

A Brief Guide to Observation Nudging in WRF
Brian Reen
Battlefield Environment Division, Army Research Laboratory
February 2016

1	Introduction.....	2
2	WRF Nudging Methodology	4
2.1	Determining If Observation Nudging is Used.....	4
2.2	Reading Observations and Calculating Innovations	4
2.3	Weighting Functions	8
2.3.1	Overall Weighting Function	8
2.3.2	Temporal Weighting.....	9
2.3.3	Vertical Weighting.....	10
2.3.4	Horizontal Weighting.....	13
3	Preparing Observations.....	15
4	Diagnostic Prints.....	18
5	Settings in WRF-ARW Observation Nudging.....	21
	Appendix A: OBS_DOMAIN format and its ingestion by WRF	26
	Appendix B: WRF Observation Nudging Code Layout.....	32
	Acknowledgements:.....	33
	References:.....	33

1 Introduction

This document describes the use of observation nudging data assimilation in the Weather Research and Forecasting model (WRF; Skamarock et al. 2008) as of WRFV3.7.1. In addition to the references cited below, and examining the source code itself, another source of information on the use of observation nudging in WRF are two files distributed with WRF: README.namelist and README.obs_fdda.

Observation nudging is a form of Newtonian relaxation wherein artificial tendency terms are introduced into the model to gradually “nudge” the model towards observations; it is also known as station nudging. The other forms of nudging present in WRF (analysis nudging and spectral nudging) relax the model towards an analysis. Spectral nudging is a special form of analysis nudging that only nudges the large-scale signal. In cases where the temporal resolution of the global model used to provide lateral boundary conditions is quite coarse, this allows one to retain the large-scale signal from the global model without retaining the smaller-scale signal from the global model (which is not sufficiently temporally resolved). Spectral nudging is normally used for very long simulations (i.e., regional climate simulations running for months or years and driven by global climate models).

Nudging is a type of four-dimensional data assimilation (FDDA), since it is not applied at a single time. Additionally, nudging is a continuous form of data assimilation since it is applied at every time step over some period. This is in contrast to some data assimilation techniques that change the model solution at certain analysis times, and are thus intermittent data assimilation techniques.

Thompson (1961) used a dynamical model to create an observation-based analysis, and Charney et al. (1969) suggested using model integration to create an analysis with observations. Some of the early nudging studies include application of nudging to a simple model (Kistler 1974), 2D hurricane modeling (Anthes 1974), 1D/2D modeling (Hoke and Anthes 1976), and 3D hurricane modeling (Hoke and Anthes 1977). Stauffer and Seaman (1990) and Stauffer et al. (1991) described the application of nudging to the Pennsylvania State University (PSU) / National Center for Atmospheric Research (NCAR) Mesoscale Model (MM4; Anthes et al. 1987), an earlier version of what eventually became MM5 (Grell et al. 1995). As detailed by Liu et al. (2005), a version of the MM5 nudging code modified by NCAR and known as Real-Time Four Dimensional Data Assimilation (RTFDDA) was the basis for the original implementation of nudging into WRF; however, not all aspects of the system described by Liu et al. (2005) were implemented in the public-release version of WRF. Some notable updates to the nudging in the public-release version of WRF, which resulted from a PSU/NCAR project funded by the Defense Threat Reduction Agency, were detailed in Deng et al. (2009).

As with any data assimilation methodology there are advantages and disadvantages of the nudging technique. One advantage is that nudging is computationally simple, since it only requires adding an additional tendency term. Another advantage is that it does not require the detailed error covariance information required by some other methods such as three-dimension variational (3DVAR). The continuous nature of nudging allows only relatively small changes to be made to the model solution at any given timestep. This makes it more likely that the model’s physical tendency terms will be able to keep the model solution in balance than when intermittent data

assimilation techniques apply much larger changes at one or only a few timesteps. Nudging also makes it rather simple to incorporate physically-based modifications to the functions used to spread information from an observation (e.g., spreading the influence of surface observations throughout the convective boundary layer due to the expected well-mixed nature of the convective boundary layer). One disadvantage of traditional nudging is that it is not able to directly ingest observations of variables not forecasted by the model. For example, radial velocity observations from radar cannot be directly used in nudging because the model does not directly predict radial wind; because there is no radial wind velocity variable in the model, we cannot nudge the model radial wind velocity directly. In contrast, other methods (e.g., 3DVAR, four-dimension variational [4DVAR], and Ensemble Kalman filter [EnKF]) often use a forward operator to ingest these observations. Another disadvantage of nudging is that the weight at which an observation is applied is generally not flow-dependent. Hybrid data assimilation methods combining nudging with other data assimilation techniques have been developed in an effort to combine the strengths of multiple techniques while mitigating the weaknesses of the individual techniques. For example, Lei et al. (2012) developed a hybrid nudging-EnKF approach (HNEnKF) and applied it to WRF, and Liu et al. (2015) developed a hybrid nudging-EnKF approach they refer to as a “4-dimensional relaxation ensemble Kalman filter” (4D-REKF). These methods allow nudging weights to be flow-dependent.

Nudging is applied to either improve the initial conditions of the forecast (dynamic initialization) or to provide an estimate of past weather conditions (dynamic analysis). When used in dynamic initialization, nudging is applied during a pre-forecast time period, and then a model forecast is launched as a result of this dynamic initialization. There will be no observations with a valid time during the forecast that are nudged towards during the forecast; observations from prior to the forecast may continue to be nudged towards during the beginning of the forecast. Dynamic initialization provides a method to improve model conditions by ingesting observations over some time period into a dynamically consistent model framework. Dynamic analysis continues to use data assimilation throughout the model simulation. This can be used to create an analysis of past conditions that includes all available observations. These analyses can be used to drive other models such as atmospheric transport and dispersion models or air quality models. Dynamic analysis can also be used to provide an analysis of the current conditions (i.e., a 0-h forecast)

Observation nudging uses the difference between the model and an observation to create an innovation that is multiplied by various factors and added to model tendency equations to gradually nudge the model towards observations (e.g., Stauffer and Seaman 1994; Deng et al. 2009). In WRF this is implemented as:

$$\frac{\partial q \mu}{\partial t}(x, y, z, t) = F_q(x, y, z, t) + \mu G_q \frac{\sum_{i=1}^N W_q^2(i, x, y, z, t) [q_o(i) - q_m(x_i, y_i, z_i, t)]}{\sum_{i=1}^N W_q(i, x, y, z, t)} \quad (1)$$

where q is the quantity being nudged (e.g., water vapor mixing ratio), μ is the dry hydrostatic pressure, F_q represents the physical tendency terms of q (for water vapor this includes advection, diffusion, conversion from water vapor to cloud water, etc.), G_q is the nudging strength for q , N is the total number of observations, i is the index to the current observation, W_q is the spatiotemporal weighting function based on the temporal and spatial separation between the observation and the current model location, q_o is the observed value of q , and $q_m(x_i, y_i, z_i, t)$ is the model value of q

interpolated to the observation location. The quantity $q_o - q_m$ is the innovation; the innovation associated with a given observation evolves with time (both before and after the time of the observation) as the model value (q_m) evolves. Thus, as the model value approaches the observed value, the nudging tendency term decreases.

Note that the innovation is spread spatially, not the observation itself; therefore, observation nudging assumes that error at the observation location is correlated with the error at other locations in the model domain, not that the field itself is spatially correlated. The inverse of the nudging weight (G_q) is the e-folding time of the innovation, the time that it takes for the innovation to decrease by $1/e$ assuming that the physical tendency terms do not exist (i.e., $F=0$). Analogous equations are used for water vapor mixing ratio and the u and v wind components. Note that the model ingests temperature observations and then converts them to potential temperature for application. The nudging tendency term should not dominate the other tendency terms, since the physically-based tendency terms need to be able to keep the model solution dynamically consistent as the observations are applied.

2 WRF Nudging Methodology

While the basic method of observation nudging is described in Equation 1, there are many other details involved in the WRF observation nudging. The following is a description of the technique, as implemented in WRF.

2.1 Determining If Observation Nudging is Used

Whether observation nudging is to be utilized is determined by *obs_nudge_opt*, and then additional switches are used to determine to which variables observation nudging is applied (*obs_nudge_wind*, *obs_nudge_temp*, and *obs_nudge_mois*).

2.2 Reading Observations and Calculating Innovations

Observations are first read into WRF from a file OBS_DOMAINX01 (where X is the domain number) in OBS_DOMAIN format. Section 3 describes the preparation of observations for use in observation nudging. Observations that are too far in the past to be used in nudging are immediately discarded and WRF stops reading when it is determined that the latest observation applicable to the current time is read. A maximum of *max_obs* observations can be available for observation nudging at a given time on each domain; if at any time any domain has more than *max_obs* observations that are considered to be neither too early nor too late WRF will exit with an error. Any observations with a pressure lower than 80 hPa are discarded. Also, any surface observation whose pressure falls outside the range 700 – 1050 hPa will have its pressure and temperature discarded. This serves as a gross error check on the pressure since the pressure will be used to convert the observed temperature to the potential temperature used for nudging. However, since there may be cases where valid surface pressure observations fall outside this range the user should be aware of this pressure check (especially when modeling in high elevation areas such as Tibet or very low elevation areas such as the Dead Sea region). If the winds are marked as earth-relative (quality control flag set to 129) they are transformed to grid-relative, otherwise WRF assumes that the observations are already grid-relative. This distinction is important because

directions within the model are not necessarily the same as directions relative to the earth due to the need to use a projection in order to transform the non-flat earth onto a flat model domain.

WRF is intended to check for new observations at every time step when the nudging tendency term is recalculated, which is designed to be every *obs_ionf* coarse-grid time steps. Although recalculating the nudging tendency terms every time step would make the nudging more responsive to model changes it increases the computational cost of observation nudging. For sufficiently frequent updating (small *obs_ionf*), the changes in the modeled conditions should be sufficiently small that whatever small improvements might be gained by recalculating nudging terms every time step are most likely outweighed by the increased computational cost.

Although *obs_ionf* is intended to be the interval between the updating of both nudging tendency terms and the interval between each check for additional observations applicable to nudging at the current time, there are two issues that may prevent this of which the user should be aware.

The first involves properly setting the frequency at which the WRF I/O checks for observation nudging input. Two settings in the *time_control* section of the namelist control when the model checks to see if *obs_ionf* indicates that one should check for additional observations: *auxinput11_interval* and *auxinput11_end_h*. The setting *auxinput11_interval* indicates how frequently the model should check to see if the *obs_ionf* setting implies that the observation files should currently be checked for additional observations. Therefore, the user should set *auxinput11_interval* to 1 so that at each time step WRF checks if *obs_ionf* indicates that additional observations should be checked for. Other similar variables that are variations of *auxinput11_interval* are also available that are based on time units rather than being in units of grid time steps; these should be avoided due to the potential for unintended consequences if the settings are not carefully thought through and updated when other aspects of the configuration are altered. The code will stop checking to see if additional observations should be read in at *auxinput11_end_h*.

Second, *obs_ionf* may or may not work as desired on the finer domains due to assumptions made in the implementation of *obs_ionf*. This variable is in terms of coarse grid model time steps and so must be translated to the time steps of the domain in question. However, the code does this translation based on the horizontal grid spacing ratio of the current domain to the next coarsest domain. Namely, the number of time steps between nudging tendency term updates on nest *n* in terms of nest *n* time steps is $obs_ionf(n) * [parent_grid_ratio(n)^{nest_level(n)}]$, where *parent_grid_ratio(n)* is the ratio between the parent domain horizontal grid spacing and the current domain horizontal grid spacing, and *nest_level(n)* is the number of parent domains (e.g., if nest *n* is a 1-km horizontal grid spacing domain that is nested within a 3-km domain that is nested within a 9-km domain then *nest_level* is 2). This method works if $parent_grid_ratio(n) = parent_time_step_ratio(n)$ (i.e., the ratio of timestep length between the current domain and the parent domain is equal to the ratio of horizontal grid spacing between the two domains) and *parent_time_step_ratio* is the same for domain *n* as for all parent domains. If this is not true, then the actual interval between nudging tendency term updates may vary significantly from that specified by the user. Until this issue is resolved, the user should ensure that the actual nudging tendency term update interval for the value of *obs_ionf* they have chosen is consistent with what they intend.

