data (e.g. correlative *in situ* sonde measurements, ground-based GPS total precipitable water-vapour, microwave radiometer…). In addition, the accuracy of other meteorological parameters as predicted under the various experiments defined in Task 1.ii shall be evaluated against ground-based observations and measurements routinely performed in meteorological stations operated by or accessible to Météo-France. This includes for instance, but not limited to, wind fields vs ceilometers data, precipitation forecasts vs rain-gauge. It is expected that these analyses enable multi-day forecast scores for the various assimilation experiments as averaged over the overall study period but also as stratified in time, e.g. on monthly or seasonal basis, and possibly also in space, depending on the variability in the domain of the study. The ability to conduct dedicated forecasts analysis on case studies shall be described, in the eventuality of relevant local weather events during the period of the study.

The reference data, methodology and metrics intended to support the forecasts skills assessment shall form part of the first report.

Input:

- Météo-France Regional model, monitoring and assimilation system
- Routine ground-based and *in situ* data managed by or accessible to Météo-France
- EARS-IASI L2 sounding products (from the operational NRT dissemination stream)
- Guidance from EUMETSAT on using the IASI L2 products, including the temperature and humidity profiles and associated quality indicators, cloud signal.

Output:

- Statistics with the IASI L2 products, as a function of the above departures with the products quality indicators and cloud signal
- Description of the assimilation baseline/experiments, including
  - Configuration of the numerical model and input data
  - Study period
  - Data selection/thinning strategy for IASI L2 (clear/cloud, spatial subsampling, QC…)
  - Observation uncertainty estimates associated to IASI L2 profiles
  - Acceptance criteria & thresholds for IASI L2 profiles
  - Strategy and metrics (reference data, forecast scores…) to evaluate the impact of the different configurations in the various assimilation experiments
- A document "Definition of the assimilation experiments" compiling the above points.

The decision to proceed with the actual experiments will be based on the above results and on acceptance of the assimilation experiments as described in the "Definition" document.

Task 2 Assimilation experiments and evaluation

During this phase, the assimilation experiments as designed and agreed in the Task 1 will be carried out during the planned period. This includes the preparatory activities to instantiate and to adapt the regional model accordingly to the assimilation experiments.

The performances of the models shall be assessed at regular intervals, with the reference data and methodology defined in the definition phase. It is expected that this will rely for a big part on pre-existing tools and data streams routinely exploited at Météo-France to monitor the operational regional model's outputs. Intermediate results will be presented to and discussed with EUMETSAT in progress meetings. In addition to regular monitoring, in the occurrence of relevant local weather events, the possibility to organize some dedicated assessments or case studies with the output of the various experiments shall be established.

Input:

- Météo-France Regional model, monitoring and assimilation system
- Routine ground-based and *in situ* data managed by or accessible to Météo-France
- EARS-IASI L2 sounding products (from the operational NRT dissemination stream)
- Definition of the assimilation experiments, from Task 1

Output:

- Forecasts and forecasts skills assessment from the various experiments
- Reports in progress meeting
- Exhaustive study report compiling the intercomparison results

To be integrated in the Draft Final Study Report

Task 3 Feed-back on the EARS-IASI L2 products

The experience of actually assimilating satellite sounding products in an operational environment like Météo-France regional model is expected to bring new perspectives and new returns from experience with the operational IASI L2 product, in practical and scientific terms.

A way-forward document shall compile the observations made throughout the study, the lessons learnt by actually using the IASI L2 products, including possible recommendations and feed-back in view of using satellite sounding products in an operational model. This may for instance address possibly the accuracy/precision of the IASI L2 products themselves, their relative applicability in certain weather or surface conditions, their vertical resolution vs sampling for assimilation, the products content and format, any needs for additional or auxiliary information, etc.

