Variational Bias Correction of Polar-Orbiting Satellite Radiances in Convective-Scale Data Assimilation

Reima Eresmaa

Finnish Meteorological Institute (FMI), Helsinki, Finland

Summary

- Functioning and reliable bias correction is important for the success of satellite data assimilation, but it is non-trivial to apply bias correction algorithms in regional NWP systems.
- A common approach is to cycle the coefficients of the Variational Bias Correction (VarBC) over 24-hour intervals such that consecutive updates deal with similar observing geometries. This implies the need to maintain multiple streams of VarBC coefficients, and each stream is applicable at only one specific daily analysis hour.
- An alternative presented here maintains VarBC coefficients in a single stream that is applied at all analysis hours and is cycled together with background field information. However, updates to the VarBC coefficients are restricted to only one per day per satellite. The timing of the update is specific to each satellite and is chosen to match the most comprehensive data coverage inside the domain.
- The alternative scheme is found to accelerate the convergence of the VarBC coefficients and to improve data fits at those analysis hours when the sampling of the VarBC predictors is compromised.

This study was published in 2024 in *Tellus A*, **76**(1), **https://doi.org/10.16993/tellusa.3259**

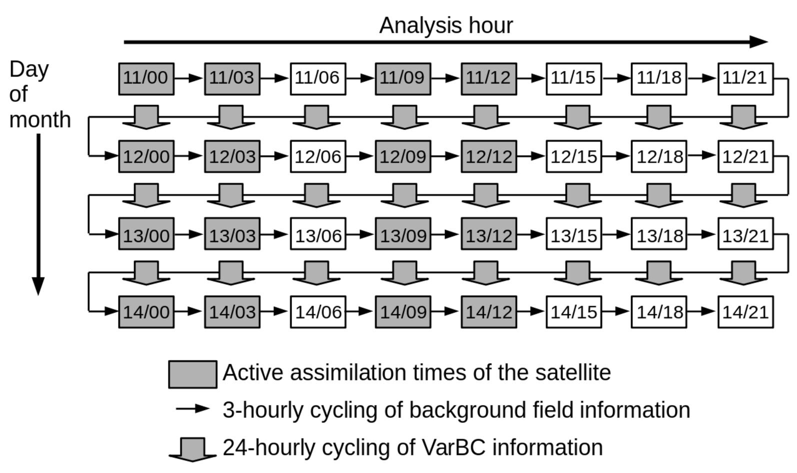
Variants of Variational Bias Correction (VarBC) in regional NWP systems

VarBC represents the observation bias as a linear combination of a set of predictors and their respective bias correction coefficients. The latter are estimated during the variational analysis such that the fit to the observations is optimized. The process requires background values for the VarBC coefficients. By default, these background values are cycled from the previous analysis time.

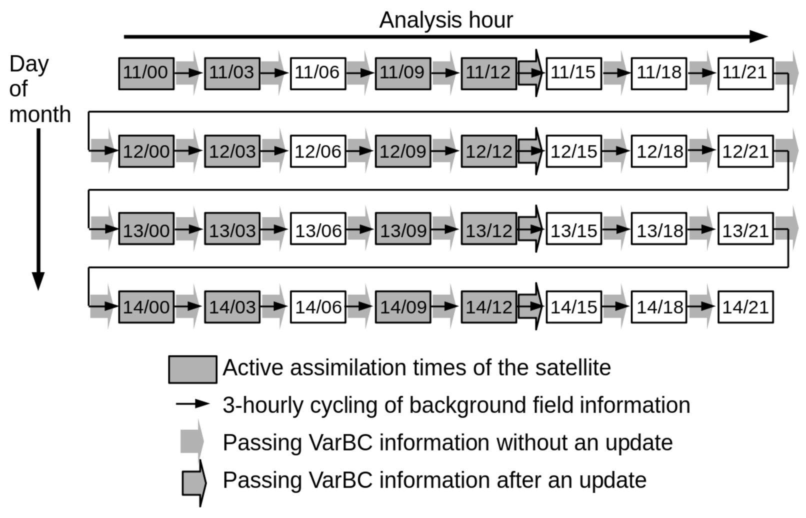
In regional NWP, the variable observing geometry results in irregular sampling of the VarBC predictors, making it problematic to apply the default cycling. Therefore, many regional NWP systems cycle the VarBC coefficients over 24-hour intervals, such that the observing geometry is (relatively) similar at consecutive updates. In this approach, each analysis hour independently maintains its own stream of VarBC coefficients. Here, this convention is referred to as ***"Multiple Stream Daily Update (MSDU)"***.