Observations are marked as surface observations if the “is_sound” field in the OBS_DOMAIN files is marked as false. Therefore, all observations that are not surface observations should have the is_sound field marked as true, even if the observation is not a sounding.

The next step is to calculate the innovations (differences between the model and the observations; $q_o(i) - q_m(x_i, y_i, z_i, t)$ in equation 1). When the observation was read in the surface observation was assigned to the first vertical model level for purposes of calculating the innovation. At this point, for all observations not classified as surface observations the pressure is used to determine the placement of the observation in model vertical-level space. If the pressure of the observation is not available, it will first use observation height and model height and pressure to determine observation pressure. Note that the placement of the observation in model vertical-level space (and where applicable the determination of pressure based on height) is made when the observation is first used and is not recalculated at future times. This means that even if the pressure changes with time, the vertical location where the innovation is calculated will not be updated. The model value is then interpolated in 3D (i.e., 8 points) to the observation location and the model value is subtracted from the observed value to find the innovation. If we are dealing with a surface temperature or surface wind observation, the observation is multiplied by a factor to account for the difference in height between the lowest model level and the height of the observation. These factors are calculated based on the model relationship between the lowest prognostic level and 2-m/10-m diagnosed values where possible. Thus these factors are case and time specific and are based on WRF’s understanding of the meteorological conditions.

The user can prevent nudging towards a negative water vapor mixing ratio by setting `obs_scl_neg_qv_innov` to “1”. Recall equation 1:

$$\frac{\partial q\mu}{\partial t}(x, y, z, t) = F_q(x, y, z, t) + \mu G_q \frac{\sum_{i=1}^N W_q^2(i, x, y, z, t) [q_o(i) - q_m(x_i, y_i, z_i, t)]}{\sum_{i=1}^N W_q(i, x, y, z, t)} \quad (1)$$

in which the innovation is calculated at time t as the difference between the i th observation [$q_o(i)$] and the modeled quantity at the observation location [$q_m(x_i, y_i, z_i, t)$], and this innovation is then applied spatially to a point (x, y, z) . If the observed value is less than the modeled value at the observation location [$q_o(i) < q_m(x_i, y_i, z_i, t)$] then the innovation is negative, and in the case of water vapor mixing ratio, the model is nudged towards drier conditions. If the location at which the negative innovation is being applied (x, y, z) is a location where the modeled water vapor mixing ratio is less than the magnitude of the innovation [$q_m(x, y, z, t) < |q_o(i) - q_m(x_i, y_i, z_i, t)|$], then the model is effectively nudged towards a negative water vapor mixing ratio. In some circumstances this can lead to excessive drying. For example, consider the case where the observed water vapor mixing ratio is 5 g kg^{-1} [$q_o(i) = 5 \text{ g kg}^{-1}$], and the model value at that location is 10 g kg^{-1} [$q_m(x_i, y_i, z_i, t) = 10 \text{ g kg}^{-1}$]. The observed value is less than the corresponding model value, and so the innovation is negative [$q_o(i) - q_m(x_i, y_i, z_i, t) = -5 \text{ g kg}^{-1}$] and the model is thus nudged towards drier conditions. However, if this innovation is spread to a location (x, y, z) where the modeled water vapor mixing ratio is 3 g kg^{-1} , this is less than the magnitude of the innovation (5 g kg^{-1}); therefore at location (x, y, z) the model is being nudged towards -2 g kg^{-1} . Nudging towards negative water vapor mixing ratios can result in rapid drying of the model and can lead to a complete removal of water vapor at that location.

To address this issue, when *obs_scl_neg_qv_innov* is set to “1”, the water vapor mixing ratio innovation is negative, and the model water vapor mixing ratio is smaller at the location the innovation is being applied than at the location the innovation was calculated, i.e.:

$$\{[q_o(i) - q_m(x_i, y_i, z_i, t)] < 0\} \text{ and } \{q_m(x, y, z, t) < q_m(x_i, y_i, z_i, t)\}$$

Then the nudging equation (1) is adjusted to be:

$$\frac{\partial q\mu}{\partial t}(x, y, z, t) = F_q(x, y, z, t) + \mu G_q \frac{\sum_{i=1}^N W_q^2(i, x, y, z, t) \max\{[q_o(i) - q_m(x_i, y_i, z_i, t)] - q_m(x, y, z, t)\}}{\sum_{i=1}^N W_q(i, x, y, z, t)} \quad (2)$$

which prevents nudging towards a negative value of water vapor mixing ratio. This fix ameliorates over-drying due to nudging. Applying this equation to the example discussed previously, the portion of the numerator of equation (2) with the maximum function is:

$$\begin{aligned} & \max\{[q_o(i) - q_m(x_i, y_i, z_i, t)] - q_m(x, y, z, t)\} \\ &= \max\{[5 \text{ g kg}^{-1} - 10 \text{ g kg}^{-1}], -3 \text{ g kg}^{-1}\} \\ &= \max\{-5 \text{ g kg}^{-1}, -3 \text{ g kg}^{-1}\} \\ &= -3 \text{ g kg}^{-1} \end{aligned}$$

Therefore, when *obs_scl_neg_qv_innov* is “1” the model will nudge towards 0 g kg⁻¹ (since the model water vapor mixing is 3 g kg⁻¹ at the location being nudged and the adjusted innovation is -3 g kg⁻¹) whereas when *obs_scl_neg_qv_innov* is “0” the model will nudge towards -2 g kg⁻¹. The use of *obs_scl_neg_qv_innov* set to “1” can thus mitigate issues with over-drying due to nudging. A description of this methodology can be found in Reen et al. (2013) and Reen et al. (2015).

As a side note, the user should be aware that assigning all surface observations to the lowest model level may result in observation nudging not performing as desired when one is using sufficiently high vertical resolution near the surface. In cases where surface observations are closer vertically to a model level above the lowest model level rather than being closest to the lowest model level one may want to calculate the innovation at whatever model level is closest to the observation, rather than calculating the innovation at the lowest model level. For example, in Gaudet et al. (2009) the lowest model layer was 4 m thick and so the lowest level with prognostic values of temperature, wind, and moisture was at 2 m above ground level (AGL) (note that these are not the diagnostic 2 m AGL temperature and moisture values WRF calculates in all WRF simulations but values from an actual prognostic model level, namely the lowest unstaggered or half level). By default, the observation nudging will compare surface observations of temperature, moisture, and wind to the lowest (unstaggered/half) model level, here 2 m AGL, to calculate innovations, and apply these innovations at this level (and then at other levels as specified by the user). For temperature and wind, before calculating the innovation the model will attempt to adjust the observed 2 m AGL temperature and the observed 10 m AGL wind to the lowest prognostic model level (here 2 m AGL) using the relationship between the model value at the lowest prognostic model level and the model surface diagnostic values (2 m for temperature and 10 m for wind). However, in Gaudet et al. (2009), since the first prognostic level with temperature and moisture was at 2 m, and the third level with wind was at approximately 10 m AGL, they avoid the adjustments to 2 and 10-m AGL, and calculate the innovations directly by comparing the lowest

model level temperature and moisture to the observation and the third model level wind to the observation. The standard version of WRF assumes all surface observations should be assigned to the lowest model level and the user should be aware of this if using unstaggered model levels below 10 m AGL.

2.3 Weighting Functions

Once the innovations are calculated, for each grid point it must be determined which observations will influence that grid point and determine all of the weights used in the application of each observation. We will discuss the overall weighting function (G in equation 1), followed by the components of the weighting functions included in W in equation 1: temporal weighting functions, vertical weighting functions, and horizontal weighting functions.

2.3.1 Overall Weighting Function

The overall weighting function (G in equation 1) is dependent only on the variable being observation nudged (*obs_coef_wind*, *obs_coef_temp*, and *obs_coef_mois*) and does not change in space nor time. As noted earlier, the inverse of G is the e-folding time of the difference between the model and the observation assuming the physical tendency terms are zero. Figure 1 illustrates the decrease in model error with time assuming the physical tendency terms are zero when nudging with various nudging weights ranging from $1e^{-4} \text{ s}^{-1}$ to $128e^{-4} \text{ s}^{-1}$. Although the stronger nudging weights clearly result in the model much more quickly approaching the observed value, with higher nudging weights the physical tendency terms will not be able to maintain a meteorologically consistent solution, since they may be overwhelmed by the tendency term introduced by the observation nudging.

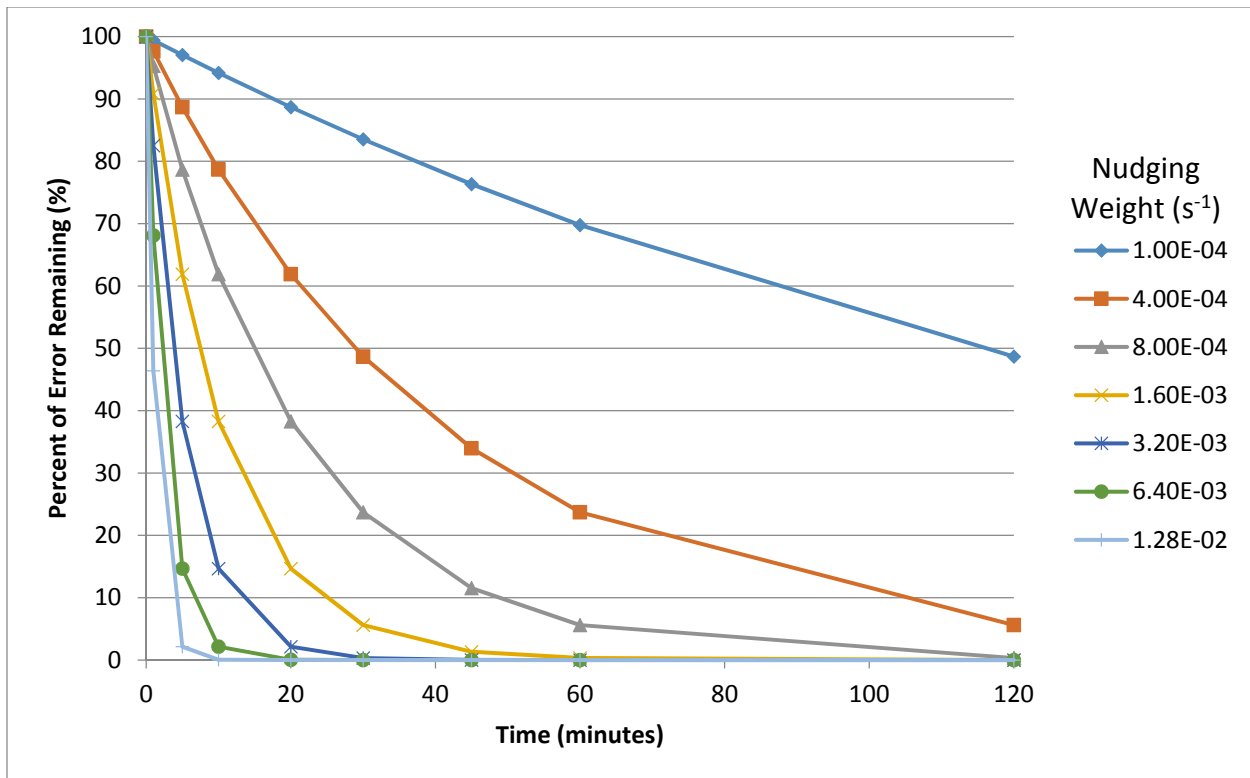


Figure 1. Decrease of model error with time when nudging with various nudging weights and assuming that all of the physical tendency terms are zero.