Input:

- Experience made during Task 1 and 2

Output:

- Way forward document

To be integrated in the Draft Final Study Report

Task 4 Additional or delta experiments (optional)

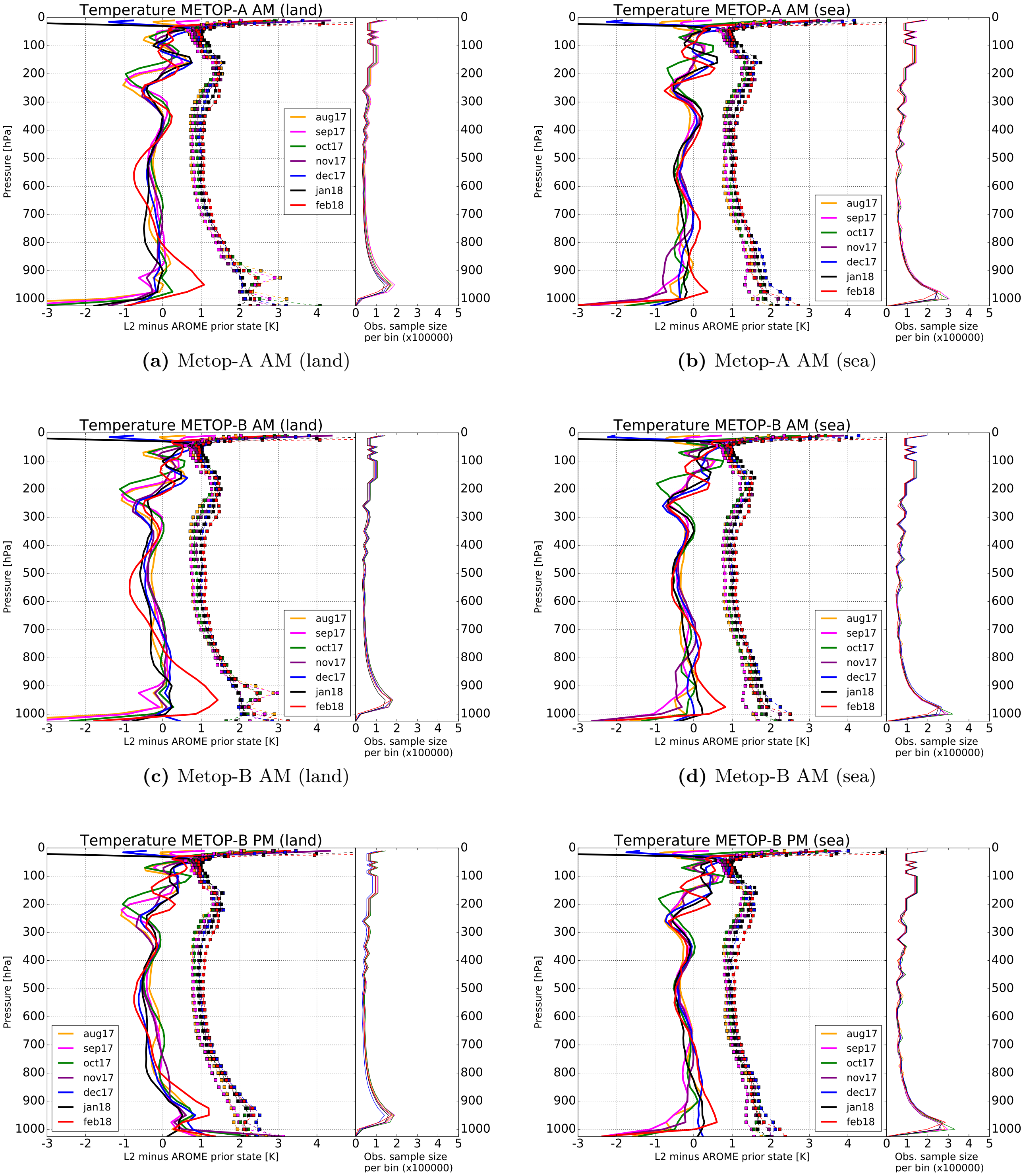
Depending on the results and observations made during the study core period, additional or delta experiments may be desirable to increase or consolidate the understanding of using the IASI L2 products. This may trigger for instance by the need for different assimilation setups or additional auxiliary information, to address specific case studies or to extend the study period. This is the purpose of this optional task, whose exact content will be specified at a later stage, if exercised. The additional work is intended to be performed over a shorter period than the overall core study, of similar complexity as an experiment carried out during the initial period and with a typical duration of 2 months.

