Guidelines on Data Assimilation

Christoph Schraff, July 2021

Summary

The **aim** of research and development in the field of data assimilation (DA) is to further improve and extend the KENDA assimilation system and use of observations in view of better convective-scale deterministic and ensemble forecasts with ICON-LAM notably for weather-related quantities. Firstly, this is addressed by extending the use of **observations**, with a focus on those related to **cloud** and **precipitation**, but also those in the **planetary boundary layer** and near the **surface**. Secondly, the current **LETKF** scheme will be complemented by a variational component for an ensemble-variational analysis (**EnVar**) and, at least experimentally, by a **particle filter** (PF) in order to address certain limitations of the LETKF. To allow for data assimilation with ICON-LAM at low cost, tools will also be provided to perform a variational data assimilation without the need to run a (convective-scale) ensemble in the assimilation cycle.

Data Assimilation Algorithms

Background and Status

In the former priority projects KENDA (Km-scale ENsemble-based Data Assimilation) and KENDA-O (KENDA for high-resolution Observations), a 4-dimensional Local Ensemble Transform Kalman Filter (**4D-LETKF**) scheme following Hunt et al. (2007) has been developed for the **COSMO** model and ported to the **ICON-LAM** model. Latent heat nudging (**LHN**, Stephan et al., 2008) of radar precipitation has been integrated into the KENDA-LETKF analysis cycle. The main purpose of this data assimilation (DA) system called **KENDA** (Schraff et al., 2016) is to provide the initial conditions both for deterministic and ensemble forecasting on the convective scale (i.e. with 1 – 3 km model mesh size). Foci of interest include improved forecasts of high-impact weather, e.g. in convective situations or for fog and low stratus conditions, and especially in the very short time range. It is running operationally now at several COSMO member states (Germany, Switzerland, Italy), providing the initial conditions for operational deterministic and ensemble forecasts.

Compared to the previous observation nudging scheme, the introduction of KENDA has resulted overall in large improvements in the EPS forecasts and more moderate but still significant improvements in the deterministic forecasts, particularly of convective precipitation in summer. This **benefit of KENDA for COSMO forecasts prevailed** throughout **except** for two aspects: a) **low stratus**, which was degraded in some cases, notably over the Swiss Plateau, and b) 2-m humidity which was found to be slightly degraded (in MeteoSwiss verifications) as 2-m humidity observations have been assimilated in the COSMO model only in the nudging scheme up to now. (For ICON-LAM however, their use in KENDA has led to clear improvements). Apart from the 2-m humidity data, these overall very positive results were achieved using a similar set of observations. The ability of the KENDA-LETKF to derive and use **flow-dependent background error covariances** for the computation of the analysis increments certainly plays a major role for this. Compared to the nudging scheme with its limitations to assimilate indirect observations, the LETKF also offers much better **prospects to use additional data**. The greatest success in this respect so far has been the recent operational introduction of radar radial velocity and reflectivity volume data in the LETKF, with clear benefits found in the main tests (at DWD and at ARPAE (Gastaldo et al., 2018)). With this, the COSMO consortium is the **first** one (in Europe and at national level worldwide to our knowledge) **to assimilate operationally 3-D radar reflectivity directly** without any retrieval approach.

Besides this great potential of the **LETKF**, there are also certain **limitations** and issues (some of which may contribute to the shortcomings with low stratus), e.g.:

- In the LETKF, the available number of degrees freedom to fit the observations within the localisation scale does not exceed the number of ensemble members. This **rank deficiency** problem poses a **limitation for** the use of **high-resolution data**. Enhancing the effective degrees of freedom in the analysis by reducing the localisation scale may lead to imbalances and incomplete consideration of real background error covariances. The variational approach and inclusion of climatological covariances may mitigate this problem.
- The LETKF deploys **localisation in observation space** by increasing the errors for observations further away from a given analysis grid point and discarding the data beyond a certain distance limit. This poses a problem for the treatment of **non-local observations** such as satellite radiances and GNSS (Global Navigation Satellite System) total delay data. It also does not allow for localisation between the analysed variables, i.e. the local cross-covariances given by the ensemble cannot be reduced in the analysis equation. (Extensions of the algorithm like multi-step approaches or (univariate) additive background covariance inflation may mitigate the latter aspect).
- **Systematic model errors** leading to model bias are difficult to account for in data assimilation, and current operational schemes do not do this (in the troposphere, except that bias corrections can be applied to observations in order to unbias the latter with respect to the model). This problem is prominent for ensemble Kalman filters (EnKF) which in this case often underestimate the background errors as they derive their covariances from the equally (or similarly) biased ensemble members.