2.3.2 Temporal Weighting

The user must determine the time period over which observation nudging is desired. One can either apply observation nudging for the entire simulation or apply it only for a specified time period at the beginning of the simulation. If one is applying observation nudging for the entire simulation (i.e., dynamic analysis) one should set *fdda_end* such that it falls after the end of the simulation. If one is applying observation nudging only for a pre-forecast at the beginning of the model integration (i.e., dynamic initialization) one should set *fdda_end* to be the end of the period where observation nudging is applied (including any rampdown period) and *obs_idynin* should also be set to 1 since this is a dynamic initialization simulation. At the end of the dynamic initialization time period, a rampdown period where observation nudging decreases from 1 to 0 can be used to decrease any noise that may occur if nudging was suddenly turned off. Assuming a rampdown period of 1 hour, if observations valid during the rampdown period should be used for observation nudging then set *obs_dtramp* to -60 , whereas if the observations valid during the rampdown period should not be used for observation then set *obs_dtramp* to 60 . The rampdown period ends at *fdda_end* and thus *fdda_end* should be set to be the end of the rampdown period. Note that the switch *fdda_start* does not appear to have any effect.

The temporal weighting function determines the time period relative to the valid time of the individual observation where the observation is applied and how the strength of this application varies with time. The user specifies the length of half of the time window in *obs_twindo*. Given an observation at time t_o the temporal weight linearly increases from zero at $t_o - \text{obs_twindo}$ to one at $t_o - \text{obs_twindo}/2$, stays constant until $t_o + \text{obs_twindo}/2$, and then linearly decreases to zero at $t_o + \text{obs_twindo}$ (Figure 2). Note that for surface observations the time window is multiplied by

obs_sfcfact, to account for the more rapid changes in meteorological conditions at the surface. Since the observations are applied even after the valid time of the observation, when preparing observations for nudging one will likely want to include observations starting *obs_twindo* prior to the start of the observation nudging period.

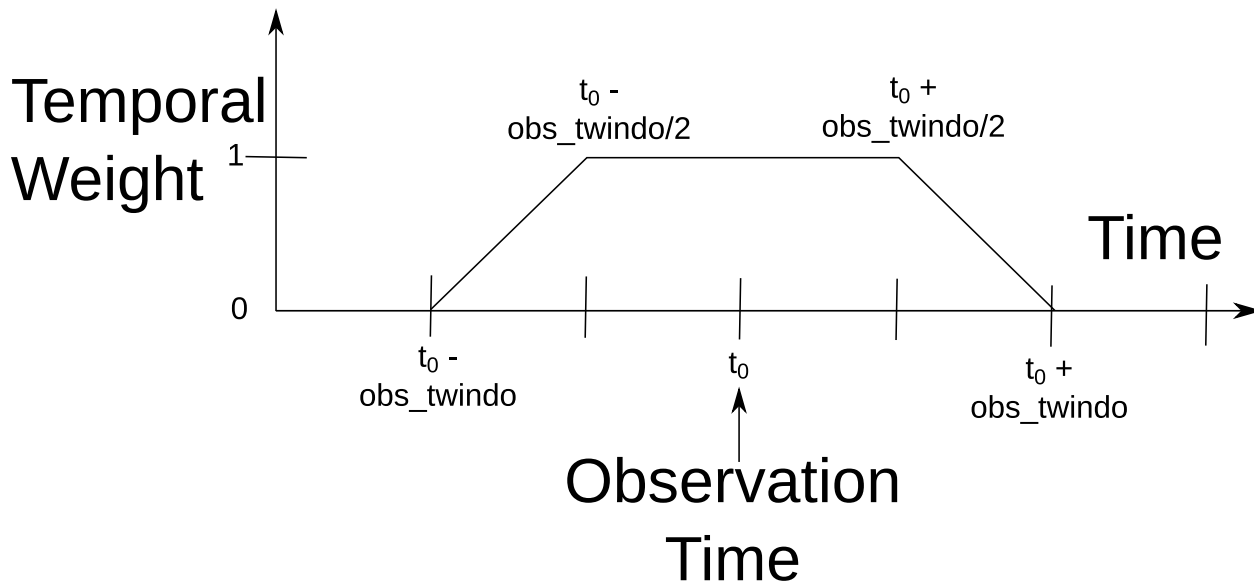


Figure 2. Temporal weighting function for an observation at time t_0 given based on the WRF namelist setting *obs_twindo*.

2.3.3 Vertical Weighting

The vertical weighting functions vary by the type of observation: surface observation, multi-level observation, or single-level above-surface observation. Note that the switch *obs_rinsig* is not used by WRF for any type of observation. In addition to the observation-type specific settings, each variable type also has a switch that can prevent observation nudging from being applied within the planetary boundary layer (PBL), which is also known as the atmospheric boundary layer (*obs_no_pbl_nudge_VAR* where VAR is uv, t, or q). One reason a user may want to invoke this option is if the user believes that observations are not sufficiently representative of the shorter temporal scales and smaller spatial scales within the PBL (compared to above the PBL).

For surface observations the user can choose from the vertical functions originally in WRF (*obs_sfc_scheme_vert* = 1) or the functions more recently introduced (*obs_sfc_scheme_vert* = 0).

The vertical functions originally in WRF (*obs_sfc_scheme_vert* = 1) attempt to spread the influence of innovations calculated from surface observations throughout the PBL with decreasing weight with height. It uses a linear decrease in weight with vertical level number from the surface to the model level where the PBL top occurs with two exceptions. If the PBL top is below the 3rd model level or above the 25th model level, then the vertical spreading will be applied through the 3rd model level above the surface instead of to the PBL top. Note that the vertical placement of the 3rd and 25th model levels will vary significantly depending on the placement of vertical levels for a given case.

The newer vertical functions for spreading surface observations (*obs_sfc_scheme_vert* = 0) are much more flexible. These functions allow the user to specify a layer in which nudging is applied at full strength and then a layer in which the nudging is linearly decreased to zero. This specification can be done in height relative to the ground or relative to the PBL top and can vary by variable and by convective regime. The inclusion of PBL-relative depth in the observation nudging vertical weights is based on the mixing that occurs within the PBL that can result in vertical error correlations within the PBL. If the user specifies that the vertical spreading of the surface innovation should be dependent on the PBL top, then the effectiveness of the nudging will be dependent on how well the model predicts the PBL and how well the PBL scheme diagnoses the top of the PBL. The three convective regimes currently recognized by the observation nudging are regimes 1, 2, and 4. The definition of convective regime varies within WRF (by surface layer scheme), but in general regime 1 is meant to represent stable conditions, regime 2 represents damped mechanical turbulent conditions, and regime 4 represents free convection. Regime 3 represents forced convection. Some PBL parameterizations define regimes, but for those that do not calculate regimes, the observation nudging code defines the regimes quite simplistically: If the PBL top is above the first half level then regime 4 is diagnosed, otherwise regime 1 is diagnosed. There is room for future improvement in the diagnosis of these regimes for use in observation nudging. If the observation nudging code encounters a grid point with regime 3 diagnosed, it will treat the grid point as being in regime 4.

For each regime and nudged variable, the user specifies the depth through which full weighting should be applied and the depth of the layer above this where weighting should linearly decrease with height to zero. The layer with full weighting is specified by switches of the form *obs_nudgezfullr#_VAR* where # is convective regime and VAR is the variable (uv, t, or q), and similarly the layer where nudging weight decreases linearly with height is specified as *obs_nudgezrampr#_VAR*. For the two preceding variables (*obs_nudgezfullr#_VAR* and *obs_nudgezrampr#_VAR*), if the value provided is positive then the values are in terms of the thickness of the layer in terms of meters. If the quantity provided is negative, 5000 will be added (to cancel the “flag” value of -5000) and the absolute value of the remainder will be considered meters relative to the PBL top. For example, specifying -5010 refers to 10 m above the top of the PBL. The minimum depth (in meters) of the full weight and ramping layers over all variables and regimes is specified by *obs_nudgezfullmin* and *obs_nudgezramprmin*. The maximum depth through which non-zero nudging weights can be applied is specified by *obs_nudgezmax* (in meters). The minimum and maximum depths can be used to prevent vertical weighting functions based on PBL-relative height from acting in an undesired manner when the PBL is very shallow or very deep. The switches used for this vertical weighting function option (*obs_sfc_scheme_vert* = 0) provide significant flexibility in setting the vertical weighting function of surface observations. Due to the complexity of these switches some examples will be given to illustrate their use.

Figure 3 illustrates four configurations for the vertical spreading of surface observations; these configurations are shown to demonstrate the how to use the switches and are not meant to suggest that the specific configurations demonstrated should be used. Reen et al. (2010) investigated the observation nudging vertical weighting function with MM5, and this research is also applicable to WRF and so can be referred to by the user to provide some background on this issue. The first example in Figure 3 shows (*obs_nudgezfullr#_VAR*=50, *obs_nudgezrampr#_VAR*=50) which

results in full weighting through the lowest 50 m above the ground and then a linear decrease in weight to zero at 100 m AGL. This is the default configuration for convective regimes 1 and 2 for all variables. The second option shows (*obs_nudgezfullr#_VAR=0*, *obs_nudgezrampr#_VAR=100*), which simply linearly decreases weighting with height through the lowest 100 m. It is important to remember that this vertical spreading function is applied at discrete model vertical levels. So if you use this vertical weighting and your first vertical prognostic level with moisture is at 40 m AGL, the maximum weighting for any vertical level will be 0.6 (the variables that are nudged are on the model unstaggered levels or half-levels and the lowest such level is not at the surface). The third option shows the convective regime 4 default setup (*obs_nudgezfullr#_VAR=-5000*, *obs_nudgezrampr#_VAR=50*) which is equivalent to (*obs_nudgezfullr#_VAR=-5000*, *obs_nudgezrampr#_VAR=-5050*). This option involves full nudging through the top of the PBL (represented by -5000) and then a 50 m rampdown or equivalently ramping down to zero 50 m above the top of the PBL (represented by -5050). The fourth option (*obs_nudgezfullr#_VAR=0*, *obs_nudgezrampr#_VAR=-5000*) linearly ramps down with height to zero at the top of the PBL. This is similar to the old WRF vertical weighting function (*obs_sfc_scheme_vert = 1*), except that here the weighting linearly ramps down with height whereas the original scheme linearly ramped with vertical level, and here we do not have the limitations of the 25th level and the 3rd level (although the new setup allows for similar limitation by height). The final example in Figure 3 shows (*obs_nudgezfullr#_VAR=-4950*, *obs_nudgezrampr#_VAR=0*) which results in full weighting up until 50 m below the top of PBL (represented by -4950) and zero weighting above that.

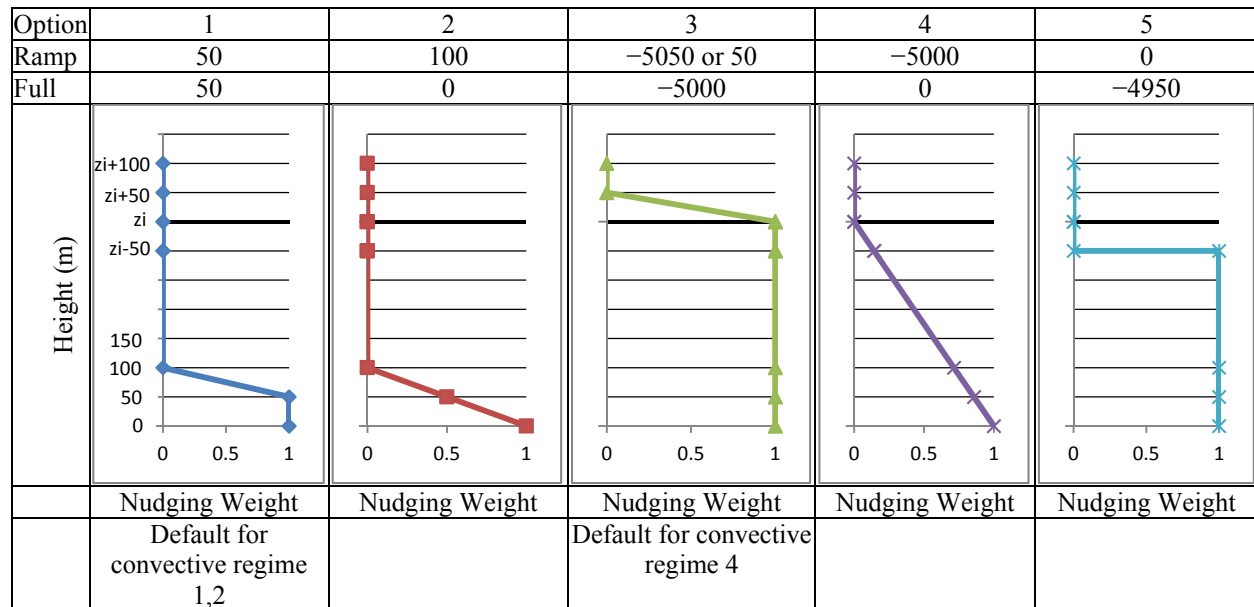


Figure 3. Examples of WRF vertical weighting functions for surface observations when *obs_sfc_scheme_vert=0*. The y-axis is in terms of meters above ground level where *zi* is the PBL depth. The switch settings that result in these shapes are indicated by the second and third rows: “Ramp” refers to namelist options *obs_nudgezrampr#_VAR* and “Full” to *obs_nudgezfullr#_VAR* where # is the convective regime and VAR is the variable.