Appendix B

Additional L2 product Statistics

B.1 Monthly variation of L2 product

The figures show the monthly variation of L2 data over land and sea. The L2 was divided in three categories based in the data source and time, e.g., Metop-A AM, Metop-B AM and Metop-B PM.



(a) Metop-A AM (land)

(b) Metop-A AM (sea)

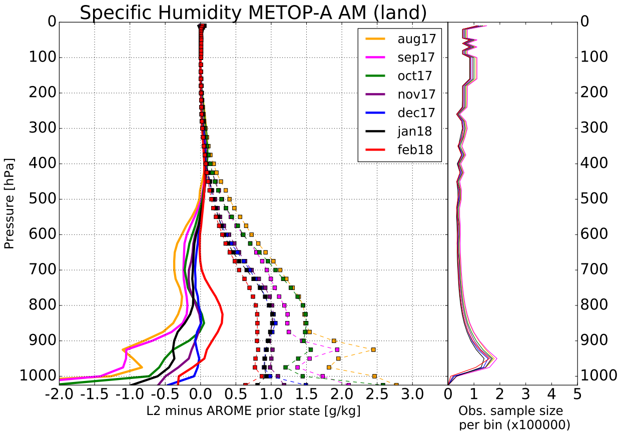
(c) Metop-B AM (land)

(d) Metop-B AM (sea)

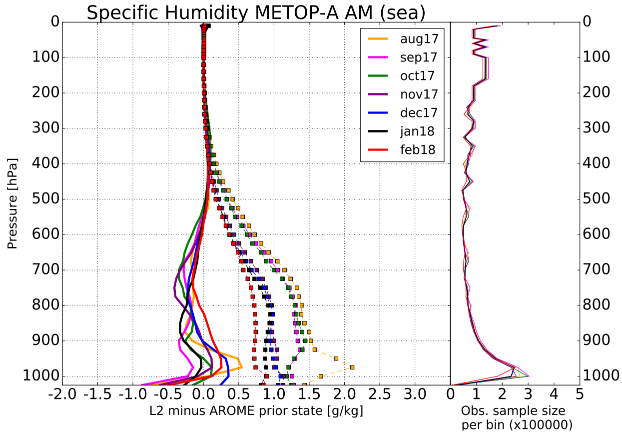
(e) Metop-B PM (land)

(f) Metop-B PM (sea)

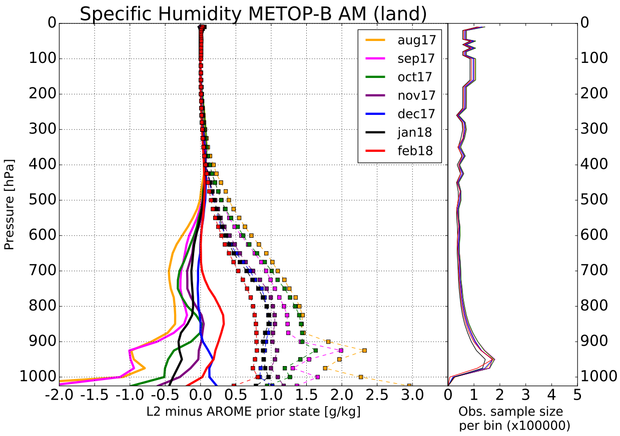
Figure B.1: Temperature profiles statistics over land.



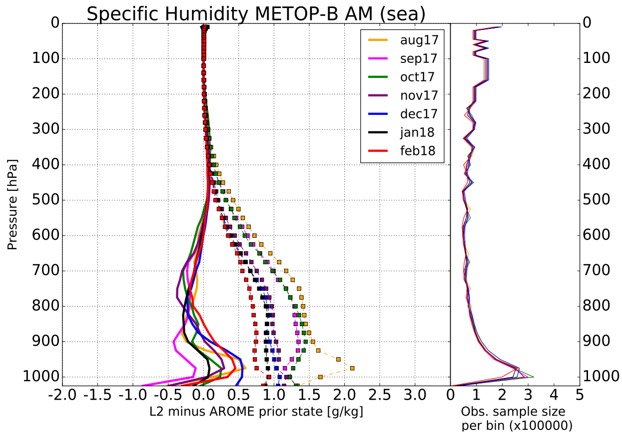
(a) Metop-A AM (land)



