Australian Government
Bureau of Meteorology

National Analysis System

Monika Krysta, Peter Steinle, Susan Rennie, Jin Lee, Imtiaz Dharssi, Andy Smith, Chris Bridge, Thomas Coleman

November 2024

Bureau Research Report - 102

National Analysis System

Monika Krysta, Peter Steinle, Susan Rennie, Jin Lee, Imtiaz Dharssi, Andy Smith, Chris Bridge, Thomas Coleman

Bureau Research Report No. 102

November 2024

National Library of Australia Cataloguing-in-Publication entry

Authors: Monika Krysta, Peter Steinle, Susan Rennie, Jin Lee, Imtiaz Dharssi, Andy Smith, Chris Bridge, Thomas Coleman

Title: National Analysis System

ISBN: 978-1-925738-88-9

ISSN: 2206-3366

Series: Bureau Research Report – BRR102

Enquiries should be addressed to:

Author: Susan Rennie

Bureau of Meteorology

GPO Box 1289, Melbourne

Victoria 3001, Australia

susan.rennie@bom.gov.au

Copyright and Disclaimer

Table of Contents

List of Figures

List of Tables

Executive Summary

The National Analysis System (NAS) has been developed to enhance the situational awareness of operational forecasters. NAS will contribute to the nowcasting of severe weather events and automated weather forecasting. NAS will also reduce discrepancies between weather forecasts as presented on Bureau's media and the actual weather.

NAS is a Numerical Weather Prediction (NWP)-like system. Its backbone is constituted by a sequence of short-term forecasts followed by 4-dimensional variational data assimilation (4D-Var) tasks which produce initial conditions for subsequent forecasts. In addition, NAS produces two additional 3-dimensional analyses (3D-Var), which are scheduled before the 4D-Var. The 3D-Var tasks generate atmospheric analyses where the physical fields more closely fit the observations.

NAS underwent subjective and objective assessment. In the subjective assessment an operational forecaster investigated NAS' usefulness in the forecasting process and indicated positive impact, especially for winds. The objective assessment was performed via cross-validation. A 3D-Var analysis that fits closely to the observations, which is destined to be NAS' product, generally led to lower values of RMSEs than two other NAS analyses. It also gives comparable results to the existing Mesoscale Surface Analysis System (MSAS), which NAS is designed to replace.

The NAS prototype described here is ready to be transferred from development to a testbed environment, where it will be subject to further subjective assessment by forecasters as well as produce hindcast simulations to enable the calibration of downstream systems.

1. Introduction

Data assimilation has been traditionally used to provide an optimal estimate of an initial condition for a weather forecast. Different applications of data assimilation are nevertheless possible. This is even more the case as the inverse problem underpinning data assimilation does not admit unique solutions. This feature gives us freedom in adjusting weights assigned to the contributions of the regularisation term and the observation term in the mathematical formulation of objective analysis. The resulting procedure may hence be giving more weights to one or other term, according to the properties required from the solution.

In this broader framework of data assimilation, the Australian Bureau of Meteorology has developed a National Analysis System (NAS). As its name indicates the main product of the system is an atmospheric analysis. The need for the National Analysis System is driven by the internal customers' requirements to provide best representation of the current state of the atmosphere. This product shall be used by forecasters to increase their situational awareness in advent of severe weather events, and to provide a reference for other downstream applications. NAS is not the first system of this type developed by the Bureau. NAS' predecessor, MSAS, Mesoscale Surface Analysis

System, has been successfully operating in the Bureau for over a decade (see Glowacki et al., 2012). MSAS is, however, a two-dimensional system producing, as its name suggests, surface analyses. Moreover, MSAS uses an orphaned code base, based on statistical interpolation and lacks the advancements of variational data assimilation methodology.

The atmospheric analysis produced by NAS uses the variational data assimilation technology of existing NWP systems. However, the NAS configuration of the variational data assimilation differs from that of NWP as the relative weights of observations are adjusted to develop a new product drawing the fields much more closely towards observations. The resulting analysis overfitting the observations includes more of the very local information from observations, and so may not be suitable for NWP. Once operationalised, these analyses provide, real-time, hourly (or more often), high spatial resolution, 3D representations of the atmosphere and ensure physical consistency across variables throughout the atmosphere.

Similar systems have been implemented internationally such as the surface Real-Time Mesoscale Analysis (RTMA) based on two-dimensional variational data assimilation, 2DVAR, developed by National Centers for Environmental Prediction (NCEP). RTMA 'produces gridded fields of 2-m temperature, 2-m specific humidity, 2-m dewpoint, 10-m U and V wind components, and surface pressure' as reported on in De Pondeca et al. (2011). It is based on NCEP Gridpoint Statistical Interpolation (GSI) used in an incremental 2DVAR mode.

In parallel, gridded observations were developed in form of the Unrestricted Mesoscale Analysis (URMA) (see Raby J.W. et al., 2020).

The current 2D RTMA/URMA system is being expanded into a fully three-dimensional analysis of the atmosphere, namely 3D-RTMA (see Weygandt, S.S. et al., 2019). In this new application, more fields such as, Planetary Boundary Layer (PBL) height, precipitable water or ceiling, are diagnosed from the three-dimensional RTMA fields. The 3D-RTMA will in addition provide full-column representation of the standard meteorological fields (temperature, water vapour, wind, hydrometeors and aerosols). Furthermore, 3D-RTMA will include land-surface diagnostics as well as convective fields.

NAS will also provide a broad range of meteorological fields, coinciding with those output by limited area NWP systems. However, in the initial interactions with forecasters the necessity of high-quality surface temperature, dewpoint and winds were emphasised as the building blocks of any forecast, regardless of what type of a severe weather event may be developing.

2. System description

NAS is a limited area modelling system covering Australia and nearby waters. Although NAS' emphasis is on atmospheric analysis, conceptually it has much in common with ACCESS-C (see Rennie et al., 2022) and only few changes have been introduced with respect to a standard NWP system. NAS has been purposely designed in this way so that existing implementations could easily be re-used in the future. The underpinning NAS workflow is an hourly cycling suite, meaning that each hour a new forecast is initialized by assimilating new observations into the existing forecast. The updated forecast is then used an hour later in the next assimilation step and so on. Since this cycling component of the NAS relies on a numerical forecast, the choice of assimilation system is driven by the need to provide the best forecast. Hence this cycling phase uses 4D-Var to produce the initial conditions for the forecast.

NAS however differs from ACCESS-C by including additional variational analyses to produce analysis products with no subsequent forecast. Due to operational requirements these additional analyses need to run close to real time, and consequently can use only a limited set of observations. However, 4D-Var is too slow to meet operational time constraints, but 3D-Var provides an affordable, and suitable approximation for single analysis. For each cycle two steps are run. The first step is to provide larger scale adjustments, with a similar configuration to that used for NWP assimilation. The second step is to draw the analysis closer towards the observations, by reducing the observation error to give the observations more weight during the analysis. This second 3D-Var analysis, by giving the observations more weight than is warranted to initialise a weather forecast, will be referred to as overfitting. The overfitting analysis better captures any very local features at a site. This analysis is specific to NAS and constitutes the main product.

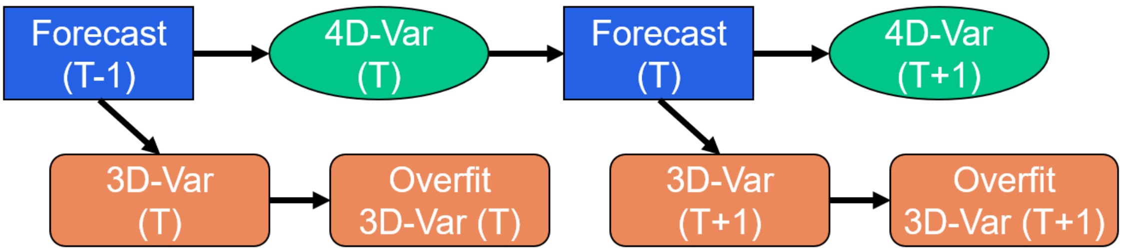


Figure 1. Simplified schematic of the sequence of forecast and data assimilation in each cycle. The forecast marks the end of the cycle (T), and provides the background for the data assimilation tasks in the next cycle (T+1).