  Generally, the ability of EnKF to correct for errors in the first guess (mean) depends on whether the assumed first guess error covariances (both variances and correlations) derived from the ensemble perturbations (i.e. the deviations of the ensemble members from the ensemble mean) reflect the true first guess errors well or not. Thus, EnKF tend to **rely** more strongly **on** the **quality of the first guess** (ensemble) than schemes that make use of flow-independent error correlations.
- The LETKF makes the Gaussian assumption. The probability density distributions of the first guess or the observation errors, however, are often non-Gaussian and (for first-guess errors) multi-modal in the convective scale and in particular for weather related variables and observations such as those related to cloud and precipitation. Furthermore, even though the LETKF requires only the application of the full non-linear observation operator[1], it takes implicitly a linear assumption. In cases of **non-linearity** and **non-Gaussianity**, the LETKF analysis is therefore not optimal.

[1] The terms ‚observation operator‘ and ‚forward operator‘ are used synonymously for the computation of a simulated observation (i.e. model equivalent related to an observation) from the model state.

Other data assimilation methods offer advantages for some of these issues but also have their own disadvantages compared to the LETKF:

- The **particle filter** (PF) addresses particularly **non-Gaussianity** and **nonlinearity**. At DWD, two PF variants, a Localized Adaptive Particle Filter (LAPF, see Potthast et al., 2019) and a Localized Markov Chain Particle Filter (LMCPF), have been developed. They compute an ensemble transform matrix locally as in the LETKF, but replace the LETKF update in ensemble space by a PF update and resampling in ensemble space. In this way, they overcome the problems of filter collapse (i.e. having far too little spread) and divergence (from the true state) that are typical for traditional PF applied to high-dimensional problems. Potthast et al. (2019) were the first to run a PF stably over a long period for a global NWP system in a quasi-operational setting and obtained forecast quality comparable to that from the LETKF. The LMCPF has also been found to work similarly for the COSMO model and ICON-LAM on the convective scale.
- In variational data assimilation (**Var**), certain aspects of non-linearity and non-Gaussianity can also be better handled than in the LETKF. The cost function may include nonlinear terms, and variational quality control (VarQC) helps to deal with observation departure statistics with fat tails by weighting down outliers. The iterative approach can improve the analysis in the presence of a non-quadratic shape of the cost function; on the other hand, the tangent linear and adjoint are required for the observation operators and in 4DVar for the NWP model. Furthermore, Var does **not** suffer from the **rank deficiency** problem and may potentially allow for using data at higher density (depending on the background error covariances applied), and it avoids the localisation in observation space.

  At DWD, an EnVar scheme following ideas from Buehner et al. (2005) has been developed and is running operationally for global DA / NWP. Technically, it can be seen as a natural evolution of the 3DVar scheme combining the climatological background error covariances (B-matrix) of 3DVar (currently with a weight of 30 %) with the B-matrix derived from the ensemble perturbations of the associated LETKF (with a weight of 70 % as obtained by tuning). In this way, this **hybrid EnVar** (‘hybrid’ commonly denotes this type of blending of climatological and flow-dependent B-matrices) combines the main advantages of the two methods, Var and LETKF: It avoids or mitigates the above-mentioned shortcomings of the LETKF and at the same time benefits from the flow-dependent ensemble B-matrix. As a result, the forecasts started from EnVar analyses have been found to be significantly superior to those from the former 3DVar and to those from the analysis mean of the LETKF (which however is run at three times lower resolution). A four-dimensional version of EnVar known as 4D-EnVar is also under development at DWD.

  Note that **EnVar** provides only a single analysis **for** a **deterministic forecast** (unless a whole ensemble of EnVar’s is run) and still relies on the LETKF for the ensemble initial conditions. However, it is possible to apply ‘re-centering’ of the analysis ensemble around the EnVar deterministic analysis (by shifting all ensemble members in the model space in the same way such that the mean of the re-centered analysis ensemble coincides with the deterministic analysis). For the global ICON system, this results in large improvements of the **EPS forecasts**, when **re-centering** was applied to the LETKF ensemble during the data assimilation cycle and to the final analysis ensemble. Small improvements are obtained even in the deterministic analysis and forecast which obviously benefit from the improved B-matrix derived from the re-centered ensemble.