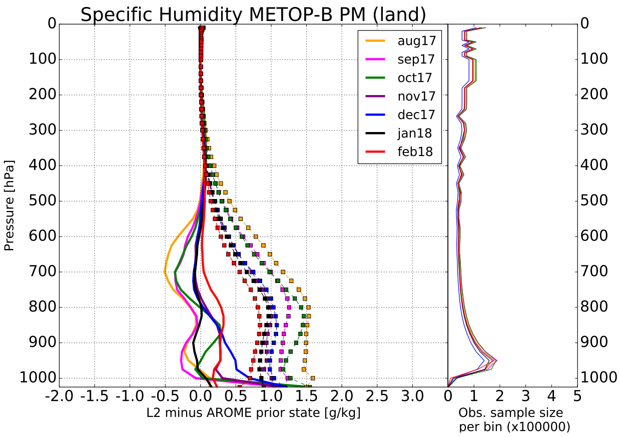
(b) Metop-A AM (sea)



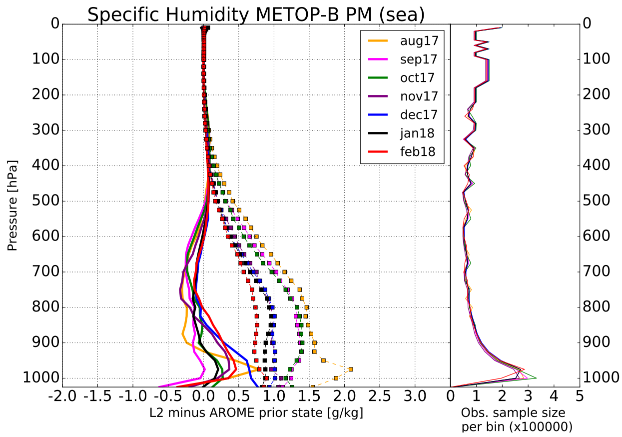
(c) Metop-B AM (land)



(d) Metop-B AM (sea)

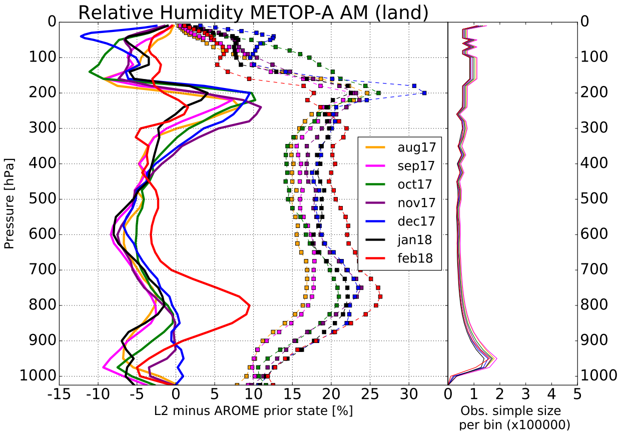


(e) Metop-B PM (land)

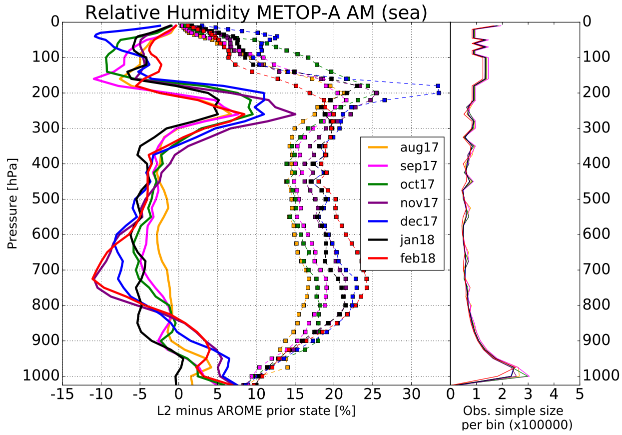


(f) Metop-B PM (sea)