2.1. Concept of NAS

2.1.1. NAS domain

The NAS domain roughly covers Australia's mainland including Tasmania. The orography field underlying the domain spans the following geographical coordinates: from 109.52 °E to 156.99 °E in the east-west direction and 46.78 °S to 8.82 °S in the north-south direction (Figure 2). For the needs of the atmospheric model, the domain is covered with a regular lat-lon mesh using high resolution horizontal grid spacing of 2.2 km. This grid spacing has been selected with consideration of available computing. In addition, Bureau had experience with this particular grid spacing as it is used in its limited area ensemble modelling, ACCESS-CE. The entire NAS domain is covered by 2400 x 1920 horizontal grid boxes. It spans 80 vertical levels up to 38.5 km.

Variational data assimilation is performed at a coarser resolution of 6.6 km and covered by 800 x 640 grid boxes.

It needs to be added that this specific domain configuration had been selected at the beginning of the developmental process and the results presented in this report use the domain and grid described above. However, it is expected that the operational version of NAS, as envisaged at the moment of typing this report, will be using the ACCESS-A grid. ACCESS-A is a currently developed regional NWP system covering Australia at a grid spacing of 1.5km.

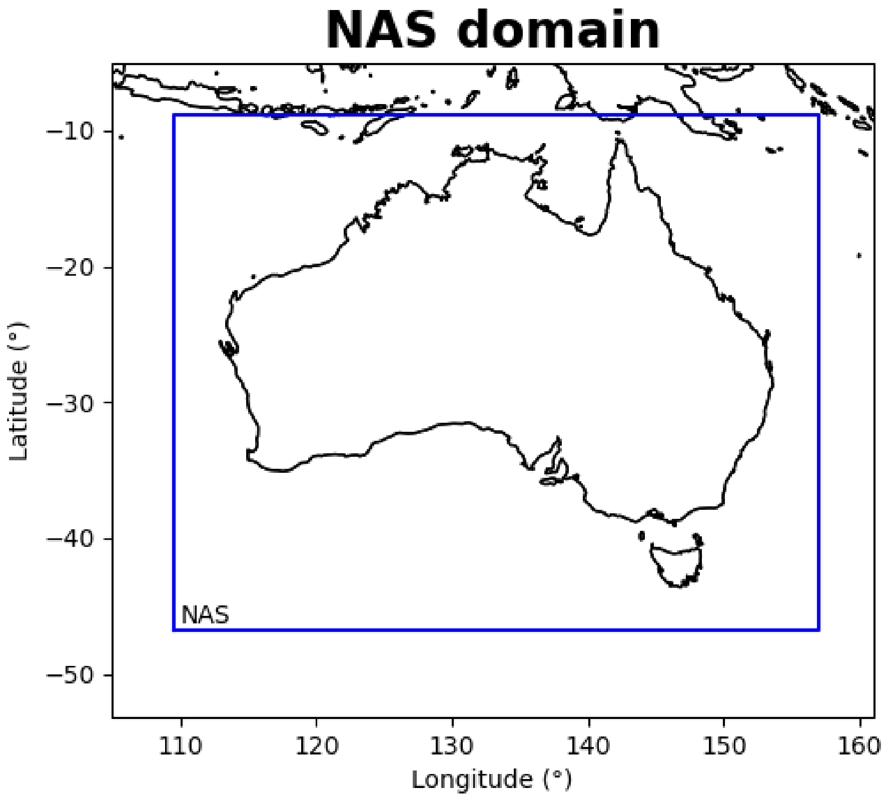


Figure 2. NAS domain.

2.1.2. Nesting and boundary conditions

NAS is a limited area model one-way nested in a global model. The global model ACCESS-G3 provides the lateral boundary conditions as well as the initial condition for any cold start. Due to the cycling of the global model the boundary conditions are updated every 6 hours.

2.1.3. Ancillaries

In addition to initialisation and boundary conditions, the atmospheric model requires supplementary information such as vegetation information, ozone fields, sea salt, black carbon, dust, organic carbon fossil fuel, sulphate, biomass burning, ozone, biogenic aerosol and others provided via ancillary files.

2.1.4. Sea Surface Temperature

Sea Surface Temperature comes from Bureau's operational, high-resolution Regional Australian Multi-Sensor Sea surface temperature Analysis (RAMSSA) described in Beggs H. et al. (2011). After reconfiguration to the NAS domain, it is used to replace the sea surface temperature in the initial conditions of the 1000 UTC run in the NAS system.

2.1.5. System components

NAS is an NWP-like system and essentially is constituted of the same elements as other NWP suites, namely ACCESS-G (see NOC Operations Bulletin Number 122) and ACCESS-C. Its backbone is a sequence of hourly data assimilation tasks and atmospheric forward modelling tasks. The variational data assimilation tasks are performed using var-2021.03.0 and the processing of observations required by data assimilation uses ops-2021.03.0. NAS' atmospheric forward model is Unified Model version 11.9 in conjunction with RA2M physics parameterisation.

For the land surface data assimilation NAS uses surf-2021.03.0. It is underpinned with the Joint UK Land Surface Model (JULES), version 5.8. It is worth mentioning that SURF is also used in the preparation of RAMSSA SST fields for NAS.

2.2. Atmospheric data assimilation in NAS

Atmospheric analyses in NAS are generated using variational data assimilation. Variational data assimilation is a well-established methodology in geoscience. In fact, the Bureau of Meteorology also uses 4D-Var (see Rawlins et al., 2007) to generate initial conditions for the weather forecast in limited area modelling systems, ACCESS-C. In both systems, and more broadly in Numerical Weather Prediction (NWP) systems in general, a new analysis is generated by data assimilation. At the Bureau, it is incremental variational data assimilation and the resulting analysis increment is added to the past forecast, generating an initial condition for the next forecast. This repeated analysis-forecast cycling across fixed windows provides the continually updated initial conditions.

2.2.1. Background error covariances

A crucial component of variational data assimilation is the background error covariance matrix. It defines the weight assigned to the prior (forecast) estimate atmospheric state. It also enables spread of information between observations. The background error covariance matrix as implemented within the Unified Modelling System is a set of transforms making the definition and application of the background error covariance matrix tractable. The background error covariance matrix is defined in three steps, the parameter transform, which maps the model variables into the control space. For the current definition of the background error covariance matrix these are: stream function, velocity potential, ageostrophic pressure and moisture control variable. These are further decorrelated using geometrical transforms, vertical and horizontal. The vertical transform computes vertical background error covariance matrix averaged over the whole domain which is then diagonalised to obtain a set of uncorrelated vertical modes. To further decorrelate the background error covariance matrix, the horizontal modes are defined with a Fourier transform. A new option has recently been added to the generation of the background error covariances, performing the horizontal transform first, followed by a vertical transform on the coefficients of the Fourier decomposition. This allows the analysis to better update smaller scales (see Wlasak and Cullen, 2014), an option that is particularly beneficial to a system, such as NAS.

