Mesoscale Surface Analysis System for the Australian Domain: Design Issues, Development Status, and System Validation

TOMASZ J. GLOWACKI, YI XIAO, AND PETER STEINLE
Centre for Australian Weather and Climate Research, Bureau of Meteorology, Melbourne, Victoria, Australia

(Manuscript received 16 December 2010, in final form 12 July 2011)

ABSTRACT

An operational surface analysis system for the continent of Australia is presented. The system is specifically designed to mitigate problems that arise when analyzing surface data with a highly inhomogeneous distribution. Hourly analyses of atmospheric pressure at mean sea level, potential temperature, 2-m dewpoint temperature, and 10-m wind components are generated on a \( \frac{1}{4} \) km grid. The system employs a statistical interpolation technique using observations of pressure, temperature, dewpoint, and wind data. The problem of data gaps in space and time is addressed by introducing pseudo-observations. For stations missing a report at analysis time, estimates are reconstructed by interpolating off-time reports. Underobserved areas in the network are identified from precalculated, gridded observation densities for each analysis time, which also yield weights to combine preliminary analysis and first-guess data into pseudo-observations. A regression-based pressure reduction technique, consistent with local reductions at observing sites and devised specifically for this system, is used for accurate and fast conversion of pressure and, indirectly, temperature variables within the system. Analysis accuracy is verified by withholding observations for specific periods. Analyzed fields are shown to be significantly more accurate than the current operational numerical model fields used as a first guess for the high-resolution surface analysis. The system design and analysis accuracies are also assessed within this context and compared with similar overseas developments.

1. Introduction

Over the last few years a high-resolution, high-frequency Mesoscale Surface Analysis System (MSAS) has been developed at the Australian Bureau of Meteorology (hereafter shortened to the Bureau), and became operational in September 2007. The aim of MSAS is to generate real-time, best-possible, gridded, high-resolution surface analyses of weather parameters. At present, analyses of mean sea level pressure (MSLP), 2-m temperature (T2), 2-m dewpoint temperature (D2), and 10-m wind components (U10, V10) are generated and archived at hourly intervals over a single domain covering Australia and the adjacent oceanic areas (Fig. 1). These analyses are designed to support the Gridded Operational Consensus Forecasts (GOCF) system and the Graphical Forecast Editor (GFE; Hume et al. 2009), as well as other applications including climate analysis. This paper describes the system and evaluates its accuracy.

Corresponding author address: Tomasz J. Glowacki, Centre for Australian Weather and Climate Research, Bureau of Meteorology, GPO Box 1289, Melbourne VIC 3001, Australia.
E-mail: t.glowacki@bom.gov.au

DOI: 10.1175/WAF-D-10-05063.1

© 2012 American Meteorological Society
systems are similar: the provision of affordable, timely, high-resolution, high-frequency analyses, with accuracies equal to or better than postprocessed analyses available from NWP systems. These systems support a range of applications: near-real-time needs of forecasters (e.g., GFE), surface data assimilation into NWP systems, verification of gridded model forecasts, fast quality control of surface observational data, climatological applications, and inexpensive generation of retrospective reanalyses (Horel and Colman 2005).

Particular implementations depend on the density of the observational data; the types of observed surface parameters; the availability of high-resolution, high-accuracy gridded NWP output; and the topographical complexity. Analyzed variables may additionally include accumulated rainfall, visibility, wind speed, cloudiness, snow precipitation, solar radiation, and other parameters that are available from observation networks. As is argued in Haiden et al. (2010), Glaahn and Im (2010), and elsewhere, a well-designed surface analysis system is able to generate more accurate analyses of surface conditions than are NWP outputs of similar resolution. Because they are constrained by model dynamics and physics in the data assimilation cycle, NWP forecasts have errors in the range up to +6 h that are commonly not significantly smaller than those at +12 or +24 h, when validated against observations. This is also a result of the sophisticated data assimilation methods using relationships among free-atmospheric variables, which are not readily applicable at the surface. Models with horizontal resolution fine enough to resolve the small-scale atmospheric conditions at specific locations are not yet available for the Australian MSAS domain, although 0.05° resolution forecast-only models covering a few selected subdomains have been implemented in a limited scope.

Various techniques have been used to analyze surface data. In situations when the observation density is low or nonuniform, the latter being a common circumstance over the domain of interest, using observations alone may not be sufficient to generate analyses of acceptable accuracy. Objective analysis schemes that blend in additional information derived from a NWP model first guess (usually a short-range forecast) are preferable for such applications. They better constrain the analysis in areas where data are sparse and also transfer a part of the NWP model dynamics and physics into the analysis. In contrast, for applications such as NWP intercomparison or verification studies it has been suggested that a model-independent analysis is preferable (Steinacker et al. 2000). Comparisons of available analysis techniques and of their relative advantages and disadvantages may be found in Koch et al. (2004), Lanciani and Salvati (2008), and Johnson and Schneider (2009), for example.

In the case of the Australian surface observation network, large areas are very poorly observed. Therefore, the analysis must rely on NWP output to deal with the underobserved areas. The applied technique should preferably also “ensure some measure of physical consistency with land-surface conditions, land-water contrasts, terrain elevation, as well as with the 3-d effects of realistic thermal stability, boundary-layer structure, and local circulations.” (Benjamin et al. 2007). For these reasons and to utilize some existing technologies, statistical interpolation (SI) (Gandin 1963; Lorenc 1981) using a NWP background was chosen. SI has also been used in the ERAMESAN and NOAA MSAS systems. To fully exploit its potential, SI requires realistic settings of observational and background errors, their correlations, quality control parameters, etc. This is more difficult in complex terrain, as it requires local adjustments of the background error correlation matrix. The scope of these adjustments is limited however by the amount of available observational data, as an attempt to enhance the level of the local analysis detail is usually in conflict with accurate spreading of information to data-sparse or data-void regions (e.g., Myrick et al. 2005; Tyndall et al. 2010).