Agenda on Algorithms

A focus will be on the development of algorithms that may complement or replace the current LETKF:

1. **EnVar**

   The above mentioned results and considerations strongly promote the development and test of **EnVar** for the convective scale (with ICON-LAM). This holds although EnVar produces only a deterministic analysis and the focus of the COSMO consortium is on convective-scale EPS which would still rely on the (possibly re-centered) LETKF. For DWD, using EnVar operationally for ICON-LAM would also further harmonize the global and regional DA systems and increase synergies in research, development, and maintenance. It is worth mentioning that this development moves in parallel to the major trend of methods used at other major centres or European consortia for convective-scale DA: The UK Met Office deploy 1-hourly 4DVar (currently with climatological B-matrix, but working towards a hybrid B-matrix, see Milan et al., 2020). MeteoFrance, HIRLAM (HARMONIE-AROME), and JMA all use 3DVar but are working towards 4DVar and testing hybrid EnVar and 4D-EnVar (Gustafsson et al., 2018). NOAA-NCEP introduced hybrid EnVar (in HRRRDAS: High-Resolution Rapid Refresh DA System) in late 2020.

   Despite DWD's encouraging experience with EnVar and recentering of the ensemble in the global system, it has yet to be seen, however, whether EnVar will really be able to improve upon the current KENDA-LETKF, for several reasons:

   – At the **convective scale**, **balances** are much **more flow-dependent** and unknown than at larger scale. Therefore, adding a climatological part to the ensemble B-matrix in a hybrid EnVar will likely have less benefit than at global scale, if at all.

   – The standard EnVar (currently used for the global system at DWD) is **3-dimensional**. The time dimension for the computation of the first guess departures and for the ensemble B-matrix in the current 4D-LETKF of KENDA would be lost. Past tests with KENDA have shown that 3D-LETKF does not perform as well as 4D-LETKF even though the difference was smaller than expected.

     It is worth mentioning however that a first version of a **FGAT-EnVar** (FGAT: First Guess at Appropriate Time) has already been developed at DWD. This computes the model equivalents to the observations by including temporal interpolation of the fields from several time slots of the first guess run to the observation times. Furthermore, a **4D-EnVar** is under development (focusing first on the global system) which will account additionally for the time dimension in the ensemble B-matrix. Note that unlike classical 4DVar, 4D-EnVar does not require the tangent linear and adjoint of the forecast model (nor an approximation of it).

   – A part of the performance gains from **re-centering** the ensemble in the global ICON system can be explained by the three times higher (horizontal) resolution of the deterministic analyses and forecasts compared to the ensemble. Current plans for convective-scale configurations however do not envisage different resolutions for the deterministic and ensemble runs. Due to the small scales, strong non-linearities and non-Gaussianities involved in situations with (explicitly simulated) deep convection, it is not clear whether re-centering of the ensemble will work beneficially at all at the convective scale (in convective situations). If not, the potential for improvements from EnVar would be limited to the deterministic forecasts (except for secondary effects e.g. from VarQC).

– Using the framework of LETKF has allowed the COSMO consortium to be the first (to our knowledge) to introduce the **direct assimilation** of **3D radar reflectivity** operationally (albeit still in combination with latent heat nudging (LHN)). The other consortia and major centers, all deploying some variants of variational DA (Var) as primary algorithm, chose to resort to other methods for their operational applications up to now. Meteo-France, HIRLAM, and JMA assimilate columns of relative humidity pseudo-observations derived from reflectivity (Wattrelot et al., 2014). NOAA-NCEP deploy a particular latent heating method for their convective-scale HRRR (High-Resolution Rapid Refresh) model (Benjamin et al., 2016). The Met Office applies LHN for their UKV model even though the development of radar reflectivity assimilation within 4DVar is at an advanced stage and set to replace LHN in autumn 2021 (Milan et al., 2020). **Issues** to assimilate reflectivity **in Var** directly include the choice / extension of the **control vector** (i.e. the 'analysed variables'), **explicit linearization** of strongly nonlinear processes and relationships, and (for classical or hybrid schemes) formulation of climatological background error covariances related to moisture and hydrometeors. COSMO will also face these issues when tempting to introduce EnVar and will need to come up with a practical solution to assimilate these data.

  This kind of issue is prominent for radar reflectivity, however it may also be present for other types of data, and quite certainly for some observation operators developed and tested in experiments with the LETKF. This includes the assimilation of nowcast objects and the assimilation of cloud top height (based on SEVIRI (Spinning Enhanced Visible and InfraRed Imager) and radiosonde data) as implemented experimentally in a former KENDA task (Schomburg et al., 2015). Since **Var requires** the tangent linear and the **adjoint of** the **observation operators**, it can make their formulation and implementation much more intricate and restricted than LETKF or PF.