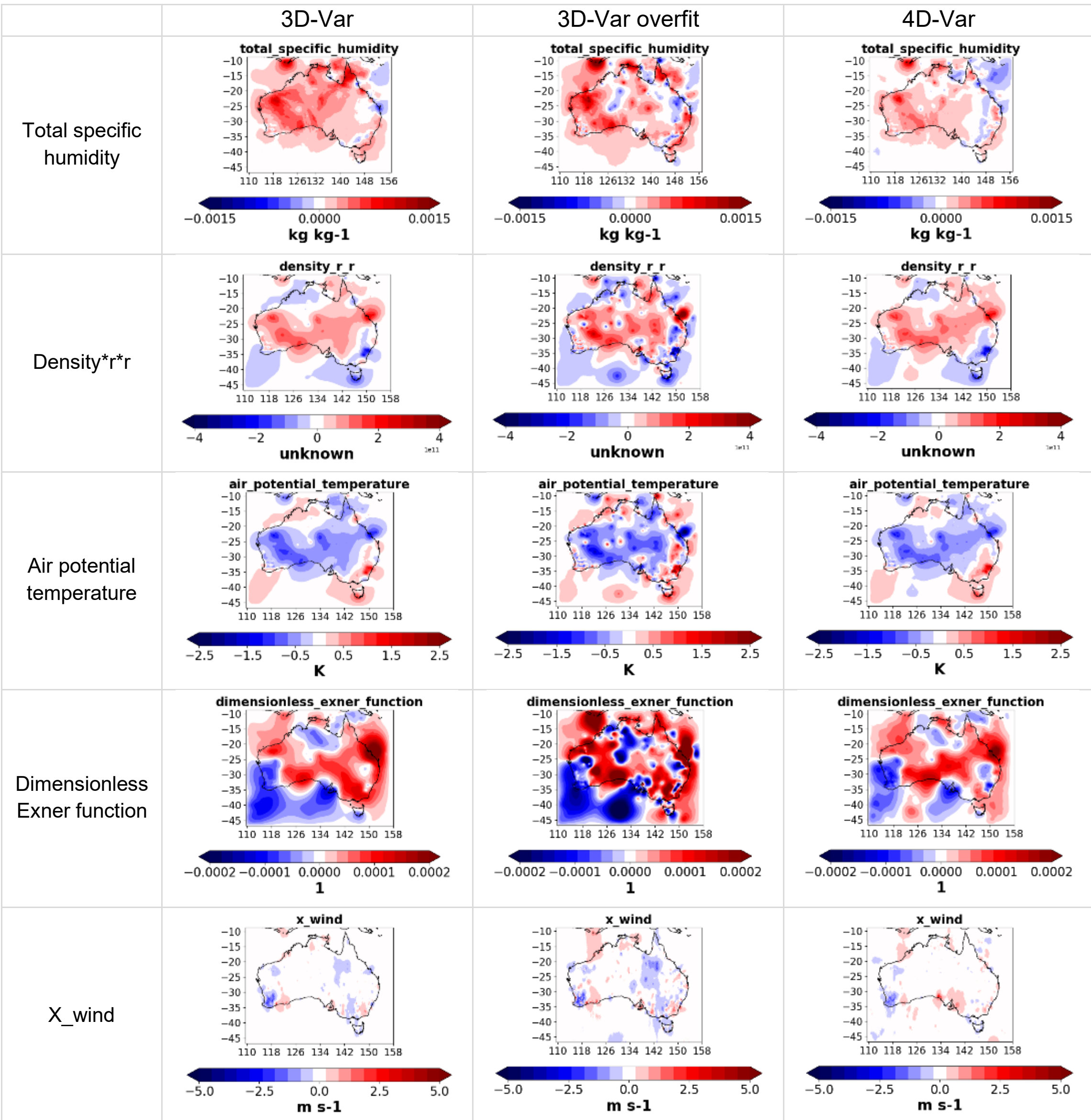
The initial version of the National Analysis System made use of a background error covariance matrix generated for earlier runs of ACCESS-C3 over the Sydney domain. Subsequently, and in view of implementation of assimilation of GeoCloud (retrievals of cloud top pressure and cloud amount from Himawari-8) as well as correlated wind errors and radar reflectivities (see section 2.2.4.2), we upgraded the background error covariance matrix. In addition to this 3D background error covariance matrix, there is also a one-dimensional background error covariance matrix used in the system, in the quality control of satellite radiance observations. There were indications from the UM Partners that using consistent background error covariance matrices, namely the 1D defined for the observation processing and 3D used in data assimilation, could have a positive impact on system performance. The consistent covariances means that both were generated using the same training set. In addition, using a training set from a downscaler from a global model rather than using a training set generated in a LAM, would lead to better results. Consequently, we performed generation of a new training set, specifically for the NAS domain. The runs were initialised by our global model and we used differences between 9h and 3h forecasts to estimate forecast error following Parrish and Derber (1992). Availability of this training set enabled generation of the background error covariances in the transform order, which favours small scales. We generated a training set for three months of summer and three months of winter and then randomly selected 100 pairs valid at the same time.

2.2.2. Modifications of the variational analysis (3D-Var and 3D-Var overfit)

The quality of variational analyses depends on the assimilated observations but also on the performance of the minimisation algorithm. 4D-Var being a costly procedure, we use a limited number of iterations in its minimisation in addition to convergence criteria. 3D-Var requires no model integration and hence is cheaper, which allows for a larger number of iterations. In the second 3D-Var analysis, that is initialised by the output from the first 3D-Var analysis and overfits the observations, we eliminated the convergence criteria but require a fixed number of iterations to be performed. Combined with the reduced observation error, this ensures bringing the resulting analysis closer to the observations. This is possible because the sequential 3D-Var analyses run relatively quickly and many iterations of the minimisation routine are allowed within the available operational time frame. In this first NAS version, only conventional observations contribute to the second 3D-Var analysis with a reduced observation error; the observation error has been reduced for surface, aircraft and sonde observations by a factor of 10 across the whole domain.

2.2.3. Analysis increments

The performance of the three analyses implemented in NAS had been assessed first. A comparison between the 3D-Var, 3D-Var overfit (main NAS product) and 4D-Var analyses is presented in Figure 3. The 3D-Var and 4D-Var analysis increments look similar, the main difference being the presence of increments over the ocean in the latter as the 4D-Var runs at a later time which allows for satellite radiances to arrive.



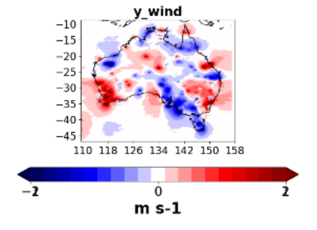
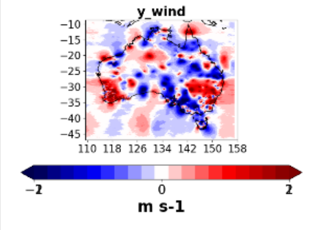
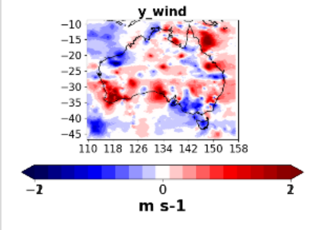
Y_wind			

Figure 3. Three analysis increments for selected fields. The name of the fields is presented in the first column, 3D-Var analysis increment is presented in the second column, 3D-Var overfit (main NAS product) in the third one. The fourth column shows 4D-Var analysis increments.

Secondly, a comparison between the two 3D-Var analyses indicates that an overfit to observations (middle column in Figure 3) results in a less smooth analysis field where the location of certain surface observations can be identified. This contrasts with a smooth analysis obtained in the 4D-Var, which is used for the initial conditions of the cycling forecast. In fact, the 3D-Var analysis ensuring a tighter fit to observations is characterised by strong gradients and would not be an appropriate initial condition for a model integration. At the same time it is representative of the current meteorological situation at the individual station locations, to give forecasters a representation of the atmosphere that agrees closely with observations.