The number of hourly surface observations available from the Australian observation network ranges from about 400 to 700 depending on the parameter and hour of the day (excluding scatterometer winds). These data come from field stations operated by the Bureau and by other authorities (water agencies, departments of agriculture, etc.), as well as private companies and individuals. Field observations are collected by the Bureau only from those networks and operators that comply with strict quality guidelines. The number of observations currently available to MSAS is not much less than what was available (600–1000 reports) over the contiguous United

![Fig. 1. MSAS domain and topography. Red circles indicate temperature observations at 0100 UTC 28 Sep 2009.](image-url)
States to the Mesoscale Analysis and Prediction System (MAPS) system (Miller and Benjamin 1992, henceforth MB92) in the early 1990s. For comparison, the number of aviation routine weather reports (METARs) or synoptic-type stations currently reporting within the contiguous United States is two orders of magnitude greater (Im et al. 2010). STMAS executes every 15 min using about 32 000 observations available within a 90-min time window (Koch et al. 2009), and the European network is also substantially denser than the Australian network. Another limitation for the Australian region in comparison with U.S. and European domains is the relatively coarse NWP fields underpinning the mesoscale analysis system. MSAS mitigates against the shortage of observations by the time interpolation of observations and by using pseudo-observations along with other modifications to the “standard” SI implementation, detailed in the following section. Similar issues with the limitations of SI with inhomogeneous coverage have been reported elsewhere (e.g., Koch et al. 2004, 2009; Horel and Colman 2005), but dealt with differently (or left unresolved), due to different operating environments. Other novel approaches to these issues have also been suggested, such as the “low-level temperature” analysis of Bica et al. (2007) for complex terrain.

An SI-based analysis scheme has been implemented that is designed to both handle the specific aspects of the Australian surface observation network, and be appropriate for operational use within the Bureau. Given the moderate amount of irregularly distributed observational data and the available supercomputing resources, it is more efficient to analyze observations over the whole Australian domain rather than over separate state-based domains. This approach, although computationally more expensive, reduces spurious noise in analyses, which can arise when the analysis is done over many small subvolumes (e.g., MB92, Koch et al. 2004)—one of the weaknesses of so-called boxed optimal interpolation (OI) schemes (Lorenc 1981).

As discussed in Koch et al. (2004), existing objective analysis approaches like successive correction, SI, or the hybrid schemes of Bratseth (1986) or Barnes (1964) are adversely affected by inhomogeneity in data distributions. For an inhomogeneous data distribution, maximizing the analyzed detail in data-dense areas comes at the cost of introducing noise in data-sparse or data-void areas, when using these analysis approaches.

The extent of the poorly observed areas in continental Australia is greater than in either Europe or the continental United States and the problem requires different treatments here. These limitations of SI are reduced by generating a more even coverage pattern using pseudo-observations in areas of low data density.

The MSAS analysis procedure employs two SI passes through the data. In the first pass, a gridded data density field is computed and pseudo-observations are generated. The proper analysis is done in the second pass, which uses shorter length scales and all reported observations, MSAS time-interpolated observations, and pseudo-observations. In contrast to other schemes that analyze data in multiple passes, here the impacts of the first pass are limited to areas of low data density, and the first guess is the same for both passes.

The accuracy of MSAS analyses is assessed by withholding observations, the only reliable approach in situations where the analyses do not initialize any NWP system. Due to logistical constraints, these tests were done for a limited period of time only and for a version of MSAS using the high-resolution version of the Local Analysis and Prediction System (MesolAPS; Puri et al. 1998, National Meteorological and Oceanographic Centre 1999) as a first guess. This system was operational until July 2010, when it was replaced by the Australian Community Climate and Earth-System Simulator (ACCESS) 0.11° resolution four-dimensional variational data assimilation (4DVAR) system (National Meteorological and Oceanographic Centre 2009). The results from the statistical performance measures indicate that MSAS should be suitable for its intended applications. Although obtained for the old system version, they are indicative of the skill expected for the new system, with slight improvements since the new background field is more accurate.

2. System description

MSAS objective analyses are generated using a univariate, two-dimensional version of the SI method, an implementation tailored here to the Australian MSAS domain and observation network. Observational data are combined with NWP model hourly forecasts, generated by the 0.125° latitude–longitude resolution MesolAPS 6-hourly cycle system, nested in the 0.375° LAPS system (Puri et al. 1998). The MesolAPS forecasts are downscaled to the MSAS 2.5 arc-minute resolution 1110 × 870 regular latitude–longitude grid for use as a background in SI. The background fields are forecasts of varying ranges and accuracies. Table 1 shows the current operational configuration of MSAS.

MSAS performs the following tasks (detailed further in the text):

- preprocessing of observational data,
- downscaling of NWP model forecasts to the MSAS grid,
- quality control of observational data,
- increment analysis (using two passes of SI and MSAS specific structure functions), and
archiving of analyzed fields using the Meteorological Archival and Retrieval System (MARS).

a. Preprocessing of observational data

The current Australian surface observation network reflects population and industry concentrations, leaving vast continental and oceanic areas underobserved relative to the prescribed resolution of the MSAS analysis grid (Fig. 1). Figure 2 shows the typical amount of observational data available for analysis, from synoptic land station, ship, buoy, and METAR reports, for each reported weather parameter and per each hour of the day. METAR data dominate land observations and, after removing duplicates, constitute 80%–90% of all land observations depending on the hour of the day. There are usually no more than a dozen ship reports while there may be over 100 buoy reports, which includes closely spaced multiple reports coming from the same platform. Scatterometer wind data (not included in Fig. 2’s observation count) are only available for a few hours per day, otherwise falling outside either the domain or the time window. Because of occasional differences between scatterometer winds and other marine data, as well as occasional inconsistencies in their reports, they have been temporarily excluded from the generation of operational wind analyses, pending more systematic verification.