In this study, we consider an alternative strategy: ***"Single Stream Daily Update (SSDU)"***. Here, the cycling of the VarBC coefficients goes together with the regular cycling of background field information from the previous analysis cycle. However, updating the VarBC coefficients is restricted in time to allow only one update per day for each satellite. The exact time of the satellite-specific updates is chosen to match the hours at which the data coverage inside the NWP system domain is the most comprehensive.

Multiple Stream Daily Update (MSDU)



Single Stream Daily Update (SSDU)



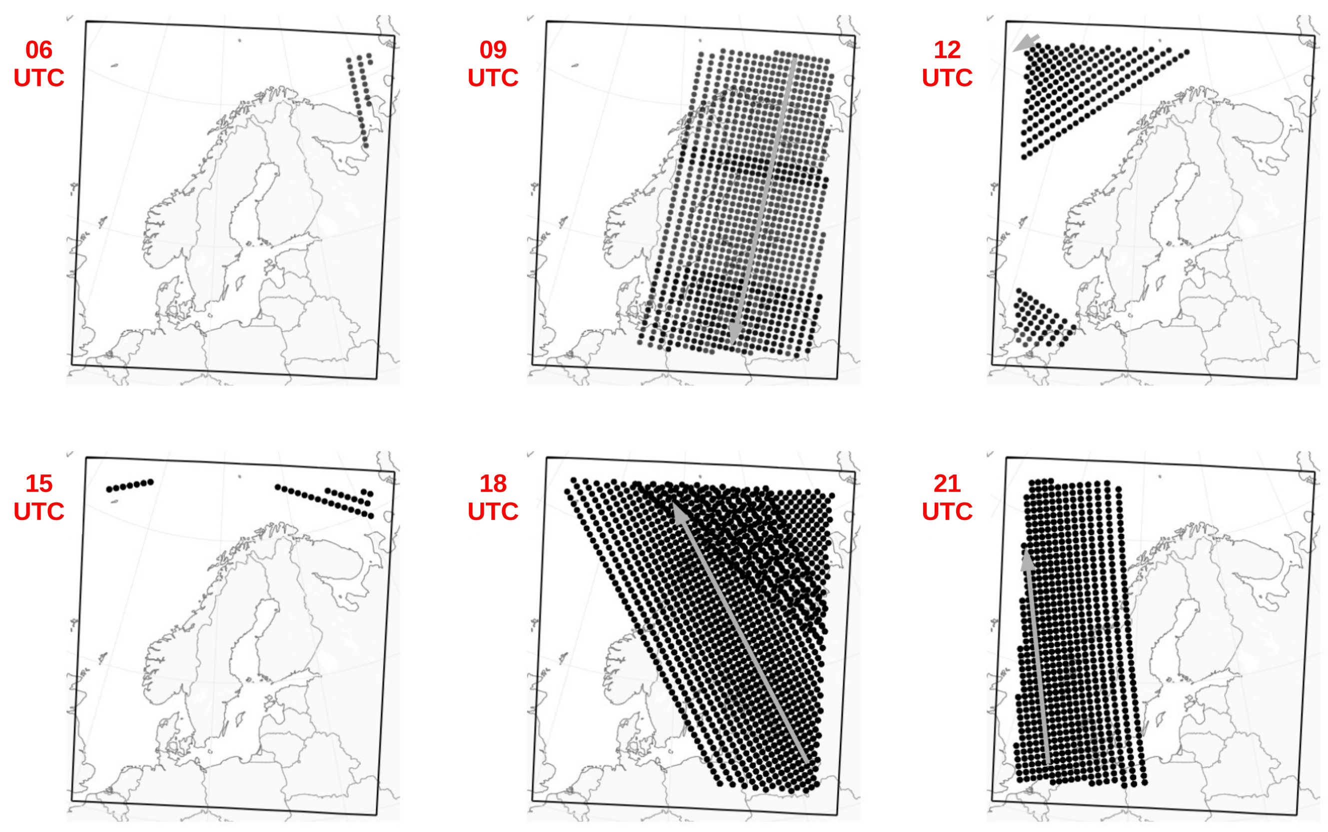
The use of satellite data in the operational NWP systems at MetCoOp

Meteorological Co-operation on Operational NWP (MetCoOp) is a Nordic-Baltic collaborative project that maintains and develops operational high-resolution NWP in Northern Europe. They produce hourly updates in two operational suites: the MetCoOp Ensemble Prediction System (MEPS) and MetCoOp Nowcasting System (MNWC), that both operate in a 2.5 km horizontal grid and perform analysis of upper-air meteorological fields using the three-dimensional variational method (3D-Var). MEPS is a lagged ensemble that refreshes 30 members over each six-hour interval.