2.2.4. Observations

The National Analysis System makes use of the observations traditionally assimilated into NWP models. These comprise surface, upper air, radar and satellite observations. Specifically, NAS assimilates in-situ observations from synoptic stations, aircraft data, sondes, wind profilers, satellite radiances (from ATMS, ATOVS, CrIS, IASI, AHI clear), atmospheric motion vectors, scatterometer winds, Doppler radial winds and radar reflectivities, and ground-based GNSS. Novelty resides in assimilation of GeoCloud, which uses cloud top pressure and effective cloud amount (the product of cloud fraction and cloud top emissivity). Assimilated into developmental NAS, GeoCloud comes from the infra-red imager on Himawari-8 and 9. More specific details for assimilation of similar cloud information from SEVIRI can be found in Renshaw and Francis, (2011).

The same observation types can in principle be assimilated for all three analyses. The difference is the cut-off time: the latest time at which observations must arrive in the Bureau in order to be assimilated. Consequently, 4D-Var assimilates more observations than its 3D predecessors and so, for example, ground-based GNSS data, are only assimilated in 4D-Var due to their latency.

The description of assimilated observations follows Operations Bulletin Number 137 (2021).

2.2.4.1. Surface-based conventional observations

NAS assimilates a broad range of conventional surface observations of temperature, pressure, humidity and wind velocity. Surface observations comprise land synoptic observations, METARs (MÉTéorologique Aviation Régulière) from AWS (Automatic Weather Stations), atmospheric component of buoy observations as well as observations from ships. An example of data coverage is shown in Figure 4.

Temperature, pressure, humidity, wind data are ingested (where available) from automatic and manual weather stations, buoys, and ships. These are standard observations used in all NWP systems.

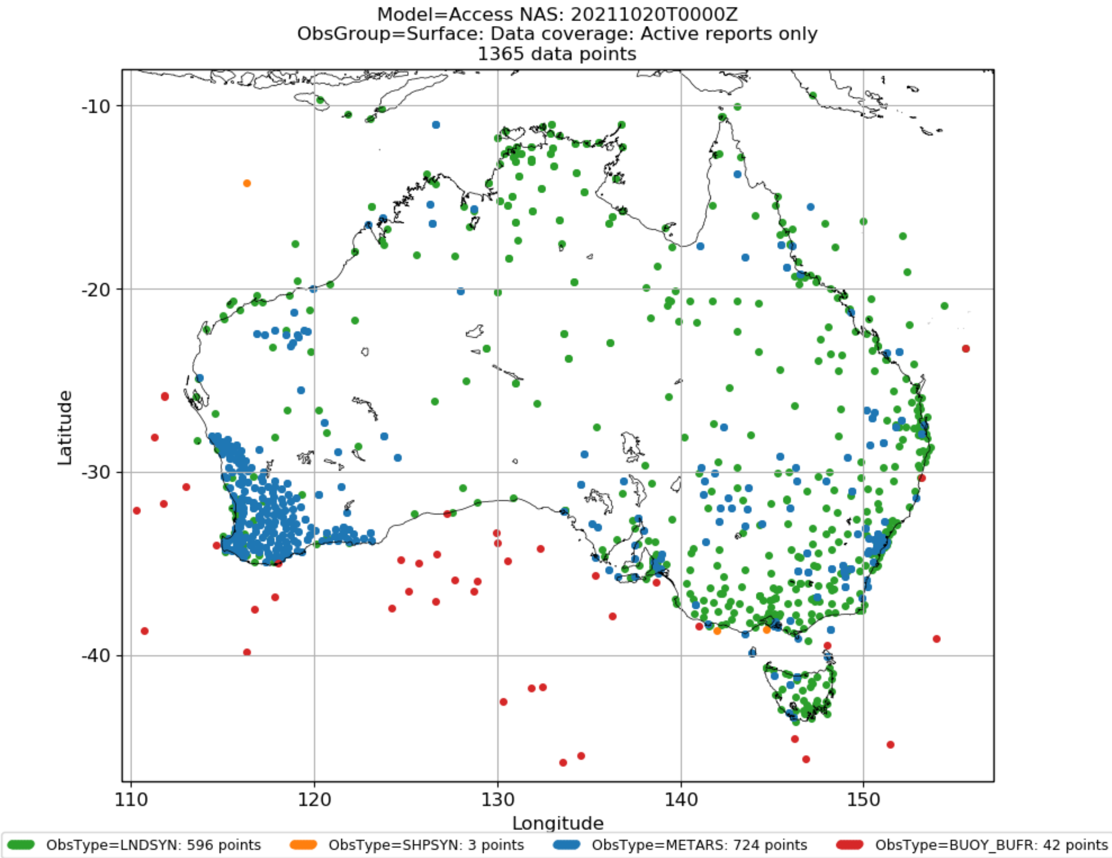


Figure 4. Surface-based conventional observations.

2.2.4.2. Surface-based remotely sensed observations

Two types of observation fall under this umbrella, namely radar observations and ground-based GNSS observations.

Doppler weather radars are a valuable source of spatially and temporally high-resolution observations of local winds, convective systems and precipitation. Two types of Doppler radar observations are assimilated in NAS. Firstly, Doppler radial winds, which are extracted from both precipitation and insect echoes (see for example, Rennie et al., 2020). The assimilation extracts horizontal wind components from the radial wind observations via an appropriate observation operator and its adjoint. Another challenge resides in the high resolution of observations which has been traditionally addressed by thinning or combination into super-observations. However, recently, Simonin et al. (2019), introduced a methodology of accounting for correlated observation error statistics allowing assimilation of a larger number of radial winds into models. This methodology, as reported in Rennie (2023), is also utilised in NAS. Secondly, direct assimilation of radar reflectivities has been implemented in NAS, (see also Rennie, 2023). This eliminates a necessity for latent heat nudging in the vicinity of radars, although this does leave large gaps not covered by radar reflectivity information. Another important feature to bear in mind is that radar observations depend on weather and consequently observation coverage is highly variable. An example of coverage is shown in Figure 5.

The second type of surface-based remotely sensed data are ground-based GNSS (Global Navigation Satellite System) ZTDs (Zenith Total Delays). In fact, the path in the zenith direction, from a satellite to a ground-based receiver is known as the Zenith Total Delay (see, for example Bennitt et al., 2017). It is sensitive to the moisture content in the atmosphere and is hence used to estimate total precipitable water (see for example Le Marshall et al., 2020). An example of data coverage can be found in Figure 5.

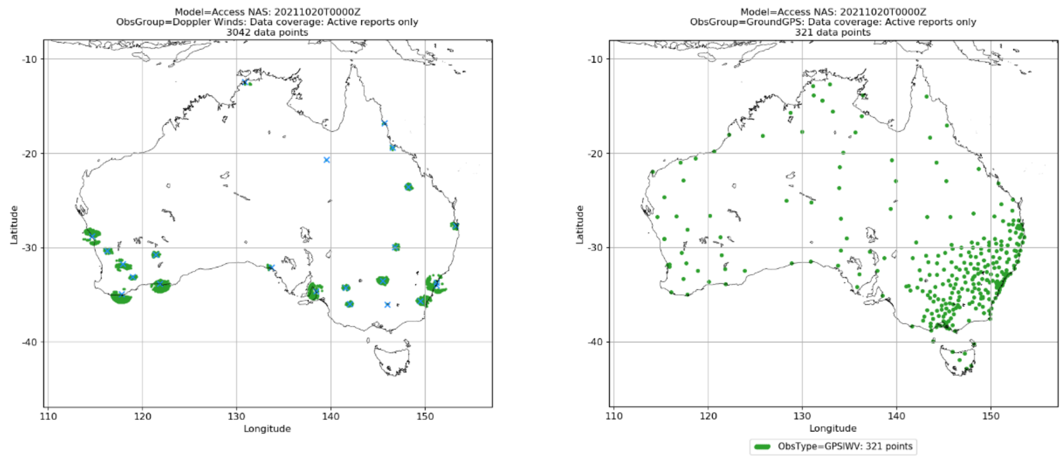


Figure 5. Examples of data coverage for Doppler weather radars on the left hand side and for ground-based GNSS stations on the right hand side.