If a report from a fixed land station at any analysis time is missed, then a reconstructed observation is interpolated to the analysis time, if other reports from the given station are available before and after analysis time within a prespecified time window of ±3 h. Otherwise a time window of ±1 h for all observations applies. Since MSAS is not cycling, the indirect multiple uses of some observations for several consecutive analyses are permissible here. This time interpolation of observations, otherwise known as the first guess at the appropriate time (FGAT) technique, has also been tested and is pending an implementation in RTMA (de Pondeca et al. 2011), allowing the extension of an assimilation window in RTMA from ±12 to ±30 min. In MSAS the analysis window must be wider as it uses a much sparser and a less frequently reporting observation network.

To constrain extrapolation errors in the SI analysis, pseudo-observations are inserted into data voids. The pseudo-observations are generated by a combination of a downscaled first guess and a low-resolution SI analysis. In addition to a coarse grid analysis, this preliminary SI application also gives an estimate of analysis error, $e^a_k$, in SI (after Lorenc 1981), used here to derive a gridded measure of observation network data density $d$:

$$\delta_k = 1 - (e^a_k)^2 \quad \text{and} \quad (e^a_k)^2 = 1 - h_k^T M^{-1} h_k,$$

where $M$ is the matrix of error correlations at observation locations and $h$ is the vector of error correlations.
between an analysis grid point at location \((k)\) and the observation locations.

This measure uses locations of observations, and an assumed background error covariance model (described later), while the innovations are not explicitly used. Since this measure is based on the error covariances, it will vary depending on the scales to be resolved by the analysis, and on the analysis variable. It is thus suitable for defining poorly observed areas, by using threshold densities \(\delta_{\text{max}}\) and \(\delta_{\text{min}}\). The pseudo-observations are linear interpolations between the background and the preliminary analysis, based on \(\delta\), with the weights for the background ranging from 1 for densities below \(\delta_{\text{min}}\) down to 0 for densities greater than \(\delta_{\text{max}}\). The number of pseudo-observations is limited so as to minimize the effects of these bogus data and for computational reasons. The pseudo-observations are only generated on the coarse grids, and even then are only used where the data density is less than \(\delta_{\text{max}}\). The pseudo-observations can also be safely ignored in data voids such as the ocean, which are far enough away from other observations that spurious extrapolation is not an issue. Figure 3 shows contours of the observation network density proxy (estimated analysis error) for pressure (top) and wind (bottom), the distribution of the observations, and the pseudo-observations for use in the second SI analysis pass.

Effectively, the first pass SI is only used to generate the data density and pseudo-observations, whose role is to constrain the following SI analysis to the background in data voids, and to control spurious noise in the transition between data-dense and data-void areas. This analysis procedure is computationally less expensive than running two full-grid SI passes with different length scales, while aiming to generate a similar result.

The preliminary SI uses broader length scales than the main analysis, depending on the variable being analyzed—600 km for pressure and 500 km for other variables; the main length scales are given in Table 2. A longer scale is used for pressure because of the smaller number of observations and the smaller horizontal variability of MSLP.

**b. Downscaling of NWP model forecasts to the MSAS grid**

NWP forecasts are produced four times per day from initial conditions at 0000, 0600, 1200, and 1800 UTC. Each forecast provides hourly fields of surface pressure, MSLP, T2, D2, and U10 and V10.

A major issue for MSAS is the quality of the first guess. Given the low observational data density, inaccuracies in the first guess carry through to the analysis in data-sparse areas. Additionally, the operational scheduling of the Bureau’s NWP systems and the MSAS cutoff time of 15 min means that the 1200, 1300, and 1400 UTC MSAS analyses must use forecasts with at least a 6-h lead time (i.e., based on a 0600 UTC analyses). Similar problems also exist at 0000, 0100, 0200, 0600, 0700, 0800, 1800, 1900, and 2000 UTC. While the longer forecast horizon may affect the accuracy of the first guess, changes between NWP initial conditions have an adverse effect on the continuity of MSAS analyses. Furthermore, the underlying NWP system is nested in a coarser system and the initial conditions are interpolated from the coarser system analyses using bicubic interpolation. For comparison, the NCEP RTMA system uses 13-km “full physics” Rapid Update Cycle (RUC) model 1-h forecasts as a first guess (Benjamin et al. 2007).

The NWP forecast resolution is more than twice as coarse as the MSAS grid. Prior to being used as a first guess, NWP grids need to be precalculated on the finer MSAS grid in a manner that accounts for the differences in the topographies of the two grids. This downscaling of the original model forecasts depends on the variable being considered.

Downscaling of NWP forecasts to the MSAS grid also uses bicubic interpolation. For pressure, the model MSLP is simply interpolated to the MSAS grid. For temperature the downscaling is done after converting T2 to potential temperature (PT), which, importantly, is independent of elevation and therefore smoother. The estimation of surface pressure involved in this conversion uses monthly gridded pressure reduction factors, smoothly varying over a larger area and following the topography, controlled by estimates of local climatologies (see section 2d).

Figure 4 shows a sample time series for root-mean-square errors (RMSEs) between the observations and the first guess in terms of both T2 and PT. It is clear that, on average, temperature downscaling using PT is more accurate than that using T2, and that for most hours of the day it is at least as accurate as for downscaling via T2 directly. Further tests have shown that this holds for larger samples.