For multiple-level observations, the innovations calculated at the observation location are vertically interpolated in pressure to the level at which they are being applied. The interpolation is done in log-pressure space.

This interpolation will not be done to model levels with pressures that are within a vertical gap in the observation greater than or equal to *obs_max_sndng_gap*. For example, consider a case with a sounding that has observations at 950, 850, 600, and 450 hPa, and the user has set *obs_max_sndng_gap* to 20 cb (200 hPa). The model will be horizontally and vertically interpolated to the observations and innovations will be calculated. The innovations at 950, 850, 600, and 450 hPa can then be vertically interpolated to model levels between 950 and 450 hPa. However, the gap between 850 and 600 hPa is 25 cb (250 hPa), which is greater than *obs_max_sndng_gap*, so no interpolation can be made to model levels with pressures between these two pressures. Therefore the innovations at 950 and 850 hPa will be interpolated to any model levels with pressures between 950 to 850 hPa where other weighting factors do not preclude nudging, and similarly between 600 and 450 hPa.

For single-level above-surface observations, the innovation is applied for model levels with pressures within 75 hPa of the observation. The vertical weighting function linearly decreases with pressure from one at the pressure of the observation to zero 75 hPa above or below the observation. Additionally, if the innovation is calculated below the model-diagnosed PBL top (at the observation location) then it is only applied below the model-diagnosed PBL top at the model point at which the innovation is being applied; similarly, if the innovation is calculated above the model-diagnosed PBL top then it is only applied above the model-diagnosed PBL top. This feature is based on the assumption that errors within the PBL are not necessarily correlated with errors above the PBL. For example, given a single-level above-surface observation at 850 hPa where the model PBL top is at 900 hPa, the innovation will be applied between 900 and 775 hPa. Since the innovation is calculated above the PBL top it is not permitted to be applied in the 925 to 900 hPa range at this point since that pressure range is within the PBL, even though the pressure range falls within 75 hPa of the observations. Similarly, if the innovation is applied to another model grid point where the model PBL top is at 930 hPa, the innovation will not have an effect on the model solution because the entire extent of the weighting function (925–775 hPa) will be above the PBL top whereas the innovation was calculated within the PBL. The restriction against single-level above-surface observations within (above) the PBL being spread above (within) the PBL can avoid possible difficulties when the elevation or PBL depth at the model location where the innovation is applied is different from where the innovation was calculated.

2.3.4 Horizontal Weighting

The horizontal spreading of innovations depends on the observation type. Horizontal spreading of surface observations is along the lowest model half-level (i.e., unstaggered level) whereas for above-surface observations the spreading is in pressure space. If surface observations were spread in pressure-space, then a surface observation might be applied 200 m above the surface at a location with a lower terrain than the observation location. Similarly at a location with a higher terrain than the observation location, the surface observation would not be able to be applied at all since the surface pressure is less than the pressure of the observation. Note that surface observations that are part of a multi-level observation (i.e., a rawinsonde) are not treated as surface observations.

There are several components involved in the horizontal spreading of innovations. For all observations, the horizontal spreading is dependent on *obs_rinxy*, the horizontal radius of influence. This value is multiplied by *obs_sfcfacr* for surface observations since there will likely

be more spatial variation at the surface than above the surface and thus the error correlation lengths will likely be smaller at the surface. For above-surface observations (single-level or multi-level), the radius of influence specified by *obs_rinxy* is multiplied by a term that linearly increases with decreasing pressure from one at the surface to two at 500 hPa, and then remains two for all observations above 500 hPa. This increase in radius of influence with height is based on increased error correlation length scales at higher altitudes.

The horizontal weighting function is calculated as: $W_{xy} = (RIN^2 - DIST^2)/(RIN^2 + DIST^2)$ where *RIN* is the effective radius of influence for this observation (based on *obs_rinxy* and the vertical placement of the observation) and *DIST* is the distance between the observation and current model point where the innovation is being applied. Close to the observation, this function decreases less quickly with increasing distance than a linear decrease (e.g., Figure 4). This allows the observations to be more fully used in model locations close to the observations.

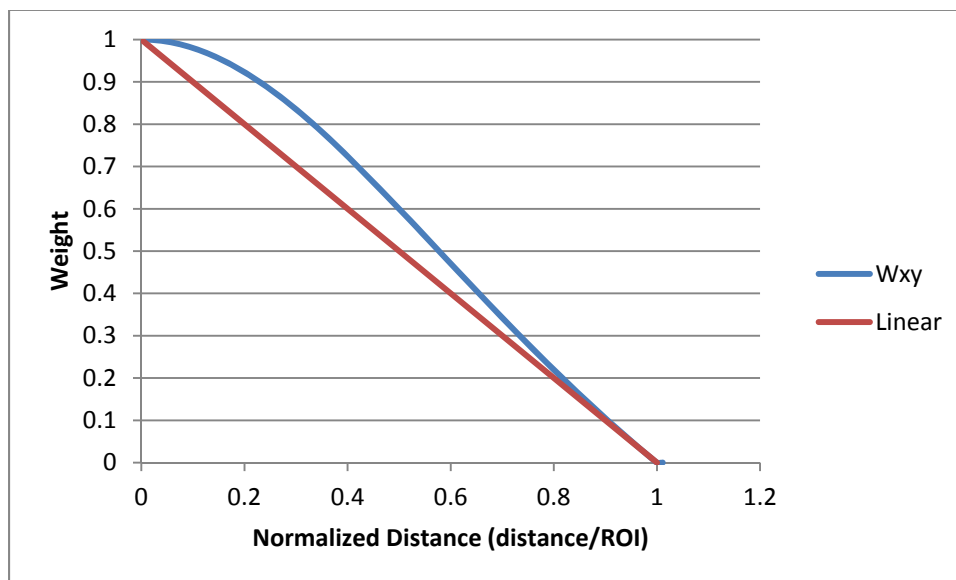


Figure 4. Decrease of horizontal weighting function with distance (relative to the radius of influence [ROI]) for the WRF function (*Wxy*) and a simple linear decrease (*Linear*).

For surface observations, an additional factor limits the application of innovations at locations whose surface pressure differs from the surface pressure at the observation location. This accounts for the decrease in error correlation length scales expected due to differences in surface elevation. For example, if surface observation A is at 500 m MSL but surface observation B is at 1000 m MSL, then the errors at those two sites are likely to be less correlated than between surface observation A and a surface observation C at 550 m MSL. In this example site A and C may be in a valley while site B is on a ridgetop and may experience different meteorological conditions. Two methods are provided to restrict nudging of surface observations at locations with a model surface pressure different than the model surface pressure at the observation location: *obs_sfc_scheme_horiz* = 0 is the scheme that was the sole scheme in early versions of WRF and *obs_sfc_scheme_horiz* = 1 is the scheme used in MM5 that is now available in WRF.

Both methods use a maximum allowable model pressure difference (obs_dpsmx) between the innovation calculation location and the innovation application location, but they implement this differently. The inverse of the maximum allowable model pressure difference is multiplied by the actual model pressure difference to determine the proportion of the maximum allowable difference that is occurring for a given case. The original WRF scheme multiplies the horizontal weighting function by this term, whereas the MM5-based method uses this term to increase the *effective* distance of the observation from the innovation application location, thereby potentially decreasing the radius of influence in areas with terrain complexity. Specifically, for the MM5-based method an effective distance ($distance_e$) is defined as:

$$distance_e = distance + obs_rinxy * \frac{|p_{m@ob} - p_m|}{obs_dpsmx}$$

Where $distance$ is the distance between the location of the observation and the model location where the innovation is being applied, $p_{m@ob}$ is the model surface pressure at the observation location and p_m is the model surface pressure at the innovation application location. Then the horizontal weighting function is:

$$w_{xy} = \frac{(obs_rinxy * obs_sfcfacr)^2 - distance_e^2}{(obs_rinxy * obs_sfcfacr)^2 + distance_e^2}$$

For the original WRF method:

$$w_{xy} = \left[1 - \min \left(1, \frac{|p_{m@ob} - p_m|}{obs_dpsmx} \right) \right] * \frac{(obs_rinxy * obs_sfcfacr)^2 - distance^2}{(obs_rinxy * obs_sfcfacr)^2 + distance^2}$$

Since the original WRF method is multiplying the horizontal weighting factor by another factor, for that method, as long as the pressure difference does not exceed obs_dpsmx , the adjustment cannot result in the horizontal weighting function becoming zero for a given point. In other words, in contrast to the MM5-based method, unless the pressure difference is greater than obs_dpsmx this method does not decrease the radius of influence in areas with terrain complexity, but instead decreases the weighting. Figure 5 illustrates the differences between these two methodologies for two cases: in Figure 5a pressure decreases with distance from the observation (consistent with terrain height gradually increasing with distance from the observation) and in Figure 5b pressure alternates between increasing and decreasing with distance (consistent with a series of ridges and valleys). This figure assumes $obs_rinxy = 100$ km, $obs_sfcfacr = 0.75$, and $obs_dpsmx = 75$ hPa. Without accounting for surface pressure differences, in Figure 5a the nudging weight decreases to zero 75 km from the observation ($obs_rinxy * obs_sfcfacr$). The original WRF methodology also decreases to zero 75 km from the observation but this nudging weight is less than if the surface pressure difference was not accounted for. In the MM5-based method the nudging weight decreases to zero around 45 km. In Figure 5b the original WRF methodology still decreases to zero 75 km from the observation but the MM5-based methodology decreases to zero about 40 km from the observation.

3 Preparing Observations

The observation nudging code reads observations from files named OBS_DOMAIN?01 where ? is the domain number. Note that only one OBS_DOMAIN file will be read per domain and that identical OBS_DOMAIN files may be used for each domain since WRF will not utilize observations outside the model domain. The OBS_DOMAIN format is described in Appendix 0.