2.2.4.3. Upper air conventional observations

Upper air observations are taken from aircraft, balloon ascents and wind profilers. AMDAR (Aircraft Meteorological Data Reports) are automated weather reports from commercial aircraft, particularly from ascent and descent stages of flight.

2.2.4.4. Satellite observations

NAS directly assimilates radiance observations from infrared sounders IASI and CrIS and from microwave sounders ATMS and AMSU. Examples of data coverages are presented in Figure 6. In addition, NAS also assimilates AHI (Advanced Himawari Imager) clear sky radiances.

Moreover, Atmospheric Motion Vectors (AMVs) are generated based on Himawari-8 imagery. In order to obtain AMVs, visible and infrared water vapour channels are used. Observation coverage for those is presented in Figure 7.

Another observation type arising from satellite detection are scatterometer winds, currently retrieved from MetOp-B and MetOp-C.

Variational bias correction (VarBC) is used to correct the radiance biases and further details on this methodology can be found in Dee (2004) and Auligné et al. (2007). In NAS, the VarBC coefficients are taken from the global system, which updates the bias corrections every 6-hour cycle.

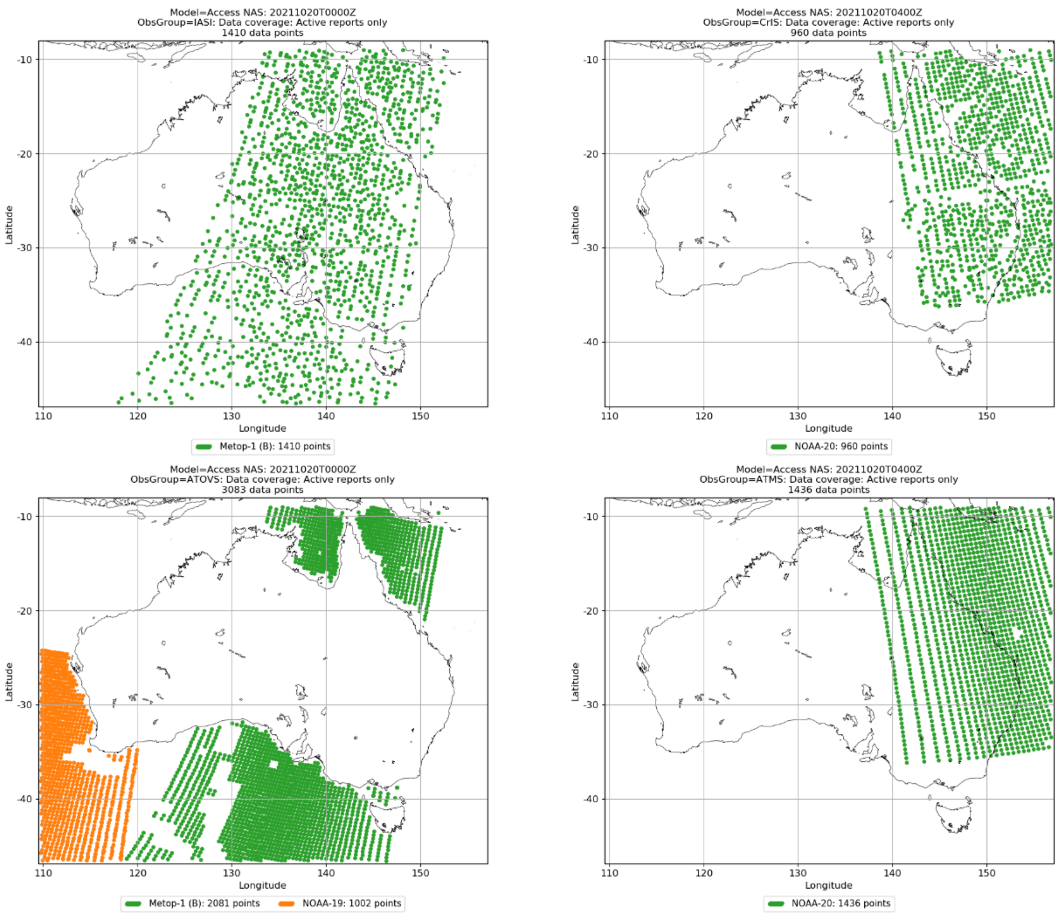


Figure 6. Examples of satellite data coverages. IASI (top left), CRiS (top-right), ATOVS (AMSU) bottom left and ATMS bottom right. Note different times, namely 20211020T0000Z for IASI and ATOVS and 20211020T0400Z for CRiS and ATMS.

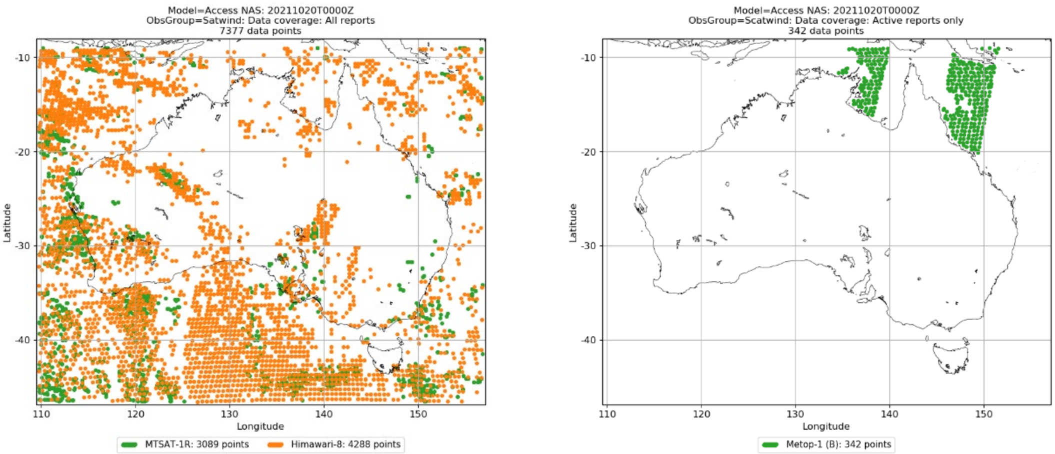


Figure 7. Observation coverage for AMVs on the left hand side and scatterometer winds on the right hand side. Both plots for 20211020T0000Z.

2.3. Land surface data assimilation

The aim of land surface data assimilation is to produce high quality fluxes from the land to the atmosphere, and so soil temperature and soil moisture are regularly updated. The importance of correct land surface information cannot be overstated. This has been exhibited in Benson et al. (2022), which describes an analysis that formed part of the assessment of the case study encompassing the New South Wales floods which occurred in 2021.

2.3.1. EKF and observations

The land surface data assimilation methodology employs a Simplified Extended Kalman Filter-based scheme (SEKF, Dharssi et al., 2012; Gómez et al., 2020) that assimilates bias corrected measurements of surface soil wetness from satellite-based microwave sensors and pseudo-observations of 2 m temperature and humidity. The pseudo-observation innovations are analysis increments produced by the 4D-Var atmosphere analysis in the same cycle. The improved land surface feeds into the background of the subsequent cycle's 3D-Var analyses.

3. Performance assessment

Being focussed on the analysis of atmosphere, it is difficult to assess the performance of NAS due to lack of tools and a lack of independent observations. Firstly, the existing verification methodology generally targets system performance based on forecast verification. NAS produces forecasts of only limited temporal horizon. Therefore, these are not appropriate for the assessment of NAS, as the forecast errors are still highly correlated with the analysis errors. Consequently, the system was assessed with subjective verification using a series of case studies. These comprise different meteorological phenomena, ranging from a tropical cyclone, through bush fires, dust storm, floods to hail and thunderstorms. The performance of these have been assessed in the meteorological tools used by forecasters. Secondly, the system was subject to objective verification performed using cross validation, where a small subset of observations was withheld form the analysis and then used for verification.