This method of downscaling T2 does have two limitations. First, low-level nocturnal inversions will not be apparent in the PT fields, which explains the reduced impacts of downscaling during nighttime and early morning in Fig. 4. Second, there are fewer pressure observations than T2 observations, particularly in higher-elevation areas of the domain. Consequently, the accuracy of analyzed MSLP, which is indirectly used in the conversion to PT, is somewhat compromised, adversely affecting the accuracy of both the downscaling and the following analysis.

A similar downscaling procedure is applied to D2, although the dewpoint analysis is done later in the original variable. A diurnal validation test for one (winter) day showed no apparent numerical gain, with the average
impact on accuracy being neutral. This method of downscaling introduces a structure related to the topography in the D2 first guess, some of which is later passed to the analysis. It has not yet been convincingly verified whether the extra structure introduced is more realistic than that from downscaling the original variable, although a very small positive impact on analysis accuracy (RMSE) has been noted. Testing of downscaling approaches for D2 continues.

Downscaling of 10-m wind is based on the zonal (U10) and meridional (V10) components using bicubic interpolation.

c. Quality control of observational data

The observations used in MSAS have only undergone the most rudimentary of quality checks, such as coding error detection. As with all objective analysis schemes, further quality control (QC) and validation is required.
This is accomplished in two steps, both within the analysis scheme:

1) verification that the observations are within prespecified thresholds of the background—the thresholds are defined in terms of expected innovation variances, based on the assumed background and observation error variances, and

2) cross validation of increments, where this routine QC application itself uses SI and allows identification and elimination of normalized innovations that are inconsistent with neighbors.

Earlier comparison of observations against the background and the verification presented later (see Fig. 8) indicate that the background accuracy is different at different hours of the day, which is likely related to using forecasts of different horizon as the background. As is discussed later (Figs. 7 and 8), it is also revealed that the model background currently used is somewhat different and somewhat less accurate than the one from the new (ACCESS) model.

The assumption of fixed background errors for the whole domain is only a rough approximation. Apart from the above-mentioned hourly changes, the relatively coarse original MesoLAPS forecasts are unable to resolve satisfactorily (when downscaled to MSAS resolution) weather features in rough terrain such as surface temperature inversions in the Australian Alps or surface winds channeled by Tasmanian valleys. In testing, the combination of this assumption and of the varying accuracy of background fields and the small ratio of observation to background error variance (ranging from 0.06 to 0.1) resulted in a relatively large number of observation rejections, when using the usual QC tolerances for the two QC steps. This led to the QC being relaxed, by increasing the tolerances in steps 1 and 2 of the QC procedure two- to threefold. These settings result in no more than a few hourly rejections on most occasions. While this allows the retention of more observations in important parts of the domain, the QC is somewhat compromised on average across the whole domain. Both the tolerances and background errors will need to be verified for the new (ACCESS model based) first guess.

d. Analysis procedure

The MSAS analysis procedure uses SI and employs two passes through the data for each analysis variable. In the first pass, gridded data density fields are computed and pseudo-observations are generated (see section 2a). The second pass provides the actual analyses and uses shorter length scales, but the same background as the first pass. Because of only a modest number of observations available within the domain, it is feasible to analyze all observations simultaneously for each analysis variable. This includes reported observations, MSAS time-interpolated observations, and the pseudo-observations. The pseudo-observations project the low-resolution first-pass analysis into the areas of low data density, and constrain the analysis to the background in data voids. This is needed, as in poorly observed areas extrapolations by the SI can generate large errors.

Depending on whether the scheme is set to analyze finer or coarser scales, the density of the pseudo-observations can be controlled by modifying respective data density thresholds and other parameters that control their generation.

MSAS uses observations of station-level pressure, T2, D2, and 10-m wind to generate normalized innovations of MSLP, PT, D2, U10, and V10. The quality-checked innovations are analyzed univariately. The assumed analysis order is the same as in MB92, allowing the use of analyzed pressure for a more accurate conversion of T2 to PT and of analyzed PT to affect the D2 and wind analyses.

The background error spatial correlation functions in MSAS are of the form:

<table>
<thead>
<tr>
<th>Parameter Settings for MSAS Background Error Correlation Functions</th>
<th>rp</th>
<th>rt</th>
<th>rd</th>
<th>rw</th>
</tr>
</thead>
<tbody>
<tr>
<td>r_p</td>
<td>450 km</td>
<td>400 km</td>
<td>300 km</td>
<td>300 km</td>
</tr>
<tr>
<td>r_t</td>
<td>400 km</td>
<td>400 km</td>
<td>300 km</td>
<td>300 km</td>
</tr>
<tr>
<td>r_d</td>
<td>300 km</td>
<td>300 km</td>
<td>300 km</td>
<td>300 km</td>
</tr>
<tr>
<td>r_w</td>
<td>300 km</td>
<td>300 km</td>
<td>300 km</td>
<td>300 km</td>
</tr>
</tbody>
</table>

FIG. 4. RMSEs of downscaled NWP temperature and potential temperature first-guess fields on 11 Aug 2009. Comparisons are between bicubic interpolations done in the original variable and a transformation to potential temperature. Downscaled fields were interpolated to observation sites and the RMSE between interpolated and observed values averaged over all observation sites within the domain.
MSLP: \( \exp(-f_{p}d^{2}/r_{p}^{2}) \),

PT: \( \exp(-f_{d}d^{2}/r_{d}^{2})[1.0 + K_{z}(\Delta z)^{2}] \),

D2: \( \exp(-f_{d}d^{2}/r_{d}^{2}) \), and

U10, V10: \( \exp(-f_{w}d^{2}/r_{w}^{2}) \),

where \( d \) (m) is the distance between an observation and a grid point; \( f_{p}, r_{p}, r_{d}, \) and \( r_{w} \) (m) are length scales for respective analysis variables; \( \Delta z \) (m) is the elevation difference between an observing site and a grid point; \( K_{z} \) (m\(^{-2}\)) is the vertical correlation function modifier for \( \Delta z \); and \( f_{p}, f_{d}, f_{w}, x_{p}, x_{r}, x_{d}, \) and \( x_{w} \) are fixed modifiers to Gaussian correlations, with the current parameter setting as in Table 2. The length-scale settings were determined after comprehensive empirical testing, which involved withholding observations, monitoring analysis errors against first-guess errors at observation locations, and visual assessment of analysis coherence. The outcomes of such testing are however adversely affected by the observation network inhomogeneity and by the varying accuracy of the first guess.