  As a result, the benefit from EnVar can be severely hampered if these types of data, often related to cloud and precipitation, are to be assimilated.

2. CEnVar / 3DVar for avoiding the need to run an ensemble

Yet, there is another important reason for the COSMO consortium to develop and provide Var. For the COSMO model, observation nudging has allowed institutes from member states or licensees with quite limited computational resources to carry out data assimilation as the nudging scheme is cheap and does not require running a (convective-scale) ensemble in the data assimilation cycle in contrast to LETKF. As observation nudging is not available for ICON, a **cheap** surrogate **approach without the need to run a convective-scale ensemble** is needed, and this can be provided by Var.

– One solution is **classical 3DVar**, using a climatological B-matrix. For convective-scale NWP, this does not appear ideal because the real background errors are very flow-dependent at that scale, notably for high-impact weather. Furthermore, the climatology needs to be computed for each new model configuration in order to derive a suitable, adjusted B-matrix unless one wants to rely on the statistics of another, possibly similar configuration.

– Another solution is to derive an ensemble B-matrix from the ensemble members of a global ensemble (data assimilation) system. NOAA-NCEP have deployed this approach for their convective-scale system HRRR (Gustafsson et al., 2018) until late 2020 and also for their regional system RAP (RApid Refresh) since it has resulted in immediate significant improvements over classical 3DVar (Wu et al., 2017). Even though convective motions are missing, the method is still able to capture the larger-scale flow dependency of the background errors. To avoid misunderstandings, we will tag this approach **CEnVar** (Coarse-Ensemble Var). This avoids the need to run an ensemble (data assimilation) at convective scale, however the prerequisite is to have enough bandwidth to transfer the ensemble fields in time from the centre (e.g. DWD for ICON) that runs the global system and provides these fields.

  Naturally, there is also the possibility then to combine this ensemble information and a climatological B-matrix for a **hybrid CEnVar**.

3. **Particle Filter**

   Non-linearity and non-Gaussianity tend to play a major role for convective-scale DA, both with respect to the involved (convective) model processes with their short time scales and with respect to the higher priority to assimilate data related to cloud and precipitation with highly nonlinear observation operators. As the Particle Filter (PF) neither takes the assumption of Gaussianity nor of linearity, it is considered particularly promising to continue the longer-term oriented research on PF in the context of ICON-LAM. This will also attain a lot of attention and visibility in the scientific community.

Refinement of LETKF and estimation of uncertainties

Not only in the context of introducing new observation types but also in general, it will be necessary to **refine** and adapt features and settings of the main DA scheme, currently the **LETKF**. This relates particularly to the ensemble generation since a good estimate of the uncertainties and possible distribution of atmospheric states is essential for a good analysis and forecast quality. An **integrated approach** is required **for data assimilation and ensemble prediction**. Within the DA cycle, there is scope for improvement e.g. by tuning and upgrading covariance inflation. This may also include perturbed physics parameters or stochastic perturbations.

Use of observations

Short- to medium-term agenda

Besides developing these algorithmic aspects, it is also very important to **extend** further the set of **assimilated observation types** for operational purposes. Much of this will **continue** the research and development carried out in previous years. Besides Mode-S aircraft data, **radar** radial velocity and **reflectivity** (for earlier studies, see Bick et al., 2016; Gastaldo et al., 2018) could be introduced operationally and successfully into KENDA in 2020. The assimilation of **2-m temperature** and **humidity** from Synop stations (used to modify the atmospheric state) in ICON-LAM followed in 2021.

It is now important to carry on maintaining, refining, improving, and (geographically) **extending** the use of these data and possibly **adjusting** it to significant changes both in the model (e.g. 2-moment microphysics) and in the assimilation scheme as mentioned above. For some other observation types such as GNSS (Global Navigation Satellite System) zenith and slant total delay (**ZTD** and **STD**) and **all-sky VIS** and **IR satellite radiances** promising results have already been obtained, and it will possibly require only fairly few development steps and more testing to allow for operational use. There are also observation types such as all-sky satellite radiances where work is ongoing and more work is needed despite the progress achieved.

Even though any type of observation can be beneficial and important to be used, there are two classes of data for which further research, development, and extended use are considered particularly relevant for convective-scale analysis and forecasting of weather, in particular locally driven (or influenced) high-impact weather such as deep convection or low stratus and fog.