3.1. Objective verification with cross-validation

3.1.1. Description

Cross-validation consists of withholding a subset of observations from assimilation and using them to assess the quality of the analysis. For that purpose, a series of 8 experiments were run. Each experiment covers a period of 4 days, 8 experiments summing up to a period of one month. Each hourly cycle withholds observations from 15 randomly selected surface stations, which amount to, roughly, 2.5% of the total surface observations. This number is a compromise between limiting the number of withheld observations in order not to significantly degrade the quality of the analysis and being large enough for the experiments to produce significant verification statistics. It is subsequently assessed with the root mean squared error calculated for the observations which were withheld from the analysis.

3.1.2. Results

In Table 1 we present the results of cross-validation experiments for selected meteorological fields. In general, the NAS product, 3D-Var overfit, leads to the smallest values of RMSE, except for wind direction and zonal wind.

Table 1. RMSE for selected meteorological fields computed against observations withheld from data assimilation for NAS analyses.

Field	3D-Var	3D-Var overfit NAS product	4D-Var
Screen temperature (ºC)	1.39	1.31	1.42
Dew point (ºC)	2.10	1.99	2.16
Relative humidity (%)	8.34	8.16	8.40
Station pressure (hPa)	0.58	0.53	0.58
Wind direction (deg)	47.27	46.78	46.50
Wind speed (m/s)	1.94	1.89	1.97
Meridional wind (m/s)	1.89	1.84	1.87
Zonal wind (m/s)	1.89	1.87	1.80

3.2. Subjective verification via case study analysis

In addition, NAS has been subject to subjective assessment via case study analysis. The main component of the analysis has been performed by a recently retired operational forecaster. In this analysis mean sea level pressure charts have been looked at to understand the general atmospheric circulation underpinning the meteorological phenomena. Subsequently, the attention was focussed on a comparison of modelled surface temperature, dewpoint temperature and surface winds with observations as well as Visual Weather OnScreen Analysis. The detailed analysis has been reported in Benson et al., 2022. The general conclusions point to good NAS performance in terms of simulating winds. The deficiencies in surface temperature have been identified to come from lack of up-to-date soil moisture fields. This was remediated by adding land surface data assimilation component to NAS, which updates soil moisture field.

3.3. METplus verification of radiance simulator

These two methodologies were used to assess performance of the entire system. Further, we performed a targeted verification of GeoCloud assimilation in NAS. For that purpose we ran three instances of NAS for a period of floods in Queensland in early 2022. These consisted of a control, which did not assimilate Geocloud, a trial in which GeoCloud assimilation was performed in a sub-optimal setting and an experiment in which GeoCloud was assimilated in an optimal setting. The suboptimal configuration used older background error covariances instead of 1D and 3D background error covariances estimated from randomly selected forecast pairs (see section 2.2.1).

For GeoCloud assimilation, an assessment against surface observations is not necessarily the best option. Therefore, we decided to assess its impact by a comparison of model outputs with the observations from Himawari-8. In fact, a new addition was supplied to NAS in the form of a radiance simulator. It produces reflectances and brightness temperatures for Himawari-8 channels. The gridded fields produced by NAS and observations are compared using METplus, see Opatz et al., (2023).

Firstly, dimensioned statistics are compared for three instances of NAS. These were produced for all 10 brightness temperature channels. Mean absolute error and root mean squared error indicate that the performance of the optimized experiment run outperformed the other two instances of the NAS suite. We illustrate these results in Figure 8. It is visible in the plots that the optimized configuration produces lowest values of error among three simulations. The illustration is for one of the atmospheric window channels but the results are similar for the remaining infra-red channels.

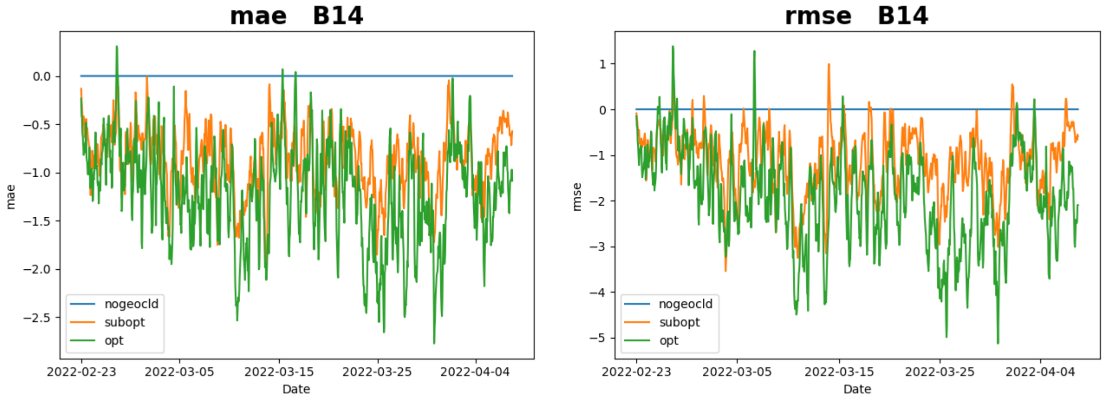
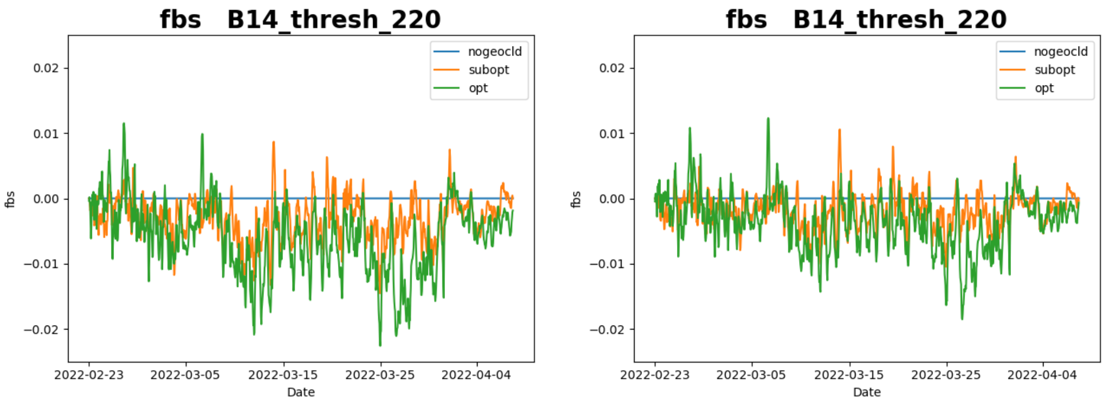


Figure 8. Assessment of GeoCloud assimilation with point-wise statistics. Time series of mean absolute error and root mean squared error illustrated for one of the atmospheric window channels, number 14 centred on 11.2 micrometres. Plots are normalised.

Secondly, neighbourhood statistics were computed. These consist of fractions Brier score (FBS) and fractions skill score (FSS). The statistics involve an arbitrary definition of a cut-off threshold re-defining the analysed quantities into a binary field. They are not performed point-wise but within a neighbourhood of a certain size around each grid point. They were calculated for a single channel only. In this case, however, different threshold values and neighbourhood sizes were tried. The statistics illustrate that the trials performed reasonably well. FBS, having an optimal value of 0, decreases with an increasing neighbourhood size. This is the expected performance since the increased neighbourhood size relaxes on the location error and produces more favourable outputs for large neighbourhoods (see Figure 9).