The role of Gaussian correlation modifiers is twofold: to allow a broadening of the peak, thus increasing the weights of observations close to a grid point being analyzed, and to control the distance at which correlation values reach machine precision. Figure 5 shows plots of MSAS structure functions and those used in MESAN (for temperature and humidity analysis, but with an omission of corrections for difference in land fraction and height, Hägmark et al. 2000) and in NOAA MSAS (MB92). The allowance for anisotropic modification of temperature background error correlation for elevation difference has been applied here after MB92 and also the current NOAA MSAS (information online at http://msas.noaa.gov/msas_descrip.html). The impacts, however, have not yet been sufficiently assessed here by withholding observations and consequently in the operational version of MSAS \( K_{z} \) is set to 0. With \( K_{z} \) greater than 0, observations at elevations similar to the analyzed grid point are given more weight. The average effect is to somewhat reduce the correlations used in SI by shortening the effective length scale. Overall, this may improve the fit of the analysis to observational data where observations are dense, as described in MB92. Within the Australian domain however, observations are much sparser, and the elevation decorrelation makes the network even sparser. This potentially causes other analysis problems in already data-sparse regions. For example, the Great Dividing Range in Victoria is considered sparsely observed in this context. Furthermore, temperature observations there are not well supported by pressure observations, while both variables are needed to analyze PT. Consequently, when using the vertical correction factor \( K_{z} \), the temperature length scale in MSAS would need to be somewhat increased, to compensate for the shortening of the effective length scale, while taking advantage of the preferential weighting. MB92 also allow the use of analyzed PT differences for the anisotropic modification of dewpoint and wind correlations. Although this option has been tested, it remains switched off in the operational version due to there being no clear positive impacts on the analyses. Further testing of this option using ACCESS-based background fields along with testing elongated wind correlations will be considered in the future.

As has been found for this system and elsewhere (e.g., MB92), a selection of a set of parameter values for the whole analysis domain is a compromise in the sense that the set is not necessarily expected to be optimal everywhere within the domain and for varying analysis times. Simultaneous optimal tuning of analysis parameters given large parameter space remains a challenging task; consequently, some degree of subjectivity in their settings is unavoidable, which shows up for example in the usual rounding of parameter values.

The analyzed MSLP, adjusted to surface pressure, is used to convert first-guess T2 to PT, and back to T2 postanalysis. Pressure conversion coefficients (an average pressure change per meter altitude at gridpoint elevations) are estimated locally, on a monthly basis from nonlinear regression equations (second-order polynomials). The regression coefficients are estimated using pressure reports from a selection of stations surrounding each grid point. Average pressure changes per meter altitude at each station are the predictands and the station elevations are the predictors. The observed surface pressure, on-site reduced MSLP, and the station elevation are used to calculate the predictands at each station.
the Australian domain pressure reductions at observing stations use monthly coefficients (based on site climatologies), so the derivation of pressure conversion coefficients for the MSAS grid needs only to be done once a month (alternatively the previous year’s coefficients may be used). Where possible, the selection of stations is done in a way that ensures horizontal proximity to respective grid points, sufficient range in station elevations, and omission of stations at low elevations for which the predictands calculated from reports are less accurate. If low-elevation (i.e., 200 m) stations are selected, the pressure changes per meter at their locations are replaced with that derived by using the Gibbs pressure reduction method (Seaman 1997), based on monthly climatologies of mixing ratio, air temperature, and geopotential heights at 850 and 1000 hPa. This technique aims to approximate local monthly reductions at pressure observing sites, in a spatially coherent manner, and will be detailed separately.

The high resolution of the analysis grid allows detailed topography to enhance PT (and thus T2) analysis details in rough terrain even in data-sparse areas. There is also an apparent detail enhancement in the D2 analysis in rough terrain, which is not necessarily supported by the observations. This is in part the result of the dewpoint downscaling technique, which uses topography. The realism of this apparent detail remains to be verified.

After the T2 and D2 analyses are completed, a smooth land—sea mask is used to replace respective analyses over ocean with the numerical model first guess. The impacts of temperature and dewpoint observations over ocean sites (ships, reef stations) are retained in the analysis over land near the coast.

The use of a persistence first guess in MSAS is an option for the infrequent situations when a model first guess is not available. However, the inhomogeneous observation network does not support the use of persistence for a prolonged period. The value of persistence as a first guess fields in data voids, subject to correct specification of observational and first-guess error variances and correlations. Precise background error specification over the domain for each hour of day, given the 6-hourly NWP cycle and fluctuations in NWP performance, is not possible. Therefore, following common practice (e.g., MB92), constant background error variances are used.