– The first class is high-resolution data related to humidity, cloud, and precipitation. Besides radar reflectivity (related to precipitation) and ZTD / STD (related to water vapour), a focus is on the **all-sky VIS** and **IR satellite radiances** (from the **SEVIRI** instrument onboard of geostationary Meteosat MSG satellites) in view of assimilation of information mainly on **cloudiness** (for an overview on all-sky satellite DA, see Geer et al., 2018; Kurzrock et al., 2018). Brightness temperatures from infrared (IR) window channels depend on cloud cover and cloud top height of all cloud types all day. IR water vapour (WV) channels are sensitive to water vapour and / or cloud cover and cloud top height of high and mid-level clouds also all day. In contrast, reflectances from visible (VIS) channels provide information only during daytime mainly on cloud cover, in particular of low clouds not seen well by other (IR) channels. The VIS reflectances are also sensitive to a wider range of the ice and liquid water path and thus contain information beyond the mere cloud top.

  For VIS data, a novel efficient forward operator (MFASIS: Method for FAst Satellite Image Synthesis) has been built, and promising results have already been obtained in a comprehensive case study for COSMO-KENDA by Scheck et al. (2020) and even for longer test periods with ICON-LAM at DWD. A novel statistical bias correction seeks to align the histograms of first-guess derived reflectances to the observed ones. Related to the assimilation of infrared (IR) **water vapour** (**WV**) **radiances**, Hutt et al. (2020) assimilated clear-sky WV radiances in COSMO-KENDA, while for cloudy (all-sky) IR WV radiances, Harnisch et al. (2016) introduced a model for observation errors, and Otkin and Potthast (2019) introduced a conditional nonlinear bias correction. Recent tests have shown clear forecast improvements mainly in the upper-tropospheric humidity from the assimilation of all-sky WV radiances despite having deployed a simplistic observation error model. Even though an operational use of these data appears to be possible in near future, more efforts e.g. on bias correction and observation error modelling will be required to exploit the potential of these data.

  In the longer term, these developments for SEVIRI will also need to be transferred to the Flexible Combined Imager (**FCI**) of **MTG** (Meteosat Third Generation) satellites the first of which is scheduled to start operations in autumn 2023.

– The second class of observations that appear particularly relevant are observations describing the planetary boundary layer (**PBL**), down to the **surface** level. These include high-frequency indirect measurements (often available every 30 min) from ground-based remote sensing devices such as microwave radiometers (**MWR** which contain low-resolution vertical profile information on temperature and humidity), **wind lidars**, **Raman lidars** (high-resolution vertical temperature and humidity profiles), and **DIAL** (Differential Absorption Lidar: humidity) but also direct measurements from **drones** and **towers**. Some of these devices are still rare and experimental, while others are increasingly available. Leuenberger et al. (2020) have shown that high-frequency profile observations from Raman lidars and meteo drones can improve analyses of the pre-convective boundary layer and the stable boundary layer favourable for fog, and subsequent forecasts of cloudiness and precipitation benefitted up to 9 hours lead time. There are national projects dealing with several of these data types which need to be coordinated and supported as part of the further KENDA development. For **screen-level observations** (2-m temperature, 2-m humidity), a novel approach for station-dependent bias correction has been devised which can also be adapted for observation error specification. It is found to require further refinements under certain condition and appears to promising to **deal with networks with very variable quality** (see also below).

Generally, bias correction and improving the specification of the random observation errors are important topics for many new or already used observation types. In the medium term, specifying correlated observation errors should be explored.

It should also be noted that the analysis step **interacts** strongly with the numerical model, particularly in an assimilation cycle with very short update frequency. Hence, introducing new observation types will often require **collaboration** between data assimilation, physical parameterization, model dynamics, and ensemble generation. Furthermore, **diagnostic** tools such as an extension of the ensemble-based variant of the FSOI (Forecast Sensitivity to Observations Impact) and their applicability to convective-scale systems (limited e.g. due to linearity assumptions, localisation, spin-up, impact of biases, etc.) should be studied.

Medium- to long-term perspective

There is a variety of novel observation systems which deliver new types of observations or will do so in the future. For some of them, only test data are currently available. To explore their potential and prepare their operational use in the long term, starting development and research on some of them soon is very important, while opportunities for exploratory research on others should be taken.