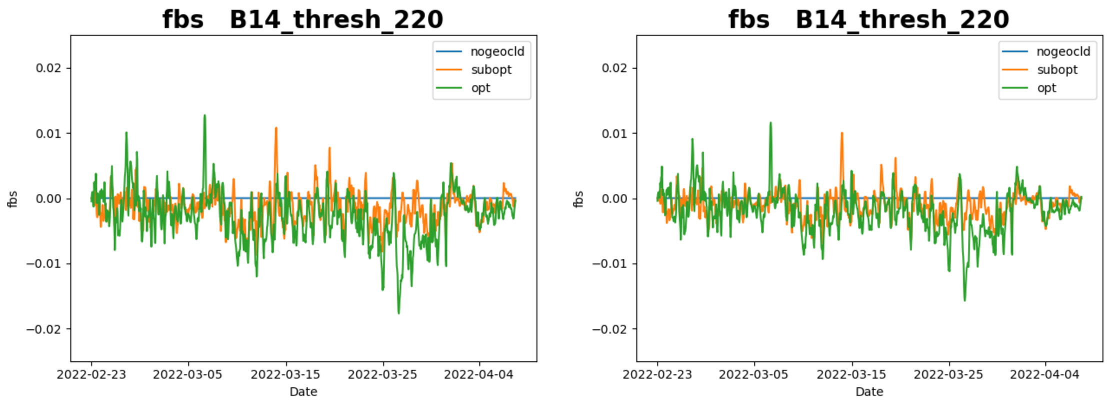


Figure 9. Time series of Fractions Brier Score for three GeoCloud experiments and different neighbourhood sizes (1, 25, 51, 101, left to right, top to bottom).

We also observe that FBS shows larger values for higher cut-off temperatures (Figure 10). This is also expected as the higher temperatures define larger cloud fields and, for a given neighbourhood size, these have larger probability of overlapping with the observations.

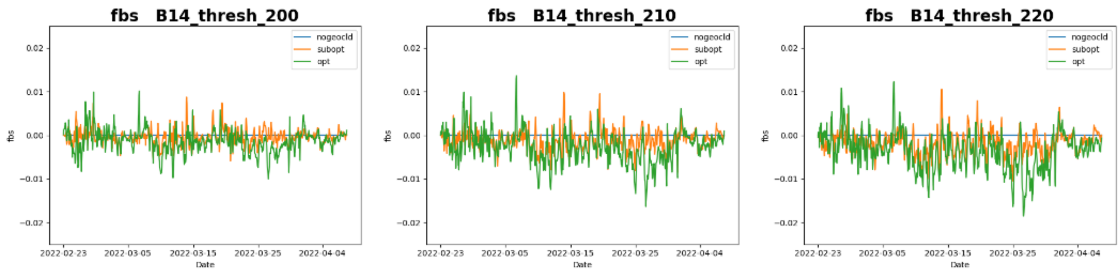


Figure 10. Time series of Fractions Brier Score for different cut-off values of brightness temperature, 200, 210 and 220K. Neighbourhood size is 25.

FSS behaves similarly, however, its optimal value being one, it grows with the neighbourhood size (Figure 11). There is not a clear difference between sub-optimal and optimal configuration experiments for FSS. For the analysed cut-off temperatures there seems no significant difference between different values of the cut-off temperature (Figure 12).

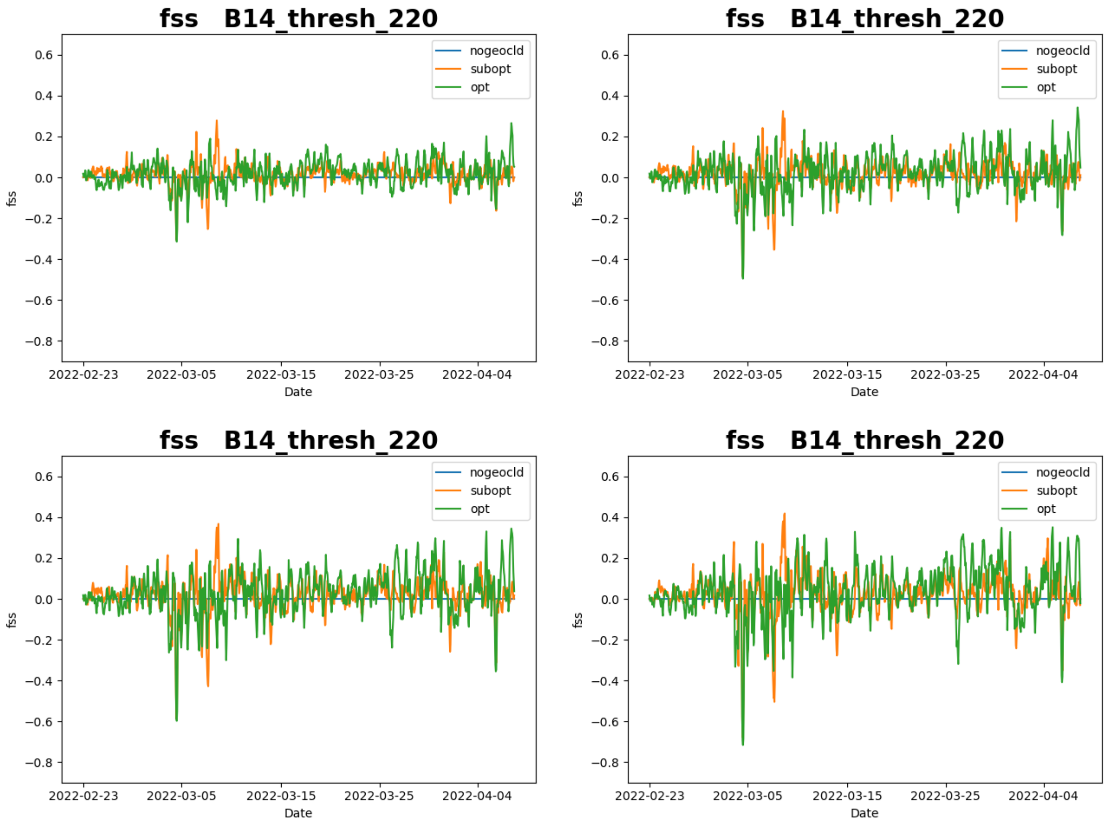


Figure 11. Time series of Fractions Skill Score for three GeoCloud experiments and different neighbourhood sizes (1, 25, 51, 101, left to right, top to bottom).

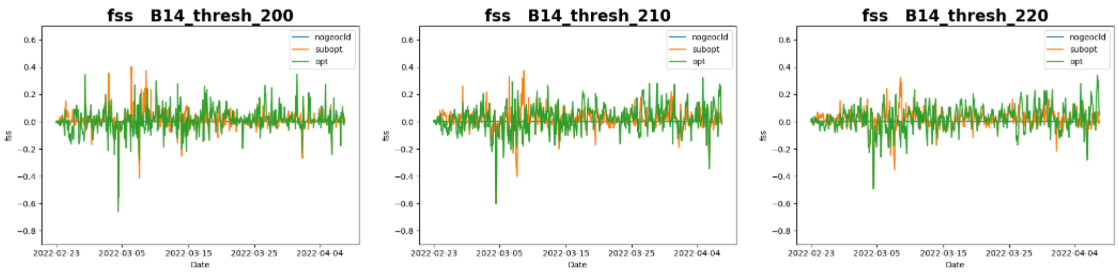


Figure 12. Time series of Fractions Skill Score for three GeoCloud experiments and different cut-of temperature values, 200, 210 and 220K. Neighbourhood size is 25.

Overall, these results indicate that there is added value in assimilating GeoCloud information into the system. Moreover, the updated background error covariances used in the optimal setting illustrate their added value with respect to the original formulation of background error covariances.

4. Comparison with MSAS

Regarding mesoscale analysis systems, Bureau currently operates a Mesoscale Surface Analysis System (MSAS). The results of cross validation experiments have been compared between NAS (the overfit product) and MSAS. The MSAS results come from a visual inspection of plots in Glowacki et al. (2012) and produce similar results (Table 2). MSAS has the lower RMSE except for screen temperature, though the values are in a similar range.

Table 2. Comparison of the RMSEs from cross-validation experiments between NAS and MSAS.