Although an attempt has been made in the design of MSAS to reduce the effects of the irregular observational data distribution, MSAS analyses are still greatly affected by the first-guess accuracy and availability at analysis time. The impacts of changes in the first guess can be seen in Fig. 7 (top), which shows the average differences in MSLP and T2 between LAPS and the newer system ACCESS-R (National Meteorological and Oceanographic Centre 2009). LAPS is used as a proxy for MesolAPS, to match the domain and resolution of the ACCESS-R system. The LAPS and ACCESS-R systems both cover the MSAS domain and are at the same resolution of 37.5 km so comparison avoids regridding and the associated interpolation errors. This impact is also seen in standard deviations of differences in D2 and T2 fields shown in Fig. 7 (bottom). For MSLP a clear bias is seen in eastern Australia, while the D2 and T2 standard deviations differ more substantially in the west. Biases in the model first guess (similar to those in LAPS) will transfer to MSAS.

3. Analysis accuracy

The analysis accuracy is determined by the accuracies of the observations, background fields, their error correlations, and a range of control parameters. The level of accuracy is expected to be higher where observational data are dense and to decrease to the accuracies of the first-guess fields in data voids, subject to correct specification of observational and first-guess error variances and correlations. Precise background error specification over the domain for each hour of day, given the 6-hourly NWP cycle and fluctuations in NWP performance, is not possible. Therefore, following common practice (e.g., MB92), constant background error variances are used.

Figure 6 shows mesoscale features captured in MSAS analyses versus larger-scale MesoLAPS forecasts during a catastrophic bushfire event on 7 February 2009. The selected area is centered on Melbourne and includes Kinglake—one of the main areas affected by the bushfires. A sequence of wind and temperature MSAS analyses (valid at 0400, 0800, 1000, and 1200 UTC), and the equivalent time MesoLAPS forecasts, are presented against a topography background, with its highest elevation exceeding 1500 m in the eastern center of the area.

Strong, hot northerly winds up to 51 km h$^{-1}$ (69 km h$^{-1}$ in gusts), brought record high temperatures reaching 46°C in Melbourne City and approaching 48°C on the city’s outskirts. As reported in Tolhurst (2009), the fires around Kinglake and nearby areas first burnt in a long narrow strip aligned with the winds. The change in wind direction associated with the passage of a cool change (0800 UTC) caused a spread of this cigar-shaped fire into a fire front several kilometers wide. MSAS wind analysis at 1000 UTC correctly captures a temporary wind reverse from southwesterlies east of Melbourne back to northerlies and an associated mesoscale wind circulation around Kinglake, not evident from the larger-scale MesoLAPS analysis. At 1200 UTC, southwesterlies dominate the flow except for the ranges to the northeast of the area. The vertical variation of surface temperature clearly follows the topography. The T2 analyses are more accurate than the respective MesoLAPS analyses, although T2 observations are not shown to avoid cluttering the charts.

e. Case study of February 2009 Victorian bushfires

Figure 6 shows mesoscale features captured in MSAS analyses versus larger-scale MesoLAPS forecasts during a catastrophic bushfire event on 7 February 2009. The selected area is centered on Melbourne and includes Kinglake—one of the main areas affected by the bushfires. A sequence of wind and temperature MSAS analyses (valid at 0400, 0800, 1000, and 1200 UTC), and the equivalent time MesoLAPS forecasts, are presented against a topography background, with its highest elevation exceeding 1500 m in the eastern center of the area.

Strong, hot northerly winds up to 51 km h$^{-1}$ (69 km h$^{-1}$ in gusts), brought record high temperatures reaching 46°C in Melbourne City and approaching 48°C on the city’s outskirts. As reported in Tolhurst (2009), the fires around Kinglake and nearby areas first burnt in a long narrow strip aligned with the winds. The change in wind direction associated with the passage of a cool change (0800 UTC) caused a spread of this cigar-shaped fire into a fire front several kilometers wide. MSAS wind analysis at 1000 UTC correctly captures a temporary wind reverse from southwesterlies east of Melbourne back to northerlies and an associated mesoscale wind circulation around Kinglake, not evident from the larger-scale MesoLAPS analysis. At 1200 UTC, southwesterlies dominate the flow except for the ranges to the northeast of the area. The vertical variation of surface temperature clearly follows the topography. The T2 analyses are more accurate than the respective MesoLAPS analyses, although T2 observations are not shown to avoid cluttering the charts.

3. Analysis accuracy

The analysis accuracy is determined by the accuracies of the observations, background fields, their error correlations, and a range of control parameters. The level of accuracy is expected to be higher where observational data are dense and to decrease to the accuracies of the first-guess fields in data voids, subject to correct specification of observational and first-guess error variances and correlations. Precise background error specification over the domain for each hour of day, given the 6-hourly NWP cycle and fluctuations in NWP performance, is not possible. Therefore, following common practice (e.g., MB92), constant background error variances are used.

Although an attempt has been made in the design of MSAS to reduce the effects of the irregular observational data distribution, MSAS analyses are still greatly affected by the first-guess accuracy and availability at analysis time. The impacts of changes in the first guess can be seen in Fig. 7 (top), which shows the average differences in MSLP and T2 between LAPS and the newer system ACCESS-R (National Meteorological and Oceanographic Centre 2009). LAPS is used as a proxy for MesoLAPS, to match the domain and resolution of the ACCESS-R system. The LAPS and ACCESS-R systems both cover the MSAS domain and are at the same resolution of 37.5 km so comparison avoids regridding and the associated interpolation errors. This impact is also seen in standard deviations of differences in D2 and T2 fields shown in Fig. 7 (bottom). For MSLP a clear bias is seen in eastern Australia, while the D2 and T2 standard deviations differ more substantially in the west. Biases in the model first guess (similar to those in LAPS) will transfer to MSAS.