As of Spring 2025, the use of microwave and infrared sounder data at MetCoOp is as follows:

* Advanced Microwave Sounding Unit -A (**AMSU-A**; channels 5–9) – *NOAA-18*; *NOAA-19*; *Metop-B*; *Metop-C*
* Advanced Technology Microwave Sounder (**ATMS**; channels 6–10 and 18–22) – *NOAA-20*; *NOAA-21*
* Microwave Humidity Sounder (**MHS**; channels 3–5) – *NOAA-19*; *Metop-B*; *Metop-C*
* MicroWave Humidity Sounder -2 (**MWHS-2**; channels 11–15) – *FengYun-3D*; *FengYun-3E*
* Cross-track Infrared Sounder (**CrIS**; a selection of 40 channels from the long-wave and mid-wave infrared absorption bands) – *NOAA-20*
* Infrared Atmospheric Sounding Interferometer (**IASI**; a selection of 55 channels from the long-wave and mid-wave infrared bands) – *Metop-B*; *Metop-C*

In regional NWP systems, the assimilation of polar-orbiting satellite data is complicated by the difficulty to maintain robust bias correction information. While the variational bias correction schemes have proven efficient and reliable in global NWP, in regional applications their performance is compromised because of the lack of consistent and comprehensive sampling of all relevant predictor variables. The figures below demonstrate typical data coverage from one satellite at different analysis hours inside the MetCoOp domain: in this example, the sampling of the full range of nadir viewing angles is almost complete at 09 and 18 UTC, but seriously limited at other analysis times.



*Dots indicate locations of **AMSU-A** observations from the Metop-C satellite in three-hour time windows centred at synoptic hours on 1 February 2023. Gray arrows indicate the approximate ground tracks of the satellite. The rectangle shows the boundaries of the computing domain at MetCoOp.*

Experiments using the HARMONIE-AROME* 3D-Var assimilation system

Data assimilation experiments are conducted using the HARMONIE-AROME reference version Cy43h2.1, that was operational at MetCoOp in 2021–2022. The experiments make use of the MetCoOp computing domain and they cover the 46-day period 1 August – 15 September 2021. The use of satellite data is made similar to the operational usage at the time.

Three pairs of runs are conducted to mimic the operational practice to launch a new set of three-hourly cycled members each hour. The notation applied to the three pairs is as follows:

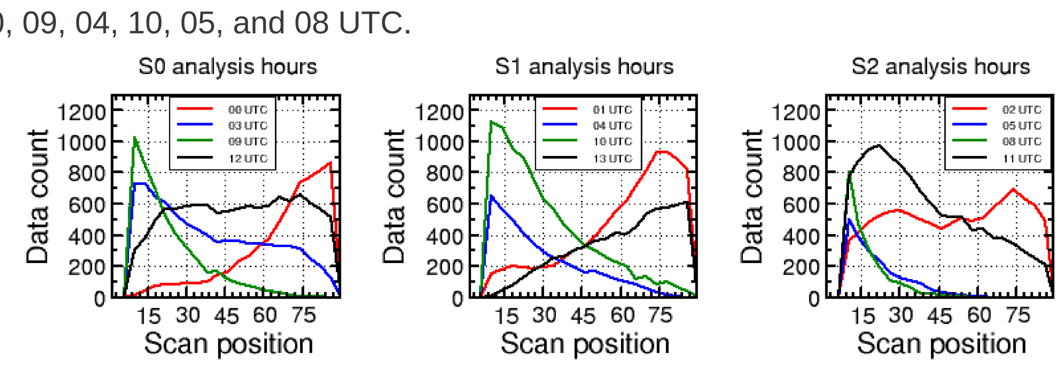
- **S0**: analyses at the synoptic hours (that is, 00, 03, …, 21 UTC)
- **S1**: analyses offset by +1 hour from S0 (that is, 01, 04, …, 22 UTC)
- **S2**: analyses offset by +2 hours from S0 (that is, 02, 05, …, 23 UTC)

Each of the three pairs includes
(1) a **control run** that applies VarBC in **the conventional MSDU setup**, and
(2) a **test run** that applies VarBC using **the SSDU method**.