Field	NAS	MSAS
Screen temperature (ºC)	**1.31**	1.5
Dew point temperature (ºC)	1.99	**1.7**
Zonal wind (m/s)	1.87	**1.7**
Meridional wind (m/s)	1.84	**1.7**

5. Conclusions

At present, NAS has reached a certain level of maturity. The initial concept of the system has been implemented. The use of two additional 3D-Var analyses in NAS has been assessed. Firstly, the analysis increments produced by three different analyses provide increments of the expected shape, i.e. a strong imprint of observations for 3D-Var overfit. Secondly, the performance of NAS has been tested via objective analysis performed using cross-validation. The root mean squared error calculated for the withheld observations generally gives slightly smaller values for the 3D-Var overfit than for the other two analyses. A crude comparison to MSAS gives smaller values of RMSE for surface temperature but slightly higher values for dew point temperature and winds. However, these are comparable, and given other advantages of NAS, namely its 3D aspect, the first version of NAS is considered to be ready for operationalisation.

Acknowledgements

The Authors, all in Research Programme of the Science and Innovation Group (SIG) at the Bureau, are grateful to the Colleagues in Research to Operation Programme, namely Tan Le, Ilia Bermous, Tom Coleman, Christopher Down, Wenming Lu and Andrew Kelly for their support during the system development and assessment phase. The Authors would like to extend their gratitude to the NAS Project Manager, Jill Finch, for coordinating the efforts of all involved.

References

Auligné, T., McNally, A.,P. and Dee, D.,P. (2007): Adaptive bias correction for satellite data in a numerical weather prediction system, *Quarterly Journal of the Royal Meteorological Society*, **133**(624), 631-642.

Beggs H., Zhong A., Warren G., Alves O., Brassington G. and Pugh, T. (2011). RAMSSA – An Operational, High-Resolution, Multi-Sensor Sea Surface Temperature Analysis over the Australian Region, *Australian Meteorological and Oceanographic Journal*, 61, 1-22, 2011

Bennitt, G.V., Johnson, H.,R., Weston, P.P., Jones, J. and Pottiaux, E. (2017) An assessment of ground-based GNSS Zenith Total Delay observation errors and their correlations using the Met Office UKV model, *Quarterly Journal of the Royal Meteorological Society*, **143**(707), 2436-2447.

Benson, S., Krysta, M. and Bannister, T. (2022) Evaluating the Advantages and Challenges of the National Analysis System, Bureau Research Report, **70**, pp 1-52

Dee, D.,P. (2004): Variational bias correction of radiance data in the ECMWF system, Conference Paper, ECMWF, pp. 97-112.

De Pondeca, M.S.F.V., Manikin, G.S., Dimego, G., Benjamin, S.G., Parish, D.F., Purser, R.J., Wu, W., Horel, J.D., Myrick, D.T., Lin, Y., Aune, R.M., Keyser, D., Colman, B., Mann G. and Vavra, J. (2011). The Real-Time Mesoscale Analysis to NOAA's National Centers for Environmental Prediction: Current Status and Development, *Weather and Forecasting*, 26(5), 593-612

Dharssi, I., Steinle, P. J., and Candy, B. (2012). Towards a Kalman Filter based land surface data assimilation scheme for ACCESS. In CAWCR Technical Report No. 054 (Issue 54). http://www.bom.gov.au/research/publications/cawcrreports/CTR_054.pdf

Glowacki, T. J., Xiao, Y. and Steinle, P. (2012). Mesoscale Surface Analysis System for the Australian Domain: Design, Issues, Development Status, and System Validation, Weather and Forecasting, 27(1), 141-157

Gómez, B., Charlton-Pérez, C. L., Lewis, H., & Candy, B. (2020). The Met Office Operational Soil Moisture Analysis System. *Remote Sensing*, 12(22), 3691. https://doi.org/10.3390/rs12223691

Le Marshall, J., Norman, R., Howard, D., Rennie, S. Moore, M., Kaplon, J., Xiao, Y., Zhang, K., Wang, C., Cate, A., Lehmann, P., Wang, X., Steinle, P., Tingwell, C., Le, T., Rohm, W., Ren, D. (2020). Using global navigation satellite system data for real-time moisture analysis and forecasting over Australian region I. The system, *Journal of Southern Hemisphere Earth Systems Science* 69(1), 161-171

NOC Operations Bulletin Number 122, APS3 upgrade of the ACCESS-G/GE Numerical Weather Prediction system, 20 August 2019

Opatz, J., J. Halley Gotway, T. Jensen, J. Vigh, M. Row, C. Kalb, H. Fisher, L. Goodrich, D. Adriaansen, M. Win-Gildenmeister, G. McCabe, J. Prestopnik, J. Frimel, L. Blank, T. Arbetter, 2023: The METplus Version 5.1.0 User's Guide. Developmental Testbed Center. Available at: https://github.com/dtcenter/METplus/releases.

Operations Bulletin Number 137, APS3 ACCESS C Upgrade, August 2021

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120, 1747–1763, https://doi.org/10.1175/1520- 0493(1992)120,1747:TNMCSS.2.0.CO;2.

Raby, J.W., Cai, H., Dawson, L. and Dumais, R. (2020). An Evaluation of the Unrestricted Mesoscale Analysis [Slides]. AD1115820.pdf (dtic.mil)

Rawlins, F., Ballard, S.P., Bovis, K.,J., Clayton,A.M., Li, D., Inverarity, G.W., Lorenc, A.C. and Payne T.J., The Met Office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, **137**, 1963-1974, (2011).

Rennie, S. (2023) Radar data assimilation upgrades for APS4, Bureau Research Report, **78**, pp 1-80

Rennie S., Cooper S., Steinle P., Dietachmayer G., Krysta M., Franklin C., Bridge C., Marshall M., Xiao Y. and Sgarbossa, D., ACCESS-C: Australian Convective-Scale NWP Hourly 4D-Var Data Assimilation, *Weather and Forecasting*, vol 37, issue7, pp. 1287-1303 (2022)

Rennie S., Rikus, L., Eizenberg, N., Steinle, P., Krysta, M., Impact of Doppler Radar Wind Observations on Australian High-Resolution Numerical Weather Prediction, Weather and Forecasting, vol **35**, issue 2, pp. 309-324 (2020)

Renshaw, R. and Francis, P.N., Variational assimilation of cloud fraction in the operational Met Office Unified Model, *Quarterly Journal of the Royal Meteorological Society*, **133**, 347-362 (2007).

Simonin, D., Waller, J.A., Ballard, S.P., Dance, S.L. and Nichols, N.K., A pragmatic strategy for implementing spatially correlated observation errors in an operational system: an application to Doppler radial winds, *Quarterly Journal of the Royal Meteorological Society*, **145**(723), 2772-2790 (2019)

Weygandt, S.S., Alexander, C.R., Ge, G., Hu, M., Ladwig, T.T., Hartsough, C., Carley, J.R., Zhao G., Pondeca, M. and Yang, R (2019). Evaluation of a Prototype Version of the 3D-Real-Time Mesoscale Analysis (3D-RTMA) for Situational Awareness and Nowcast Applications [abstract]. In: AMS 99th Annual Meeting, Phoenix, 6-10 January 2019. Abstract: Evaluation of a Prototype Version of the 3D-Real-Time Mesoscale Analysis (3D-RTMA) for Situational Awareness and Nowcast Applications (99th American Meteorological Society Annual Meeting) (confex.com)

Wlasak M.A. and Cullen M.J.P., Modelling static 3-D spatial background error covariances – the effect of vertical and horizontal transform order, *Advances in Science and Research*, 11, pp. 63-67, (2014) Yin, J., Han, W., Gao, Z. and Chen, H., Assimilation of Doppler radar radial wind data in the GRAPES mesoscale model with observation error covariances tuning, *Quarterly Journal of the Royal Meteorological Society*, **147**(737), 2087-2102, (2021)