- A future data source considered highly important due to its vast amount and high horizontal and temporal resolution is the **MTG IRS** (Meteosat Third Generation Infrared Sounder) which is scheduled to be operationally available in 2024. Even though the ultimate goal might be the use of all-sky data, the main benefit from these hyper-spectral radiances on top of the cloud information from the imagers (SEVIRI, FCI) is expected from **profile** information on **temperature**, **humidity**, and vertical atmospheric stability. Preparations for the use of these data have started with a focus on the clear-sky radiances. Later, this can be extended to radiances sensitive to regions above cloud or affected (slightly) by clouds, keeping the focus on deriving information of temperature and humidity.
- **Lightning and polarimetric parameters** from dual-polarisation radars, **weather cameras**, IR cameras, web cams are among a variety of **novel data types** which are completely different from those already assimilated. Their use typically requires large research efforts, for instance for the development of suitable observation operators. Taking opportunities for such efforts (e.g. involving third party funding) is encouraged.
- Weather-related data from sources not being operated by national or public regional weather services will increasingly be produced and potentially available. This includes data from devices or networks operated by **public entities** or **private companies** and so-called **citizen data** (or **crowd-sourced** data, i.e. data produced

by (typically many) individual entities). Some of these data sources provide quantities similar to the ones already assimilated, such as **roadside sensors**, **Netatmo** private weather stations, wind speeds measured at **wind mills**, etc. These data can greatly enhance the data density from the network operated by the weather services near the surface, and investment in their development is therefore clearly recommended. However, their **quality** is often poorer or at least **more variable** in terms of biases, random errors, outliers, meta data, availability, etc. This demands careful **bias correction**, **estimation of** random **observation errors**, and **quality control**, possibly based on machine learning approaches, and should be complemented by automated **monitoring**.

Further aspects

Machine learning

For only the past few years, neural networks and machine learning (ML) in general have been gaining rapidly increasing attention in the international NWP community. While such methods are not expected to replace data assimilation schemes (or even numerical models) for operational NWP as a whole in the foreseeable future, they are beginning to be explored and used successfully for certain aspects. In the context of KENDA, **neural networks** are used in a recent version of the MFASIS **observation operator** for VIS satellite radiances, and are also explored as part of an observation operator (or pre-processor) for images from weather cameras (in the ICamCloudOps project at DWD). Exploratory research (in collaboration with DWD) has started on **estimation of observation errors** by machine learning. The conditional online **bias correction** based on nonlinear regression and tested for 2-m temperature and humidity data may also be seen as a (simpler) example of ML. Another promising application of ML is **quality control**. It is expected that in particular (certain) citizen data and data from external providers will require or at least benefit from machine learning applied to some of the mentioned processing steps. Therefore, research on ML is highly recommended.

Addressing observation and model biases, and link to model physics

An assumption for optimality of current data assimilation schemes is that there are no systematic differences, hereafter called '**relative biases**', between observations and the model first guess (in observation space, i.e. after having applied the observation operators). Assimilating data with relative biases (i.e. with biased observation departures) can often lead to spin-up effects and degrade subsequent forecasts. The **source** of a relative bias may be systematic errors (biases) in either the **observations**, the **model** state, the **observation operator**, or a combination of them. It is often very **difficult to** identify the source(s) and **discriminate between** the **biases**. Collaboration with experts on the model and its parameterizations will be further enhanced to help identify and reduce such biases in the model and observation operators, and to help understanding the effects of such biases on the assimilation.

Within DA, it is recommended to further develop (also by the above-mentioned advanced techniques), refine and much more widely apply **bias correction** schemes **to observations**, based on observation minus first guess differences. This helps avoiding imbalances and spin-up effects. However, it will not correct biases in the model state which in some cases might even be gradually enlarged due to feedback effects with the bias correction.

Model biases themselves may also be addressed by applying DA techniques for online estimation of parameters in physical parameterizations. Particular **physical** (or external) **parameters** with very strong, **direct relationship to well-observed quantities** may be estimated, preferably by dedicated schemes. Such parameters are often related to the surface and soil. An example of this is the dynamic adaptation of certain external parameters (leaf area index, stomata resistance, fraction of sealed surfaces) in ICON which heavily influence 2-m temperature forecast errors in certain conditions. Implementing further similar schemes appears feasible.

Generally, the **control vector** (i.e. the ‘analysed variables’) of the main DA scheme (currently KENDA-LETKF) might be **augmented** with any uncertain parameters. This will update the parameters such that the whole augmented state is optimized with respect to the observations. However in such an all-in-one-step approach, the estimated parameter values can be influenced by all kinds of errors, and past idealised studies have shown that the true parameter values will normally not be retrieved successfully. Due to these imponderables, this method is likely **beyond the focus** of research by COSMO at least in the short to medium term.