Unauthenticated | Downloaded 07/11/22 03:38 PM UTC
FIG. 6. Comparison of (a)–(d) wind and T2 MSAS analyses and (e)–(h) MesoLAPS forecasts valid at (a),(e) 0000, (b),(f) 0800, (c),(g) 1000, and (d),(h) 1200 UTC 7 Feb 2009 against topography background (m) over central Victoria. The red barbs mark wind observations: Melb, Melbourne City; Kl, Kinglake.
analyses in data voids. The large standard deviations and biases in the first guess also indicate that significant noise may be present in observed increments from the first guess.

The suite of ACCESS systems also includes a 0.11° ACCESS-A system covering Australia and 0.05° ACCESS-C systems covering smaller domains of Australia’s main population centers (National Meteorological and Oceanographic Centre 2009). Initial verification against observational data has shown members of the ACCESS suite to be more accurate than relevant members of the LAPS suite, as the ACCESS systems make better use of observational data with their 4DVAR assimilations. The ACCESS-based systems are the newly available source of first guesses for MSAS—initially ACCESS-A and after that ACCESS-A with ACCESS-C inserts. Figure 8 shows the gains from using ACCESS-A over MesoLAPS as a first guess for MSAS. The improved first guess will particularly benefit data-sparse and data-void areas in the domain.

Quantitative estimates of analysis accuracy are determined by withholding observations. The SI equations are derived and optimized using observational data, including their accuracies and their locations and, therefore, are expected to be most accurate at these locations. They are valid at other locations in between (or close to) observations, but may be less accurate there. Randomly withholding observations from analysis and verifying it at their sites against their values allows an independent verification of the analysis technique. This verification technique is well established (e.g., Gandin 1963; Thiebaux 1975). For the tests here, 2.5% of the observations for each variable were randomly withheld over the period 18–27 July 2009 (216 consecutive hours, with a different observation set for each analysis hour), corresponding to approximately 2300–2800 observations, depending on the variable. The hourly average fit to withheld observations is measured in terms of RMSE, bias, mean absolute deviation (MAD), and the correlation between observed and analyzed deviations.

Withholding observations from the analysis however compromises the analysis technique. For this reason and because of logistical considerations, the verification period has not been extended to the whole year. The fit of the analysis to the observations is compared with a respective fit obtained for the first guess to show the gain in accuracy achieved by the SI analysis (Fig. 9). With all of these measures and considering the sample size, the gains in accuracy for MSAS analyses relative to the first guess are statistically significant, for each analysis variable.

It follows from the SI theoretical formulation that optimization of SI equations should result in zero bias,
FIG. 8. Comparison of accuracies for MesoLAPS and ACCESS-A based MSAS analyses of (a),(f) MSLP, (b),(g) T2, (c),(h) D2, (d),(i) U10, (e),(j) V10. (a)–(e) RMSE and (f)–(j) bias of hourly averaged first guess and analysis fit to observations for 10–18 Jul 2010 (navy, MesoLAPS first guess; green, ACCESS-A first guess; pink, MesoLAPS-based analysis; blue, ACCESS-A-based analysis).
when analysis is validated over all observations. Thus, the smaller the bias at withheld observation sites, the better performing the analysis is. The positive bias in V10 may indicate a need of some adjustments to the V10 structure function parameters (e.g., an elongation in the meridional direction). While useful for the assessment of analysis accuracy of U10 and V10, the statistics in Fig. 9 are less suitable for direct inference about wind speed performance, which has not been assessed in this study. The correlation coefficient as a single normalized measure is useful for comparing the analysis performance of different analysis variables. It gives an indication of the analysis accuracy gain over the first guess similar to the gain inferred from the MAD chart. Jointly, Fig. 9 statistics are helpful when tuning system parameters.

For temperatures, the RMSE statistics are also compared on an hourly basis. An example is shown in Fig. 10. It gives a rough idea of the varying analysis performance over the 24-h period and of the impacts of the first guess on this daily variation. For logistical reasons this test is for 1 day only. A more comprehensive evaluation is postponed until the system uses a first guess from the 0.11° ACCESS-A system (with 0.05° ACCESS-C inserts). For an ideal system, the differences between the RMSE at withheld observations and at used observation sites should stay small. The main contributor to the observed spikes in the time series is the variation in the first-guess hourly accuracy, as both this and withheld observation lines follow a similar pattern. Neglecting the spikes, the average magnitude of the gap between the lines for withheld and used observations is also (in a small part) contributed to by an indiscriminate sampling of withheld observations, some of which would lie on the edges of data voids. These observations have a large impact when included (Seaman and Hutchinson 1985), so withholding these observations has a large impact on the analysis of nearest-neighboring observations. Another possible contribution could be from SI overfitting at observing sites, due to the imperfect specification of error statistics, length scales, or other structure functions parameters.

FIG. 9. Fit of MSAS analyses interpolated to withheld observation sites in the domain: (a) RMSE, (b) MAD, (c) bias [MSAS (first guess) — observations], and (d) correlation of observed and analyzed deviations, for test period 18–27 Jul 2009 (216 consecutive hours). For T2 the statistics were obtained in PT.

FIG. 10. RMSE of temperature first guess and MSAS analyses interpolated to observation locations on 5 Aug 2009.
While MSAS parameters can benefit from further tuning, a plausible explanation is that a large part of the difference is due to low average observational data density, and that the analyses are already close to the limits of accuracy. An RMSE difference of a similar magnitude, roughly 1°C, is also reported by Kahler and Myrick (2007), who give an evaluation of the RTMA temperature analysis (with a RUC first guess) over complex terrain, using similar performance measures. The line for direct interpolation in T2 has the largest average RMSE, confirming (cf. Fig. 4) that a substantial part of the RMSE reduction in MSAS temperature analysis is due to downscaling in PT. Finally, the apparent cyclic variation in accuracy may be attributed to varying accuracies of model forecasts used as the first guess, which are updated by the 6-hourly NWP cycle. This cyclic variation in RMSE would also to some extent be contributed to by variations in observation availability and by nocturnal temperature inversions. Although the latter factor was not necessarily present in the test case, surface temperature inversions are a relatively frequent phenomenon.