All the runs are performed in the deterministic mode.

The success of the SSDU approach will depend on the ability to select the satellite-specific daily update hours appropriately. Ideally, the updates should happen at times when the data coverage inside the domain is comprehensive and substantial amounts of data are collected at all scan positions of the satellite instruments.

In the example of NOAA-20 shown here, the sampling is the most homogeneous at 12, 01, and 02 UTC; these are the analysis hours chosen for the daily updates in the SSDU test runs. The sampling is particularly skewed at 00, 09, 04, 10, 05, and 08 UTC.



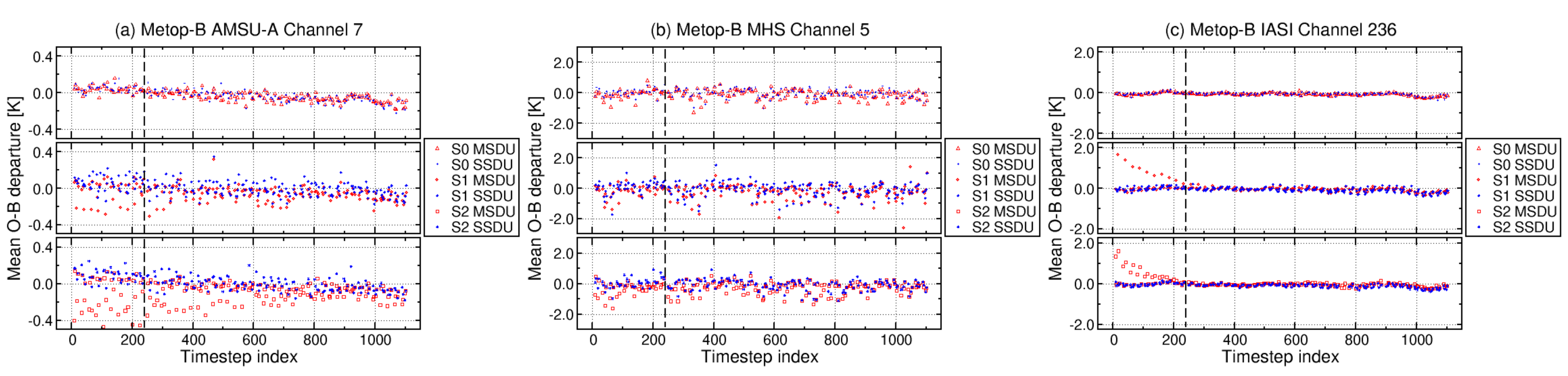
Data count inside the MetCoOp domain as collected from the NOAA-20 ATMS instrument in August 2021. The data are shown as a function of the instrument scan position index. Different line colours represent different analysis hours as shown in the legends.

The VarBC update hours selected for each satellite in the SSDU test runs are as shown in the Table below.

Satellite	***Update time*** S0 (UTC)	***Update time*** S1 (UTC)	***Update time*** S2 (UTC)
NOAA-18	12	19	20
NOAA-19	06	07	17
S-NPP	12	01	02
NOAA-20	12	01	02
FengYun-3D	12	01	02
Metop-B	09	19	20
Metop-C	09	19	20

* HARMONIE-AROME: HIRLAM-ALADIN Research on Mesoscale Operational NWP in Europe – Applications of Research to Operations at Mesoscale
HARMONIE-AROME reference versions are provided by the international High-Resolution Limited Area Model (HIRLAM) programme.

Outcome #1: VarBC coefficients converge faster when using the SSDU method instead of the conventional MSDU setup