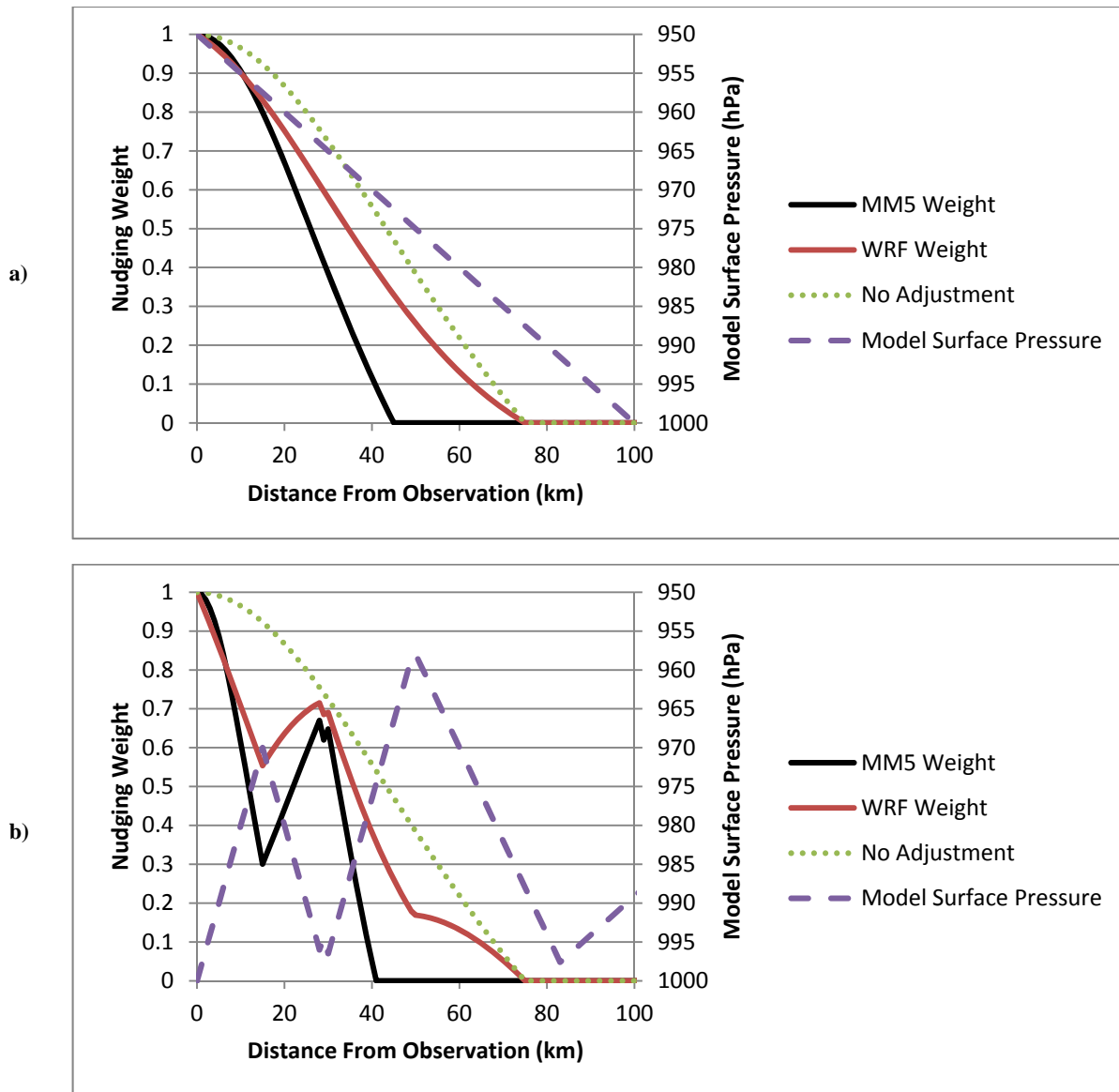


Figure 5. Illustration of the nudging weights that result from the term designed to limit the spreading of observations from a location with a given elevation to other locations with elevations that differ from those at the observation location for two cases (a, b). Shown are the MM5-based method (`obs_sfc_scheme_horiz = 1`), the original WRF method (`obs_sfc_scheme_horiz = 0`), and a line showing the result if neither adjustment is used. The variation of model surface pressure with distance from the observation is also shown. This figure assumes `obs_rinxy = 100` (km), `obs_sfcfacr = 0.75`, and `obs_dpmsx = 75` (hPa).

It is crucial that observations be quality controlled. If observations are not quality controlled, “bad” observations are more likely to be present in the OBS_DOMAIN file. Nudging toward bad observations could limit the benefit provided by observation nudging and may cause observation nudging to degrade the model solution, potentially significantly. NCAR distributes Obsgrid (http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/users_guide_chap7.htm), which is software that can be used to quality control observations for use in WRF. Obsgrid in this document refers to both Obsgrid V3.7 and V3.8 unless specifically noted. Obsgrid performs gross error checks to find entries which are clearly in error based on the normal range of a quantity. A buddy check is also performed by Obsgrid; this checks an observation against the surrounding

observations (“buddies”) in order to determine which observations are not consistent with the other observations. Obsgrid also checks observations against a background field which consists of output from whatever model the user chooses to initialize the model initial conditions and boundary conditions. The observation format required for ingestion by Obsgrid is termed “little_r” (or little-r or littler); this format is described in the Obsgrid documentation (http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/users_guide_chap7.htm#format). Since observations are utilized by observation nudging *obs_twindo* before the valid time, the user will likely want to include observations at least *obs_twindo* prior to the start of nudging.

There are some important details about formatting observations into little_r format that the user should be aware of.

- The *is_sound* switch should only be set to FALSE if the observation is a surface observation. If an observation has a single-level but is above the surface then *is_sound* should be set to TRUE even though the observation is not necessarily from a radiosonde.
- Although the little_r format includes u-component winds, v-component winds, and relative humidity, these fields are not ingested by Obsgrid. The user should set wind information via the speed and direction fields, and moisture via the dewpoint field.
- Obsgrid expects incoming winds to be earth-relative, but will output grid-relative winds in the OBS_DOMAIN file.
- If observations do not include pressure, but do include height, Obsgrid will attempt to calculate the pressure using the height from the observation and the pressure and height fields from the non-surface levels in the first-guess field. Prior to V3.8, if this fails then it uses the height of the observation along with the standard atmosphere to calculate the pressure of the observation. Starting in V3.8, while standard atmosphere is still used for observations with a pressure lower than the top first-guess level, a different methodology is used for observations with pressures higher than the lowest non-surface first-guess level:
 - If the height of the observation indicates that it is below the lowest non-surface first-guess field but at or above sea level, then interpolate between the first-guess sea level pressure and the lowest non-surface first-guess level's pressure
 - If the height of the observation indicates that it is below the lowest non-surface first-guess field and below sea level then use the first-guess sea-level pressure to find the temperature at the height of the observation
 - Use the temperature reported in the observation if available, if not available...
 - Use the temperature of the first guess field interpolated to this point, if not available...
 - Use standard atmosphere
 - Find the temperature at sea level using the temperature we found at the height of the observation and use the standard atmosphere lapse rate. With these we can now calculate the pressure at the observation.

Users who desire to have WRF assign pressure to observations based on the WRF fields should be especially aware that Obsgrid will likely have already filled in the pressure field prior to the file being ingested by WRF.

- Quality control of single-level above surface observations can be challenging:

- Prior to Obsgrid V3.8, single-level above-surface observations will most likely be rejected by Obsgrid unless the platform field is “FM-97 AIREP” or “FM-88 SATOB”. If it is one of these two values of platform, then Obsgrid will attempt to adjust the observation to a level where a first-guess field is available for the background check. In this adjustment certain observation fields are dropped or marked as bad. The rejection of most single-level above-surface observations occurs because unless the observation has a pressure exactly matching one of the first-guess fields (e.g., from the Global Forecast System), no background check can occur in Obsgrid prior to V3.8.
- Starting in Obsgrid V3.8, the user can choose an alternate way to quality control single-level above-surface observations against the first-guess field by setting *use_p_tolerance_one_level* to TRUE. This allows quality control of single-level above-surface observations against a pressure level in the first guess as long as the pressure difference between the observation and the first-guess level is within a user prescribed tolerance. This prevents the observation from being adjusted to a different pressure level and prevents certain observation elements from being discarded as occurs with *use_p_tolerance_one_level* to FALSE. The user will likely need to provide Obsgrid with first-guess fields vertically interpolated to additional pressure levels to ensure first-guess levels are closely enough spaced vertically such that the pressure tolerance can be set to reasonable values.
- Some single-level above-surface observations may be marked as duplicate in cases where the user would not expect this to occur. Obsgrid tries to prevent duplicate observations by checking for observations whose time and horizontal location are sufficiently similar. However, for an observation like an ascending aircraft, the time and horizontal location of two consecutive observations may be very close, while the vertical placement is different enough that the user would prefer to keep both observations, rather than marking one as duplicate. In this case, the user may want to adjust the distance threshold for determining duplicates in Obsgrid in this case (see function *loc_eq*).

Regarding obtaining observations in *little_r* format, one option is to obtain the observations from the Meteorological Assimilation Data Ingest System (MADIS; madis.noaa.gov) and then convert them into *little_r* format by using the MADIS2LITTLER tool (<http://www2.mmm.ucar.edu/wrf/users/wrfda/download/madis.html>). However, note that MADIS2LITTLER may set the *is_sound* switch to “false” for single-level above-surface observations, whereas Obsgrid/WRF will expect this to be set to “true”. To fix this, modify MADIS2LITTLER to mark such observations as soundings (in *module_output.F*, subroutine *write_littler_onelvl* must be modified to set *is_sound = .TRUE.*).

4 Diagnostic Prints

A variety of diagnostic prints are available that are associated with observation nudging and function to improve the user’s understanding of how the code is functioning. The presence of some diagnostic output, as well as the quantity of the diagnostic output, is dependent on how various switches are set. The diagnostic prints will be placed in *rsl.out.####* or *rsl.error.####* (or potentially both), where *####* is a four digit number ranging from 0000 to the largest number necessary for each processor core to have an individual output file.

If all of the diagnostic print switches are effectively turned off, diagnostic prints will still occur that describe the observation nudging settings utilized for this run. This includes details on the temporal and spatial weighting functions used. The diagnostic output begins with a statement like “OBSERVATION NUDGING IS ACTIVATED FOR MESH 1” and ends with a statement like “*** END SETUP DESCRIPTION FOR SURFACE OBS NUDGING ON MESH 1”. This diagnostic output can be used to help determine whether the namelist settings you provided to WRF result in the desired observation nudging configuration.

The interval between diagnostic prints is purportedly controlled by *obs_npf*, however it does not appear that this switch actually affects the diagnostic output. Note that if the *obs_npf* switch were implemented the current coding would result in the same issues described for *obs_ionf* in section 2.2, wherein *obs_npf* would not necessarily be in terms of coarse grid domain timesteps.

If *obs_ipf_in4dob* is activated certain details of the nudging configuration are listed at the beginning of the simulation (e.g., nudging strength, spatial and temporal weighting function parameters) as well as details about reading in observations for data assimilation. The time window of observations being read in is listed (in model-relative time) as well as the number of observations currently available for observation nudging. If an observation is read in with a platform name not recognized by the observation nudging code, it will print an error that includes the message “unknown ob of type XXX” where XXX is the platform. However, the observation will still be used. The code indicates that the platform type may be used in the future to create platform-specific nudging weight adjustments, but this is not done at this time. The platform type is currently only used to do a small number of adjustments for specific platforms (e.g., moisture observations with a type of “SHIP” are marked as missing if the relative humidity is less than 70%).