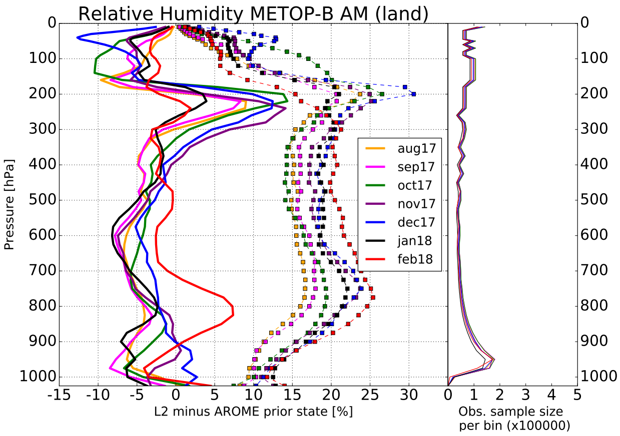
Figure B.2: Temperature profiles statistics over land.



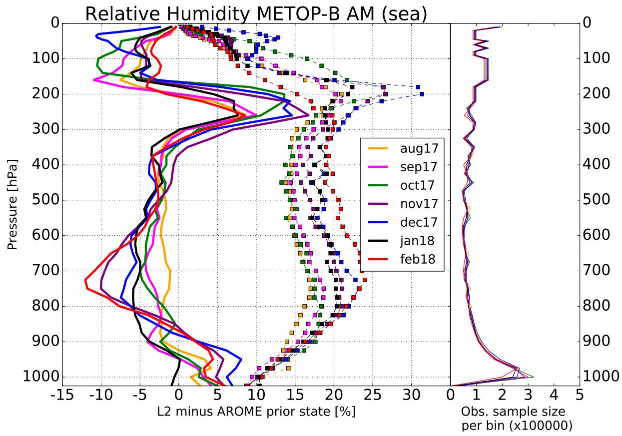
(a) Metop-A AM (land)



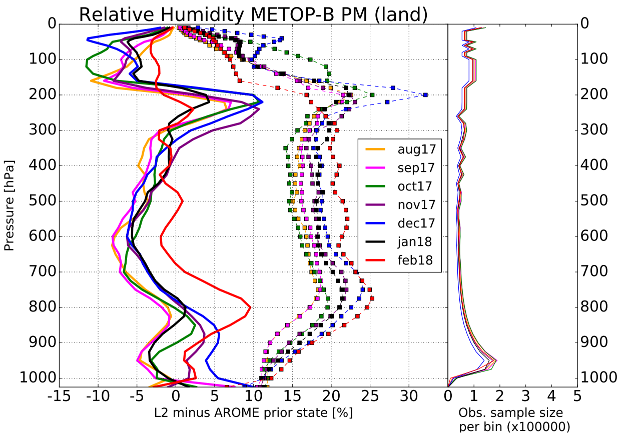
(b) Metop-A AM (sea)



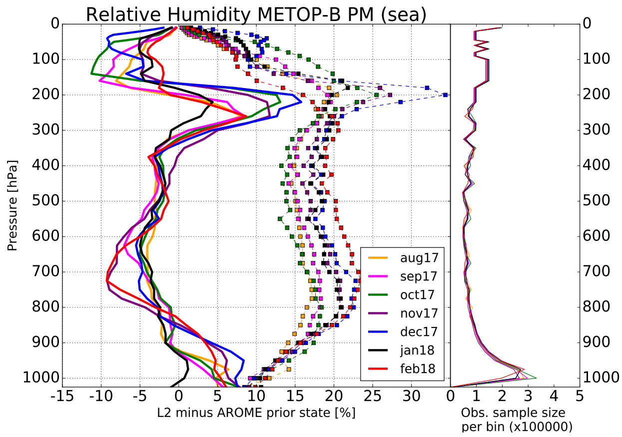
(c) Metop-B AM (land)



(d) Metop-B AM (sea)



(e) Metop-B PM (land)



(f) Metop-B PM (sea)

Figure B.3: Temperature profiles statistics over land.