Lower Boundary Conditions

Some existing analysis tools related to the **lower boundary conditions** need to be consolidated and refined, including the soil moisture analysis using soil moisture data. An important step will be their port to resp. re-write **in the framework of** the **DACE** code (Data Assimilation Code Environment) which will make their maintenance and further development (including upgrades to more advanced analysis methods) much more efficient. This refers to the **sea surface temperature** (**SST**) analysis, the **snow depth** analysis, the **soil moisture analysis** and the related 2-m temperature analysis. To meet these objectives, variational schemes based on the framework used for the atmospheric analysis will be adopted for SST and snow depth. For snow depth however, a complementary approach based on the new multi-layer SNOWPOLINO snow module will also be developed. With a longer-term perspective, efforts should also go into a **coupled data assimilation** of the **soil** (moisture and possibly temperature) **and atmosphere**.

Computational aspects: GPU

Several NWP centres in COSMO already use HPC (High Performance Computing) platforms that use GPU’s (Graphics Processing Units) and run the NWP model on it. It is expected that the role of GPU’s will further increase in the future. In order to prepare the operational usability of the data assimilation codes based on DACE on new computer architectures, it is therefore indispensable to **explore** and understand which parts of the DACE code are feasible to be ported to **GPU** and how. An exploratory project has started at DWD in 2021 but more efforts will likely be required in the longer term for the actual port.

Risks

The objectives in data assimilation do not consist of one main single task or aim which can be fulfilled or not. Therefore, and since the current KENDA system provides a good basis, the risk of an overall failure is low. Instead, there is a variety of individual tasks, each one dealing with its own scientific issues (many of which are mentioned above or in the respective tasks) and risks. For instance, there is a general risk that the **gain in forecast quality** from the use of any of the addressed observation types may be **difficult to prove** and does not meet the expectations. As mentioned, there is e.g. also a risk that EnVar does not perform better than LETKF, in particular if observations related to cloud and precipitation, such as radar reflectivity or VIS radiances, are used.

Besides the scientific aspects, there is always a risk that developments are slowed down or even ceased in case of cut downs on human resources.

Links to related topics

It is noted that e.g. at DWD there are activities or plans on further topics in the field of data assimilation. However, as they do not aim to provide initial conditions for short-range NWP they are currently seen **beyond the scope** of development in the context of the COSMO consortium and hence beyond the objectives in these guidelines. These topics include **re-analyses**, (weakly or strongly) **coupled DA** for coupled atmospheric and **ocean** modelling, (DA for) seamless integrated prediction from nowcasting to very short-range forecasting, ultra-rapid data assimilation (URDA) for **nowcasting**, and DA for **aerosols**.

References

Benjamin SG, Weygandt SS, Brown JM, Hu M, Alexander CR, Smirnova TG, Manikin GS. 2016. A North American hourly assimilation and model forecast cycle: the Rapid Refresh. *Mon. Wea. Rev.,* **144**: 1669–1694, doi: 10.1175/MWR-D-15-0242.1.

Bick T, Simmer C, Trömel S, Wapler K, Hendricks Franssen H, Stephan K, Blahak U, Schraff C, Reich H, Zeng Y, Potthast R. 2016. Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale. *Q. J. R. Meteorol. Soc.* **142**: 1490–1504, doi: 10.1002/qj.2751

Buehner M. 2005. Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Q. J. R. Meteorol. Soc.,* **131**: 1013–1043, doi: 10.1256/qj.04.15

Gastaldo T, Poli V, Marsigli C, Alberoni PP, Paccagnella T. 2018. Data assimilation of radar reflectivity volumes in a LETKF scheme. *Nonlin. Processes Geophys.,* **25**: 747–764, doi: 10.5194/npg-25-747-2018.

Geer AJ, Lonitz K, Weston P, Kazumori M, Okamoto K, Zhu Y, Liu EH, Collard A, Bell W, Migliorini S, Chambon P, Fourrié N, Kim M-J, Köpken-Watts C, Schraff C. 2018. All-sky satellite data assimilation at operational weather forecasting centres. *Q. J. R. Meteorol. Soc.* **144**: 1191–1217, doi: 10.1002/qj.3202.

Gustafsson N, Janjic T, Schraff C, Leuenberger D, Weissmann M, Reich H, Brousseau P, Montmerle T, Wattrelot E, Bučánek A, Mile M, Hamdi R, Lindskog M, Barkmeijer J, Dahlbom M, Macpherson B, Ballard S, Inverarity G, Carley J, Alexander C, Dowell D, Liu S, Ikuta Y, Fujita T. 2018. Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Q. J. R. Meteorol. Soc.* **144**: 1218–1256, doi: 10.1002/qj.3179.