If *obs_ipf_error* is activated then diagnostic output is created regarding observations ingested for data assimilation. The maximum number of observations to print information regarding is specified by *obs_prt_max*, and the interval (in terms of observation number) between printed observations is specified by *obs_prt_freq*. For example, if the user specifies *obs_ipf_error* = .true., *obs_prt_max* = 10, and *obs_prt_freq* = 2, the output could be:

```
+++++CALL ERROB AT KTAU =      0 AND INEST =  1:  NSTA =    55 +++++
REPORTING OBS MASS-PT LOCS FOR NEST    1 AT XTIME=      0.0 MINUTES
FREQ=   2, MAX=   10 LOCS, NEWLY READ OBS ONLY, -999 => OBS OFF PROC

OBS#      I          J          K      OBS LAT  OBS LON  XLAT(I,J)  XLONG(I,J)  TIME(hrs)  OBS
STATION ID
  1  37.333 148.827   1.000   37.760 -122.220   37.760   -122.220    0.00    KOAK
  3  37.333 148.827   5.306   37.760 -122.220   37.760   -122.220    0.00    KOAK
  5  37.333 148.827   7.911   37.760 -122.220   37.760   -122.220    0.00    KOAK
  7  37.333 148.827   9.879   37.760 -122.220   37.760   -122.220    0.00    KOAK
  9  37.333 148.827  11.506   37.760 -122.220   37.760   -122.220    0.00    KOAK
 11  37.333 148.827  18.683   37.760 -122.220   37.760   -122.220    0.00    KOAK
 13  37.333 148.827  19.518   37.760 -122.220   37.760   -122.220    0.00    KOAK
 15  37.333 148.827  20.329   37.760 -122.220   37.760   -122.220    0.00    KOAK
 17  37.333 148.827  27.005   37.760 -122.220   37.760   -122.220    0.00    KOAK
 19  37.333 148.827  27.745   37.760 -122.220   37.760   -122.220    0.00    KOAK
```

This lists the index number WRF is using to refer to the observation, the latitude and longitude of the observations (OBS LAT and OBS LON), where this places the observation in the domain in terms of grid points (i, j, and k), the latitude and longitude of the model grid points specified by (i,j,k), the model-relative time of the observations, and the name of the observation. Note that each time the print is done it is only printing from the observations just read in. As specified by *obs_prt_max* = 10 and *obs_prt_freq* = 2, diagnostic information about every other observation is printed and total of 10 observations are printed. As noted in the printout, if an observation is not on the part of the domain that is being processed by the current processor core (the core whose rsl.out or rsl.error file you are currently viewing) then the model latitude and longitude fields (XLAT, XLONG) on that line will be denoted as -999, since these values are not known by that processor.

If *obs_ipf_nudob* is activated the rsl.out.* file will contain prints similar to the following:

```
OBS NUDGING is requested on a total of 1 domain(s).
DYNINOBS: IN,KTAU,XTIME,FDAEND,DTRAMP,DTR,TCONST,TFACI: 1 0 0.000000
360.00000 60.000000 60.000000 1.66666675E-02 1.0000000
OBS NUDGING FOR IN,J,KTAU,XTIME,IVAR,IPL: 1 10 0 0.00 3 3 rindx= 1.1
OBS NUDGING FOR IN,J,KTAU,XTIME,IVAR,IPL: 1 10 0 0.00 4 4 rindx= 1.1
```

The line beginning DYNINOBS informs the user about details of the rampdown at the end of the observation nudging (e.g., *obs_idynin*). This specific example is indicating that:

- Nest 1 is currently being processed (IN=1)
- The current timestep is number 0 (KTAU=0)
- It is now 0.0 minutes into the model integration (XTIME=0.0000000)
- Observation nudging will end at 360.0 minutes into the model integration (FDAEND=360.000000)
- The user chose a rampdown period of 60 minutes (DTRAMP=60.000000; *obs_dtramp*)
- The absolute value of the rampdown period is 60 minutes (DTR=60.000000; $|obs_dtramp|$)
- The inverse of the absolute value of the rampdown period is 1.67E-02 minutes (TCONST=1.66666675E-02; $1/|obs_dtramp|$).
- The value of the element of the temporal weighting factor that is accounting for the rampdown period is currently 1.0 (TFACI=1.0000000).

```
OBS NUDGING FOR IN,J,KTAU,XTIME,IVAR,IPL: 1 10 0 0.00 3 3 rindx= 1.1
```

The first line beginning “OBS NUDGING FOR” indicates that:

- Nest 1 is currently being processed (IN=1)
- The 10th west-east row from the southernmost edge of the domain is currently being processed (J=10) (this diagnostic printout is only printed for J=10)
- The current timestep number is 0 (KTAU=0)
- It is now 0.0 minutes into the model integration (XTIME=0.0000000)
- Temperature is the variable now being nudged (IVAR=3; IPL=3)
- The horizontal radius of influence in terms of grid lengths is 1.1 (RINDEX=1.1)

If *obs_ipf_init* is set to true, printouts regarding the initialization of observation nudging will be included in the rsl.out.* files.

5 Settings in WRF-ARW Observation Nudging

Settings which can be specified for each domain separately will have “(max_domains)” appended to their name. Settings with a single setting applicable to all domains will have “(1)” appended to their name.

Although the majority of settings for observation nudging are in the “fdda” portion of the namelist, there are two settings in the time_control portion of the namelist (*auxinput11_interval* and *auxinput11_end_h*). The section of the namelist in which each variable exists is noted by the text in brackets following each variable name.

obs_nudge_opt(max_domains) [fdda]

Determines whether observation nudging is turned on (=1) or turned off (=0) for each domain. There are separate switches of the form *obs_nudge_VAR* (where VAR is the variable to nudge) that must be set to determine whether observation nudging is turned on and off for each variable.

obs_nudge_wind(max_domains) [fdda], *obs_nudge_temp*(max_domains) [fdda],
obs_nudge_mois(max_domains) [fdda], *obs_nudge_pstr*(max_domains) [fdda]

Determines whether observation nudging based on the specified variable should be turned on (=1) or turned off (=0) for each domain. Note that surface pressure nudging is unavailable and so *obs_nudge_pstr* is not used.

max_obs [fdda]

Specifies the maximum number of observations that may be available for nudging at any given time. The model will stop if it finds more than *max_obs* observations to be applicable at a given time.

obs_ionf(max_domains) [fdda]

Interval (in units of coarse grid time steps) between recalculations of the innovations and checking if there are observations to be read in. A smaller value for *obs_ionf* makes the observation nudging more responsive to model changes, but also increases computational demands. Note that this switch does not necessarily function as desired on subdomains, as detailed in section 2.2.

auxinput11_interval(max_domains) [time_control]

Determines how frequently the model checks to see if it needs to read in additional observations for observation nudging. Note that this is not the interval at which it reads in the observation, but just the interval at which it checks to see if it is time to read in observations based on the namelist setting *obs_ionf*. It is safest to always set this to one, to ensure that it checks to see if it needs to read in observation at every time step. Also, it is best to set this switch (*auxinput11_interval*) based on timesteps and not one based on time (e.g., via *auxinput11_interval_s*). If this switch is not set to 1, your setting for *obs_ionf* may not have the effect that you intend.

auxinput11_end_h (max_domains) [time_control]

Determines when the code should stop checking to see if it needs to read in observations for observation nudging.

obs_coef_wind(max_domains) [fdda], *obs_coef_temp*(max_domains) [fdda],
obs_coef_mois(max_domains) [fdda], *obs_coef_pstr*(max_domains) [fdda]

The strength at which observation nudging is applied to the specified variable in units of inverse seconds (s^{-1}). This is the G term in equation 1. Note that surface pressure nudging is unavailable and so *obs_coef_pstr* is not used.

fdda_start(max_domains) [fdda]

The length of time (in minutes) after the start of the model run when observation nudging should start to be applied. This switch is not used by WRF and so nudging will always start at the beginning of the model run if it is activated.

fdda_end(max_domains) [fdda]

The length of time (in minutes) after the start of the model run when observation nudging should end. The switches *obs_idynin* and *obs_dtramp* effect the methodology used to end nudging. Note that any rampdown ends at *fdda_end*. This switch may not be effective unless *obs_idynin* is set to 1.

obs_idynin(1) [fdda]

Whether this is a dynamic initialization model run (1=yes, 0=no). This enables observation nudging to stop before the time at which the model ends, and should be enabled for all runs with observation nudging that ends prior to the model ending. This switch must be set to 1 in order to use *obs_dtramp* and may need to be set to 1 for *fdda_end* to work correctly.

obs_dtramp(1) [fdda]

The length (in minutes) of the rampdown period at the end of the dynamic initialization period where a factor linearly decreasing from 1 to 0 over the rampdown period is multiplied by the nudging terms. The goal of this factor is to allow the nudging term to be gradually decreased at the end of the pre-forecast to minimize noise that might be caused by turning observation nudging off abruptly. The rampdown period linearly decreases from 1 at $fdda_end - abs(obs_dtramp)$ to 0 at *fdda_end*. If *obs_dtramp* is ≤ 0 then the observations with a timestamp after *fdda_end* will not be used by the observation nudging (i.e., observations with timestamps during the rampdown period will be assimilated). If *obs_dtramp* is > 0 then the observations with timestamps after $fdda_end - obs_dtramp$ will not be used by the observation nudging (i.e., observations with timestamps during the rampdown period will not be assimilated). Note that *obs_idynin* must be set to 1 in order to use *obs_dtramp*.

Example 1: For a run starting at 1200 UTC, if the user sets *fdda_end* = 420, *obs_idynin* = 1, and *obs_dtramp* = 60 then the nudging will end at 1900 UTC. All observations up until 1800 UTC will be used by the nudging, and a rampdown factor that linearly decreases from 1 at 1800 UTC to 0 at 1900 UTC will be applied to all observations.

Example 2: For a run starting at 1200 UTC, if the user sets *fdda_end* = 420, *obs_idynin* = 1, and *obs_dtramp* = -60 then the nudging will end at 1900 UTC. All observations up until 1900 UTC will be used by the nudging, and a rampdown factor that linearly decreases from 1 at 1800 UTC to 0 at 1900 UTC will be applied to all observations.

obs_twindo (max_domains) [fdda]

Specifies half the length of the time period over which observations are applied (in hours). As illustrated in Figure 2, given an observation at time t_o the temporal weight (part of W in equation 1) linearly increases from zero at $t_o - \text{obs_twindo}$ to one at $t_o - \text{obs_twindo}/2$, stays constant until $t_o + \text{obs_twindo}/2$, and then linearly decreases to zero at $t_o + \text{obs_twindo}$. Note that for surface observations the time window is multiplied by obs_sfcfact .

obs_sfcfact(1) [fdda]

The length of the time period over which observations are applied (*obs_twindo*) is multiplied by this factor. This allows time windows to be shorter for observations at the surface where the temporal correlation length may be shorter.

obs_rinsig (1) [fdda]

Purportedly specifies the vertical radius of influence in terms of sigma. However, this setting is not used by WRF.

obs_no_pbl_nudge_uv (max_domains), *obs_no_pbl_nudge_t* (max_domains), and *obs_no_pbl_nudge_q* (max_domains) [fdda]

Specifies whether observation nudging should be excluded from within the PBL (1=exclude nudging from the PBL, 0=include nudging in the PBL). If observation nudging is excluded from within the PBL then surface observations are not utilized for nudging, single-level above-surface observations are only used for nudging if they are above the model PBL at the observation location, and the innovations from multi-level observations are only applied above the model PBL.

obs_sfc_scheme_vert (1) [fdda]

Specifies which class of vertical weighting functions will be used for nudging based on surface observations. If *obs_sfc_scheme_vert* = 1, the original WRF vertical weighting functions are used wherein the innovation is applied with linearly decreasing weight with model level throughout the model PBL, except for PBL tops below the 3rd model level or above the 25th model level will be treated as if the PBL top were at the 3rd model level. If *obs_sfc_scheme_vert* = 0, the newer, more flexible vertical weighting functions will be used as described in Section 2.3.3.

obs_nudgezfullr#_VAR (1) [fdda]

The layer above the surface where full nudging of surface observations should be applied for convective regime # (1, 2, or 4; regime 3 is treated as if it were 4) and variable VAR (uv, t, or q). If positive it refers to meters above ground level. If negative, then the value is transformed to meters relative to the PBL top by adding 5000 to remove the -5000 “flag”, and taking the absolute value of the remainder. See Section 2.3.3 for more detail. This setting is only relevant for the newer vertical weighting functions (*obs_sfc_scheme_vert* = 0).

obs_nudgezrampr#_VAR (1) [fdda]

The layer above the layer with full weighting (see *obs_nudgezfullr#_VAR*) wherein nudging of surface observations linearly decreases with height from one to zero for convective regime # (1, 2, or 4; regime 3 is treated as if it were 4) and variable VAR (uv, t, or q). If positive it refers to meters above ground level. If negative, then the value is transformed to meters relative to the PBL top by adding 5000 to remove the -5000 “flag”, and taking the absolute value of the remainder. See