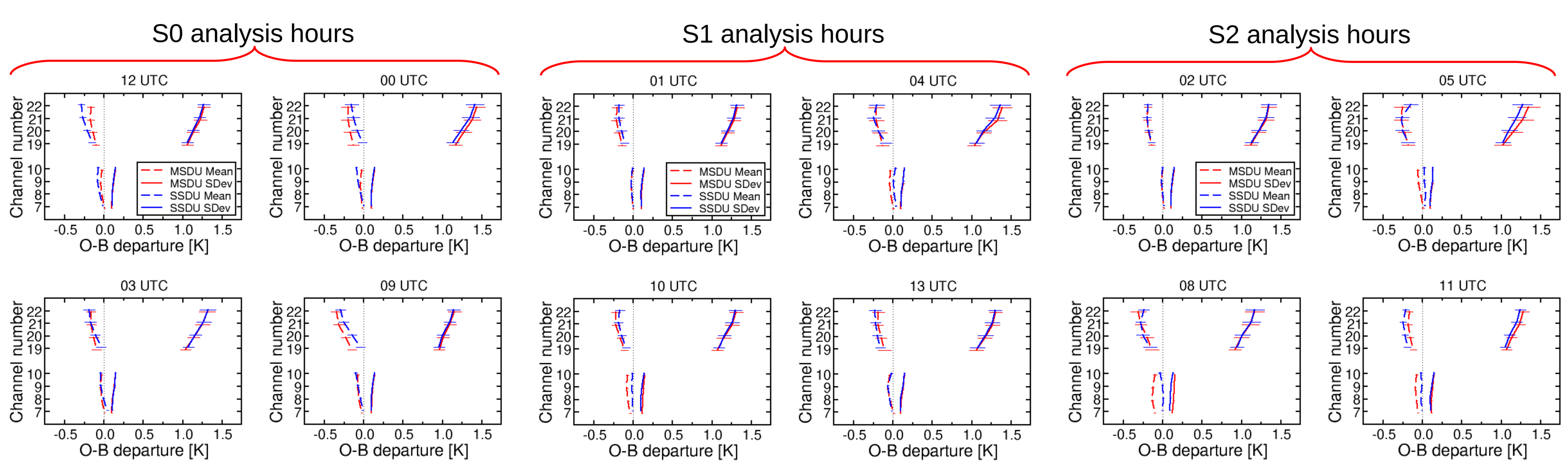


Time series of mean observation minus background (O-B) departure in selected sounder channels. Each panel shows the data fit at the S0 (top), S1 (middle), and S2 (bottom) analysis hours. Red symbols indicate the scores in the MSDU-based control runs. Blue symbols indicate the scores in the SSDU-based test runs.

The figures show time series of the mean observation minus background departure in representative sounder channels, as computed from the conducted data assimilation experiments. We conclude the following:

- At the S0 analysis hours (top panels), the statistics in the MSDU (red) and SSDU (blue) runs are similar. However, the humidity-sounding channel (panel (b)) shows isolated signals of poor data fit. Closer inspection reveals that these tend to occur mostly in the MSDU run and only at either 12 or 21 UTC. At these analysis hours the data coverage from Metop satellites in the MetCoOp domain is relatively poor.
- Both S1 (middle) and S2 (bottom panels) analysis hours indicate numerous cases where the MSDU run fails to provide a good data fit. These occur mostly during the first half of the experiment time period. The data fit in the SSDU run appears consistent throughout the time series.
- The data fit is exceptionally poor at time steps 1–200 specifically in the infrared sounder channel (panel (c)) in the MSDU run at the S1 and S2 analysis hours. This feature originates from the initialization of the VarBC coefficients to the state valid at the S0 hours at the start of the experiment runs. It is remarkable that this initial shock appears only in the MSDU runs. There are no such indications in the SSDU test runs.

Outcome #2: The SSDU method improves the data fit statistics at those analysis hours when the sampling of instrument scan positions is most skewed



O-B departure statistics in the actively-assimilated channels of NOAA-20 ATMS. The scores are aggregated separately for each analysis hour over the last 36 days of the experiment runs. Mean (dashed) and standard deviation (solid) of the departure are shown for the MSDU-based control runs (red) and the SSDU-based test runs (blue).

The figures show the data fit statistics for the NOAA-20 ATMS instrument:

- Channels 7–10 are sensitive to the temperature in the troposphere and stratosphere
- Channels 19–22 are sensitive to the humidity in the troposphere

For most parts, the data fit statistics are similar in the MSDU-based control runs (red lines) and the SSDU-based test runs (blue lines). Slight differences in the performance can be pointed out as follows:

- The SSDU-based test run shows superior performance at 00, 09, 04, 10, 05, and 08 UTC. Most often the superiority appears in the mean departure being closer to zero.
- The SSDU-based test run performs worse than the control run at 12 and 13 UTC.
- At 01, 02, 03, and 11 UTC, there either is very little difference or the superiority is controversial.

It is noteworthy that the daily analysis hours, at which the SSDU method shows an improvement in the data fit, match those hours at which the data coverage from the NOAA-20 satellite is the most skewed towards either very small or very large scan position indices. In contrast, there is little to no benefit from applying the SSDU method at those hours when the sampling of instrument scan positions is homogeneous.