Harnisch F, Weissmann M, Periáñez Á. 2016. Error model for the assimilation of cloud-affected infrared satellite observations in an ensemble data assimilation system. *Q. J. R. Meteorol. Soc.*, **142**: 1797–1808, doi: 10.1002/qj.2776.

Hunt BR, Kostelich EJ, and Szunyogh I. 2007. Efficient data assimilation for spatiotemporal chaos: a Local Ensemble Transform Kalman Filter. *Physica D*, **230**: 112 – 126, doi: 10.1016/j.physd.2006.11.008.

Hutt A, Schraff C, Anlauf H, Bach L, Baldauf M, Bauernschubert E, Cress A, Faulwetter R, Fundel F, Köpken-Watts C, Reich H, Schomburg A, Schroettle J, Stephan K, Stiller O, Weissmann M, Potthast R. 2020. Assimilation of SEVIRI water vapour channels with an ensemble Kalman filter on the convective scale. *Frontiers in Earth Science*, **8**, 70, doi: 10.3389/feart.2020.00070.

Kurzrock F, Cros S, Ming F, Otkin J, Hutt A, Linguet L, Lajoie G, Potthast R..2018. A review of the use of geostationary satellite observations in regional-scale models for short-term cloud forecasting. *Meteorol. Zeitschrift* **27**: 277 – 298, doi: 10.1127/metz/2018/0904.

Lei L, Whitaker JS, Bishop C. 2018. Improving assimilation of radiance observations by implementing model space localization in an ensemble Kalman filter. *J. Adv. Model. Earth Syst.* **10**: 3221 – 3232, doi: 10.1029/2018MS001468

Leuenberger D, Haefele A, Omanovic N, Fengler M, Martucci G, Calpini B, Fuhrer O, Rossa A, 2020. Improving high-impact numerical weather prediction with lidar and drone observations, BAMS, in press doi: 10.1175/BAMS-D-19-0119.1.

Milan M, Macpherson B, Tubbs R, Dow G, Inverarity G, Mittermaier M, Halloran G, Kelly G, Li D, Maycock A, Payne T, Piccolo C, Stewart L, Wlasak M. 2020. Hourly 4D-Var in the Met Office UKV operational forecast model. *Q. J. R. Meteorol. Soc.* **146**: 1281 – 1301, doi: 10.1002/qj.3737.

Otkin J, Potthast R. 2019. Assimilation of all-sky SEVIRI infrared brightness temperatures in a regional-scale ensemble data assimilation system. *Mon. Wea. Rev.* **147**: 4481 – 4509, doi: 10.1175/MWR-D-19-0133.1

Potthast R, Walter A, Rhodin A. 2019. A Localized Adaptive Particle Filter (LAPF) within an operational NWP framework. *Mon. Wea. Rev.* **147**: 345 – 362, doi: 0.1175/MWR-D-18-0028.1.

Scheck L, Weissmann M, Bach L. 2020. Assimilating visible satellite images for convective-scale numerical weather prediction: A case study. *Q. J. R. Meteorol. Soc.* **146**: accepted, doi: 10.1002/qj.3840.

Schomburg A, Schraff C, Potthast R. 2015. A concept for the assimilation of satellite cloud information in an Ensemble Kalman Filter: single-observation experiments. *Q. J. R. Meteorol. Soc.* **141**: 893 – 908, doi: 10.1002/qj.2407.

Schraff C, Reich H, Rhodin A, Schomburg A, Stephan K, Periáñez A, Potthast R. 2016. Kilometre-scale ensemble data assimilation for the COSMO Model (KENDA). *Q. J. R. Meteorol. Soc.* **142**: 1453–1472, doi: 10.1002/qj.2748.

Stephan K, Klink S, Schraff C. 2008. Assimilation of radar-derived rain rates into the convective-scale model COSMO-DE at DWD. *Q. J. R. Meteorol. Soc.* **134**: 1315–1326, doi: 10.1002/qj.269.

Wattrelot E, Caumont O, Mahfouf JF. 2014. Operational implementation of the 1D+3D-Var assimilation method of radar reflectivity data in the AROME model. *Mon. Wea. Rev.* **142**: 1852–1873, doi: 10.1175/MWR-D-13-00230.1.

Wu W-S, Parrish DF, Rogers E, Lin Y. 2017. Regional ensemble-variational data assimilation using global ensemble forecasts. *Wea. Forecasting,* **32**: 83–96, doi: 10.1175/WAF-D-16-0045.1.