Section 2.3.3 for more detail. This setting is only relevant for the newer vertical weighting functions (*obs_sfc_scheme_vert* = 0).

obs_nudgezfullmin (1) [fdda]

The minimum depth (in meters) for which full weighting of surface observations must be applied for all regimes and variables. See Section 2.3.3 for more detail. This setting is only relevant for the newer vertical weighting functions (*obs_sfc_scheme_vert* = 0).

obs_nudgezrampmin (1) [fdda]

The minimum depth (in meters) for which the linear rampdown of weighting with height of surface observations must be applied for all regimes and variables. See Section 2.3.3 for more detail. This setting is only relevant for the newer vertical weighting functions (*obs_sfc_scheme_vert* = 0).

obs_max_sndng_gap (1) [fdda]

Specifies the maximum gap (in cb) between pressure levels in a multi-level observation over which interpolation of the innovations is allowed.

obs_rinxy (max_domains) [fdda]

Specifies the maximum distance (in km) from an observation that its innovation can be applied. This value is multiplied by *obs_sfcfacr* for surface observations. For above-surface observations, this value is multiplied by a term that linearly increases with pressure from one at the surface to two at 500 hPa, and then remains two for all observations above 500 hPa.

obs_sfcfacr (1) [fdda]

Multiplied by *obs_rinxy* to find the maximum distance from a surface observation that its innovation can be applied.

obs_sfc_scheme_horiz (1) [fdda]

Specifies which methodology will be applied to limit the spreading of innovations from surface observations to locations with sufficiently differing terrain height. The original WRF scheme is specified by *obs_sfc_scheme_horiz*=0, while the MM5-based scheme is specified by *obs_sfc_scheme_horiz*=1. See Section 2.3.4 for details.

obs_dpsex (1) [fdda]

The maximum allowable surface pressure difference used by the scheme specified by *obs_sfc_scheme_horiz* to limit spreading of surface observations to locations with differing terrain height.

obs_npfi (1) [fdda]

Purportedly the interval (in coarse grid timesteps) between observation nudging diagnostic print outs. However, this switch does not appear to have any impact on diagnostic output. Also, if this switch were implemented, the current coding would result in the interval not necessarily being in terms of coarse grid domain timesteps due to the same issues as documented for *obs_ionfi* in Section 2.2.

obs_prt_max (1) [fdda]

Maximum number of observations included in diagnostic print outs.

obs_prt_freq (1) [fdda]

Specifies interval (in terms of number of observations) between observations included in diagnostic print outs.

obs_ipf_XXX (1) [fdda]

Specifies whether diagnostics should be printed for various observation-nudging related subroutines (this setting is a logical variable). The “XXX” stands for in4dob, errob, nudob, or init.

obs_scl_neg_qv_innov (1) [fdda]

Specifies whether negative water vapor mixing ratio observation nudging innovations should be scaled to prevent nudging towards a negative values of water vapor mixing ratio. To use this scaling set the parameter to “1”, to avoid this scaling set the parameter to “0”.

Appendix A: OBS_DOMAIN format and its ingestion by WRF

OBS_DOMAIN format is the format for observation files ingested by WRF observation nudging. It is described on NCAR's WRF website (http://www2.mmm.ucar.edu/wrf/users/wrfv3.1/How_to_run_obs_fdda.html).

WRF will read a single OBS_DOMAIN file OBS_DOMAIN?01 for each domain ?. If your method to create OBS_DOMAIN files creates multiple files per domain (as Obsgrid does) then the OBS_DOMAIN files must be concatenated before use. However, the observations within OBS_DOMAIN files must be in chronological order.

Each observation contains a 4-line header (shading of the variable name in the table below indicates that the variable is not used by WRF):

OBS_DOMAIN Observation Header Format				
Line #	Name	Variable Type	Format	Description (units)
1	SPACES		1x	Blank spaces
1	<i>date char</i>	character*14	a14	Date/time string (YYYYMMDDHHMMSS)
2	SPACES		2x	1 space
2	<i>latitude</i>	Real	f9.4,1x	Latitude (south=negative)
2	<i>longitude</i>	Real	f9.4,1x	Longitude (west=negative)
3	SPACES		2x	
3	<i>id</i>	character*40	a40,3x	Station ID
3	<i>namef</i>	character*40	a40,3x	Data type
4	SPACES		2x	
4	<i>platform</i>	character*40	a16,2x	Platform name
4	<i>source</i>	character*40	a16,2x	Source of ob
4	<i>elevation</i>	Real	f8.0,2x	Terrain height (m) at ob location
4	<i>is_sound</i>	Logical	l4,2x	Is a non-surface ob?
4	<i>bogus</i>	Logical	l4,2x	Is a bogus ob?
4	<i>meas count</i>	Integer	i5	# of levels in sounding

If it is a surface observation (i.e., *is_sound* = FALSE), then there is one line in the observation itself containing pairs of observations of various fields and their respective quality control (QC) flags. The variable name for each quantity can be determined by adding “_data” to the name base in the table, and the associated quality control flag by adding “_qc”. Shading of the variable name in the table below indicates that the variable is not used by WRF.

OBS_DOMAIN Surface Observation Format			
Name Base	Variable Type	Format	Description (units)
SPACES		1x	
<i>slp</i>	real	f11.3,1x	Sea-level pressure (Pa)
<i>ref pres</i>	real	f11.3,1x	Reference pressure (Pa)
<i>height</i>	real	f11.3,1x	Height (m MSL)
<i>temperature</i>	real	f11.3,1x	Temperature (K)
<i>u met</i>	real	f11.3,1x	U wind component (m s ⁻¹)

<i>v met</i>	real	f11.3,1x	V wind component (m s ⁻¹)
<i>rh data</i>	real	f11.3,1x	Relative humidity (%)
<i>psfc data</i>	real	f11.3,1x	Surface pressure (Pa)
<i>precip data</i>	real	f11.3,1x	Precipitation info

If the observation is not a surface observation (i.e., *is_sound* = TRUE) then there is one line in the observation itself per vertical level in the observation. The number of vertical levels in a given observation is determined by the header component *meas_count*. Each line contains pairs of observations of various fields and their respective quality control (QC) flags. The variable name for each quantity can be determined by adding “_data” to the name base in the table, and the associated quality control flag by adding “_qc”.

OBS_DOMAIN Non-Surface Observation Format			
Name Base	Variable Type	Format	Description (units)
SPACES		1x	
<i>pressure</i>	real	f11.3,1x	Pressure (Pa)
<i>height</i>	real	f11.3,1x	Height (m MSL)
<i>temperature</i>	real	f11.3,1x	Temperature (K)
<i>u met</i>	real	f11.3,1x	U wind component (m s ⁻¹)
<i>v met</i>	real	f11.3,1x	V wind component (m s ⁻¹)
<i>rh data</i>	real	f11.3,1x	Relative humidity (%)

In the OBS_DOMAIN files missing values are marked as -888888.000.

For surface observations, if the surface pressure (*psfc_data*) is missing then WRF will attempt to calculate surface pressure based on the sea-level pressure and the station elevation of the observation. If sea-level pressure is missing then it will assume the sea level pressure is 1000 hPa and calculate surface pressure based on that value and the station elevation. If surface pressure is still missing after these attempts then temperature will be marked as missing since surface pressure is needed to convert from temperature to potential temperature. The reference pressure does not appear to be used by WRF; neither does the height.

For non-surface observations, temperature will be marked as missing if neither pressure nor height is available, since it relies on these variables to convert temperature to potential temperature. If height is available but not pressure, then the pressure will be calculated based on the height and the model pressure and model height field.

For surface and non-surface observations, calm winds will be marked as missing winds because many surface stations report 0 wind speed and 0 wind direction for missing data.

The quality control flags follow the *little_r* format and are additive. For example, if you have a calm wind (flag=4) and no buddies were found for the buddy check (flag=16384) the quality control flag is 4+16384=16388.

Quality Control Flags		
Flag	2 ^x	Description
2	1	Pressure derived from the first-guess height field (e.g., GFS)
4	2	Pressure derived from standard atmosphere assumption and height
8	3	Pressure derived from standard atmosphere assumption and temperature
16	4	Temperature/Dewpoint are zero
32	5	Calm wind
64	6	Negative wind speed
128	7	Wind direction did not fall between 0 and 360 degrees
129		WRF uses this to indicate winds that are earth-relative (instead of grid-relative)
256	8	Value was vertically interpolated (to get to level with first-guess field to QC against)
512	9	The observation was assigned a nearby pressure where a first-guess field was available for qc*
1024	10	Temperature appeared to have the wrong sign based on adjacent temperatures in the vertical profile (Obsgrid flipped the sign of the temperature and adjusted dewpoint/RH)
2048	11	A superadiabatic temperature remains after convective adjustment (Obsgrid marked temperature, dewpoint, and relative humidity as missing)
4096	12	Spike in wind (Obsgrid marked winds as missing)
8192	13	Dry convective adjustment used in attempt to remove superadiabatic level
16384	14	No buddies (nearby observations) were found to do a buddy check
32768	15	Data outside normal analysis time and not QC-ed
65536	16	Fails error max check (check against first guess field)
131072	17	Fails buddy check (check against nearby obs)
262144	18	Observation outside of domain

Note that in old versions of Obsgrid, the “data outside normal analysis time and not QC-ed” flag did not exist and so those conditions now assigned to 2¹⁶, 2¹⁷, and 2¹⁸, were assigned 2¹⁵, 2¹⁶, and 2¹⁷, respectively.

WRF will mark as missing observations whose quality control flag is greater than or equal to 30000 or whose quality control flag is negative. Essentially, these observations are dropped by WRF from being nudged. The quality control flag of 129 for wind is a special flag indicating that the winds are earth-relative and thus must be rotated by WRF to be grid-relative. Unless this flag is set, WRF assumes that the winds ingested from OBS_DOMAIN files are grid-relative.

An example OBS_DOMAIN with only a surface observation is shown below followed by an explanation.

```

20120207084800
  35.2700 -113.9500
  KIGM
  NETWORK
FM-15 ASOS          1033.    F    F    1
-888888.000 -888888.000 -888888.000 -888888.000 1033.000 0.000 277.150
0.000 -4.077 0.000 2.194 0.000 35.343 0.000
90063.531 0.000 -888888.000 -888888.000

```

For this surface observation:

- Line 1
 - $data_char = 20120207084800$ → Observation taken at 0848 UTC on February 7, 2012
- Line 2
 - $latitude = 35.2700, longitude = -113.9500$ → This is the location of the observation.
- Line 3
 - $id = KIGM$ → The name of this observation location is KIGM.
 - $namef = NETWORK$ → The data type is a fictional value NETWORK, for illustration purposes. The only way the WRF observation nudging code uses $namef$ is to alter the contents of $platform$ in certain cases.
- Line 4
 - $platform = FM-15 ASOS$ → The platform of this observation is FM-15 ASOS.
 - $source =$ → The source of the observation was not specified.
 - $elevation = 1033.$ → The KIGM observation station is at 1033 m MSL.
 - $is_sound = F$ → This observation is a surface observation.
 - $bogus = F$ → This observation is not a bogus observation.
 - $meas_count = 1$ → There is one level in this observation.
- Line 5
 - $slp_data/slp_qc = -888888.000/-888888.000$ → Sea-level pressure data is missing.
 - $ref_pres_data/ref_pres_qc = -888888.000/-888888.000$ → Reference pressure data is missing.
 - $height_data/height_qc = 1033.000/0.000$ → The observation took place at 1033 m MSL, which matches the value set to elevation, consistent with this being a surface observation. The QC flag indicates that there are no known issues with the height data.
 - $temperature_data/temperature_qc = 277.150/0.000$ → The temperature observed was 277.15K and there are no known QC issues.
 - $u_met_data/u_met_qc = -4.077/0.000$ → The u-wind component observed was -4.077 m s^{-1} and there are no known QC issues.
 - $v_met_data/v_met_qc = 2.194/0.000$ → The v-wind component observed was -2.194 m s^{-1} and there are no known QC issues.
 - $rh_met_data/rh_met_qc = 35.343/0.000$ → The observed relative humidity was 35.343% and there were no known QC issues.
 - $psfc_data/psfc_qc = 90063.531/0.000$ → The observed pressure was 900.63531 hPa and there were no known QC issues.
 - $precip_data/precip_qc = -888888.000/-888888.000$ → The precipitation data is missing.