The MSAS performance relative to the first guess is shown in Fig. 11 as a function of elevation. Considerable gains are apparent in both the RMSE and bias reductions for all analysis variables; however, these gains are an overestimate, as they involve observations used in the analysis. Tests where observations are withheld from each elevation category in turn have not yet been performed. Such tests were done by Kahler and Myrick (2007) for the RTMA system although for a much smaller domain of ~3600 km² and for a considerably denser observation network. While not directly comparable, there are apparent similarities when considering MSAS assessment in Figs. 9–11. Both MSAS and RTMA have biases (first guess–observations) in their respective PT background fields (−0.90°C in RUC and −1.20°C in MesoLAPS; see Fig. 8) that are considerably reduced in the analyses. In the RTMA case, RMSE also increases with elevation and the reported bias is largest for the highest-elevation range. In MSAS (Fig. 8) the bias is negligible when averaged over all withheld observation locations. In the comparisons stratified by elevation, the MSAS RMSE and bias increase in magnitude with increasing elevation (with an exception for dewpoint biases), following the accuracy of the respective background fields. Different atmospheric variables change with altitude in a different way, which may be nonlinear, and their observations are usually less representative of the surrounding environment in complex terrain, hence the growth of the analysis error with altitude. The temperature bias in MSAS at higher elevations has been noted by users, but has not yet been quantified by withholding observations tests versus biases in the first guess. The fact that there are fewer pressure observations than temperature observations in the domain contributes to PT biases.

The evaluation of MSAS performance described above is predominantly statistical in nature and limited in extent due to computational and other considerations. While it is possible to demonstrate significant improvements in statistical measures of analysis accuracy relative to downscaled NWP forecasts, it is more difficult to achieve a stable, uniform level of performance for the system over the whole domain and for consecutive analysis hours. This difficulty is attributed to gaps in the observation network (in particular over central Western Australia), varying accuracies of the NWP first-guess fields, and limitations of the analysis technique. An adverse impact of surface temperature inversions on analyses, especially in regions dominated by topography, remains a problem that needs to be further investigated.

4. Ongoing and future work

The most immediate change in MSAS is the use of an improved first guess by moving from MesoLAPS as input to ACCESS-A combined with ACCESS-C inserts. Other potential improvements to the analysis scheme should be identified from testing of alternative downscaling procedures, including those using low-level NWP model outputs and grid topography to improve the first-guess vertical structure (RTMA; Benjamin et al. 2007; Huth 2005). Also, a comparison will be made of MSAS performance against ACCESS NWP outputs of equivalent resolution.

Further investigation of background error correlations and of the use of anisotropic functions for higher elevation terrain and coasts is planned. In particular, testing of the $K_z$ factor, scaled by the observation density factor, in PT correlation function is an area of interest.

5. Summary

The design and performance of the Mesoscale Surface Analysis System developed for the Australian domain have been presented and compared with similar developments in North America and Europe. Significant reductions in the error statistics in the analyses versus the NWP model derived first guesses have been confirmed by observation-withholding tests. The system description captures a stage in MSAS development rather than a detailed documentation of a finished product. The MSAS operating environment—available observations, background fields, specification of topography, domain extension, computing environment, etc.—has undergone many changes over the years, forcing continual validation and tuning of the system parameters. MSAS
FIG. 11. (a)–(e) RMSE and (f)–(j) bias in first guess and MSAS analyses interpolated to vertically stratified observation locations for (a),(f) MSLP, (b),(g) PT, (c),(h) D2, (d),(i) U10, and (e),(j) V10, averaged for June 2009 over all days and all analysis hours. The parentheses give the average number of observations per analysis.
uses a version of the SI scheme that has been tailored for use over Australia, with specific background error correlations and an atmospheric pressure reduction technique, prescribed for the selected analysis variables. Issues associated with a sparse observation network are mitigated by using time-interpolated observations and pseudo-observations. While the MSAS design may be considered validated and the system is operational, further tuning will be required when sourcing background fields from ACCESS 4DVAR NWP.

A regression-based pressure reduction technique, consistent with local reductions at observing sites, was developed for and used in MSAS. It is spatially coherent, uses gridded monthly coefficients based on relevant site climatologies, and may be applied to gridded or observing site pressure data alike. The technique has been indirectly verified here by application to pressure and temperature analysis.

Downscaling of the temperature field with detailed topography and pressure information was shown to be more accurate when done in the PT field rather than directly in the T2 field, at least in the absence of surface temperature inversions.

Acknowledgments. This communication arises from the involvement of TG in MSAS development for the needs of data assimilation and in support of the GOCF and GFE projects. MSAS is complemented by a display system developed by YX. The authors wish to acknowledge the contributions of and offer their thanks to the following CAWCR–Bureau of Meteorology staff members: Tan Le for assistance with MARS archiving, Ilia Bermous for improvements to MSAS code performance, and Michael Foley (Northern Territory Regional Office) for sharing the results of indirect verification of MSAS within the context of GFE. Thanks are extended to Jim Fraser from the National Meteorological and Oceanographic Centre (NMOC) for enabling verification of the CAWCR version of MSAS against the NMOC parallel trial and Robert Seaman, Elizabeth Ebert (CAWCR/BoM), and the anonymous journal reviewers for thorough review and helpful comments, which improved the paper.

REFERENCES


Unauthenticated | Downloaded 07/11/22 03:38 PM UTC