An example OBS_DOMAIN with only a single rawinsonde is shown below followed by an explanation.

```

20120207120000
  37.7600 -122.2200
  KOAK
  FM-35
  100800.000    0.000    6.000    0.000    284.750    0.000    -2.939
0.000    0.943    0.000    80.004    0.000
100000.000    0.000    70.000    0.000    284.950    0.000    -3.725
256.000    2.611    256.000    75.761    16384.000
  98800.000    256.000    170.522    256.000    285.114    256.000    -4.904
256.000    5.114    256.000    68.341    16640.000

```

For the rawinsonde:

- *Line 1*
 - *data_char* = 20120207120000 → Observation taken at 1200 UTC on February 7, 2012
- *Line 2*
 - *latitude* = 37.7600, *longitude* = -122.2200 → This is the location of the observation.
- *Line 3*
 - *id* = KOAK → The name of this observation location is KOAK.
 - *namef* = NETWORK → The data type is a fictional value NETWORK, for illustration purposes. The only way the WRF observation nudging code uses *namef* is to alter the contents of *platform* in certain cases.
- *Line 4*
 - *platform* = FM-35 → The platform of this observation is FM-35.
 - *source* = → The source of the observation was not specified.
 - *elevation* = 6. → The KOAK observation station is at 6 m above MSL.
 - *is_sound* = T → This observation is not a surface observation.
 - *bogus* = F → This observation is not a bogus observation.
 - *meas_count* = 3 → There are three levels in this observation.
- *Line 5*
 - *pressure_data/pressure_qc* = 100800.000/0.000 → The first level in this observation is at 1008 hPa. The QC flag indicates that there are no known issues with the pressure data at this level.
 - *height_data/height_qc* = 6.000/0.000 → The observation took place at 6 m MSL, which matches the value set to elevation, consistent with this being a surface observation. Rawinsondes are generally constructed by combining a surface observation with data from the balloon for all the other levels. The QC flag indicates that there are no known issues with the height data at this level.
 - *temperature_data/temperature_qc* = 284.750/0.000 → The temperature observed was 284.75K and there are no known QC issues.
 - *u_met_data/u_met_qc* = -2.939/0.000 → The u-wind component observed was -2.939 m s⁻¹ and there are no known QC issues.
 - *v_met_data/v_met_qc* = 0.943/0.000 → The v-wind component observed was 0.943 m s⁻¹ and there are no known QC issues.

- $rh_met_data/rh_met_qc = 80.004/0.000$ → The observed relative humidity was 80.004% and there were no known QC issues.
- Line 6
 - $pressure_data/pressure_qc = 100000.000/0.000$ → This level is at 1000 hPa. The QC flag indicates that there are no known issues with the pressure data at this level.
 - $height_data/height_qc = 70.000/0.000$ → The observation took place at 70 m MSL, which is 64 m AGL. The QC flag indicates that there are no known issues with the height data at this level.
 - $temperature_data/temperature_qc = 284.950/0.000$ → The temperature observed was 284.95K and there are no known QC issues.
 - $u_met_data/u_met_qc = -3.725/256.000$ → The u-wind component observed was -3.725 m s^{-1} . The QC flag indicates that the value was vertically interpolated to get to this level in order to allow for the observation to be quality controlled against the first-guess field (e.g., GFS).
 - $v_met_data/v_met_qc = 2.611/256.000$ → The v-wind component observed was 2.611 m s^{-1} . The QC flag indicates that the value was vertically interpolated.
 - $rh_met_data/rh_met_qc = 75.761/16384.000$ → The observed relative humidity was 75.761% and no nearby observations were found to serve as buddies for the buddy check.
- Line 7
 - $pressure_data/pressure_qc = 988000.000/256.000$ → This level is at 988 hPa. The QC flag indicates that the value was vertically interpolated to get to this level in order to allow for the observation to be quality controlled against the first-guess field (e.g., GFS).
 - $height_data/height_qc = 170.522/256.000$ → The observation took place at 170.522 m MSL. The QC flag indicates that the value was vertically interpolated.
 - $temperature_data/temperature_qc = 285.114/256.000$ → The temperature observed was 285.114K. The QC flag indicates that the value was vertically interpolated.
 - $u_met_data/u_met_qc = -4.904/256.000$ → The u-wind component observed was -4.904 m s^{-1} . The QC flag indicates that the value was vertically interpolated.
 - $v_met_data/v_met_qc = 5.114/256.000$ → The v-wind component observed was 5.114 m s^{-1} . The QC flag indicates that the value was vertically interpolated.
 - $rh_met_data/rh_met_qc = 68.341/16640.000$ → The observed relative humidity was 68.341%. The QC flag indicates that the value was vertically interpolated (256) and that no nearby observations were found to serve as buddies for the buddy check (16384).

Appendix B: WRF Observation Nudging Code Layout

The observation nudging capabilities in WRF are mostly carried out by: `share/wrf_fddaobs_in.F`, `phys/module_fddaobs_rtfdda.F`, and `module_fddaobs_driver.F`. The basic program flow is shown below. Each line lists the subroutine name followed by the file in which the subroutine resides. Subroutines that do not have observation-nudging specific code are italicized; these are included to show how the observation nudging code relates to the overall WRF code.

```
start_domain (share/start_domain.F)
-start_domain_em (dyn_em/start_em.F)
--phy_init (phys/module_physics_init.F)
---fdob_init (phys/module_physics_init.F)
----fddaobs_init (phys/module_fddaobs_rtfdda.F)
        Initialize observation nudging. Set variables based on namelist variables and do some error
        checking on the namelist variables.

wrf_run (main/module_wrf_top.F)
-integrate (frame/module_integrate.F)
--med_before_solve_io (share/mediation_integrate.F)
---med_fddaobs_in (share/mediation_integrate.F)
----wrf_fddaobs_in (share/wrf_fddaobs_in.F)
-----in4dob (share/wrf_fddaobs_in.F)
        Read observations from file.
--solve_interface (share/solve_interface.F)
---solve_em (dyn_em/solve_em.F)
----first_rk_step_part2 (dyn_em/module_first_rk_step_part2.F)
-----fddaobs_driver (phys/module_fddaobs_driver.F)
-----errob (phys/module_fddaobs_rtfdda.F)
        Calculate the innovation (difference between observed value and model value)
-----nudob (phys/module_fddaobs_rtfdda.F)
        Calculate nudging tendency terms.
```


Acknowledgements:

The author thanks Yubao Liu, Al Bourgeois, and Jason Knievel of the National Center for Atmospheric Research (NCAR) and Bob Dumais of the Army Research Lab for reviewing this document.

References:

Anthes, Richard A., 1974: Data Assimilation and Initialization of Hurricane Prediction Models. *J. Atmos. Sci.*, **31**, 702–719.

Anthes, R.A., E.-Y. Hsie, and Y.-H. Kuo, 1987: Description of the Penn State/NCAR mesoscale model version 4 (MM4). NCAR Tech. Note NCAR/TN-282+STR, 66 pp. [Available from the National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307].

Charney, J., M. Halem, and R. Jastrow, 1969: Use of Incomplete Historical Data to Infer the Present State of the Atmosphere. *J. Atmos. Sci.*, **26**, 1160–1163.

Deng, A., D. Stauffer, B. Gaudet, J. Dudhia, C. Bruyere, W. Wu, F. Vandenberghe, Y. Liu, and A. Bourgeois, 2009: Update on WRF-ARW end-to-end multi-scale FDDA system. 10th WRF Users' Workshop, NCAR, 23–26 June, Boulder, CO. 1.9. 14 pp. Available: <http://www2.mmm.ucar.edu/wrf/users/workshops/WS2009/WorkshopPapers.php>

Gaudet, B., D. Stauffer, N. Seaman, A. Deng, K. Schere, R. Gilliam, J. Pleim, and R. Ellemen, 2009: Modeling Extremely Cold Stable Boundary Layers over Interior Alaska using a WRF FDDA System. 13th Conference on Mesoscale Processes, American Meteorological Society, 17–20 August, Salt Lake City, Utah. 18.1. 8 pp. Available: https://ams.confex.com/ams/13Meso/techprogram/paper_155043.htm

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1995: A Description of the Fifth-Generation Penn State / NCAR Mesoscale Model (MM5), NCAR/TN-398+STR, National Center for Atmospheric Research, Boulder, CO, 122 pp.

Hoke, James E., and Richard A. Anthes, 1976: The Initialization of Numerical Models by a Dynamic-Initialization Technique, *Mon. Wea. Rev.*, **104**, 1551–1556.

Hoke, James E., and Richard A. Anthes, 1977: Dynamic Initialization of a Three-Dimensional Primitive-Equation Model of Hurricane Alma of 1962. *Mon. Wea. Rev.*, **105**, 1266–1280.

Kistler, R. E., 1974: A Study of Data Assimilation Techniques in an Auto-barotropic Primitive Equation Channel Model. M. S. thesis, The Pennsylvania State University, 84 pp. [Available from the Dept. of Meteorology, The Pennsylvania State University, University Park, PA 16802]

Lei, L., D. R. Stauffer, and A. Deng, 2012: A hybrid nudging-ensemble Kalman filter approach to data assimilation in WRF/DART. *Quart. J. Roy. Meteor. Soc.*, **138**, 2066–2078.

Liu, Y., A. Bourgeois, T. Warner, S. Swerdlin, and J. Hacker, 2005: Implementation of observation-nudging based FDDA into WRF for supporting ATEC test operations. 6th Annual WRF and 15th Annual MM5 Users' Workshop, NCAR, 27–30 June, Boulder, CO, 10.7. 4 pp. Available: <http://www2.mmm.ucar.edu/wrf/users/workshops/WS2005/WorkshopPapers.htm>

Liu, Y., Y. Wu, L. Pan, A. Bourgeois, G. Roux, J. C. Knierel, J. Hacker, J. Pace, F. W. Gallagher, and S. F. Halvorson, 2015: Recent developments of NCAR's 4D-Relaxation Ensemble Kalman Filter system. *19th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, American Meteorological Society, 5–8 January, Phoenix, AZ. 5.3.

Reen, B. P., and D. R. Stauffer, 2010: Data assimilation strategies in the planetary boundary layer. *Boundary-Layer Meteorology*, **137**, 237–269.

Reen, B. P., R. E. Dumais Jr., J. E. Passner, 2013: Mitigating Excessive Drying Caused by Observation Nudging. 14th Annual WRF Workshop, NCAR, 24–28 June, Boulder, CO, P13. 8 pp. Available: <http://www2.mmm.ucar.edu/wrf/users/workshops/WS2013/WorkshopPapers.htm>

Reen, B. P., R. E. Dumais, Jr., and J. E. Passner, 2015: Mitigating Excessive Drying From the Use of Observations in Mesoscale Modeling. Early online release. *J. Appl. Meteorol. Climatol.*, <http://dx.doi.org/10.1175/JAMC-D-14-0301.1>

Stauffer, David R., and Nelson L. Seaman, 1990: Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part I: Experiments with Synoptic-Scale Data. *Mon. Wea. Rev.*, **118**, 1250–1277.

Stauffer, David R., Nelson L. Seaman, and Francis S. Binkowski, 1991: Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part II: Effects of Data Assimilation within the Planetary Boundary Layer. *Mon. Wea. Rev.*, **119**, 734–754.

Stauffer, David R., and Nelson L. Seaman, 1994: Multiscale Four-Dimensional Data Assimilation. *J. Appl. Meteor.*, **33**, 416–434.

Thompson, P. D., 1961: A Dynamical Method of Analyzing Meteorological Data. *Tellus*, **13**, 334–349.