Assessment of Cloud Retrieval for IASI 1D-Var Cloudy-Sky Assimilation and Improvement with an ANN Approach

AHREUM LEE,^a BYUNG-JU SOHN,^a ED PAVELIN,^b YOONJAE KIM,^c HYUN-SUK KANG,^d
ROGER SAUNDERS,^b AND YOUNG-CHAN NOH^{a,c}

^a School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea

^b Met Office, Exeter, United Kingdom

^c National Institute of Meteorological Sciences, Korea Meteorological Administration, Jeju, South Korea

^d Numerical Modeling Center, Korea Meteorological Administration, Seoul, South Korea

^e Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin–Madison, Madison, Wisconsin

(Manuscript received 21 October 2019, in final form 12 March 2020)

ABSTRACT

The Unified Model (UM) data assimilation system incorporates a 1D-Var analysis of cloud variables for assimilating hyperspectral infrared radiances. For the Infrared Atmospheric Sounding Interferometer (IASI) radiance assimilation, a first guess of cloud top pressure (CTP) and cloud fraction (CF) is estimated using the minimum residual (MR) method, which simultaneously obtains CTP and CF by minimizing radiance difference between observation and model simulation. In this study, we examined how those MR-based cloud retrievals behave, using “optimum” CTP and CF that yield the best 1D-Var analysis results. It is noted that the MR method tends to overestimate cloud top height while underestimating cloud fraction, compared to the optimum results, necessitating an improved cloud retrieval. An artificial neural network (ANN) approach was taken to estimate CTP as close as possible to the optimum value, based on the hypothesis that CTP and CF closer to the optimum values will bring in better 1D-Var results. The ANN-based cloud retrievals indicated that CTP and CF biases shown in the MR method are much reduced, giving better 1D-Var analysis results. Furthermore, the computational time can be substantially reduced by the ANN method, compared to the MR method. The evaluation of the ANN method in a global weather forecasting system demonstrated that it helps to use more temperature channels in the assimilation, although its impact on UM forecasts was found to be near neutral. It is suggested that the neutral impact may be improved when error covariances for the cloudy sky are employed in the UM assimilation system.

1. Introduction

Satellite observations have become indispensable in the current numerical weather prediction (NWP) systems (Bauer et al. 2015; McNally et al. 2014). Among satellite observations, hyperspectral infrared sounders are a single largest contributor to forecast accuracy (Hilton et al. 2012; Joo et al. 2013; McNally et al. 2006). However, this contribution is largely from the clear-sky

radiance assimilation. Because of the difficulty to correctly simulate upwelling infrared radiances for cloudy conditions, cloudy-sky data assimilation using the hyperspectral infrared sounders is very limited in terms of data volume used and area applied. Considering that approximately 60% of the globe is cloud-covered at a given moment (Wylie and Menzel 1999), additional use of hyperspectral sounder data over the cloudy areas in the data assimilation must have a potential to improve forecasting skill. Despite its necessity, there seem to be few studies of actively using hyperspectral infrared measurements over the cloudy area for data assimilation.

Li et al. (2005) developed a “cloud-clearing” method for the Atmospheric Infrared Sounder (AIRS) radiance data. This method constructs clear-sky radiances for the cloudy area that would be measured if it were clear. In doing that, cloud-affected hyperspectral infrared radiances were altered with the aid of clear-sky information

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/WAF-D-19-0218.s1>.

Corresponding author: Byung-Ju Sohn, sohn@snu.ac.kr

DOI: 10.1175/WAF-D-19-0218.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

from the Moderate Resolution Imaging Spectroradiometer (MODIS) within an AIRS pixel. The resultant cloud-cleared radiances (i.e., “presumed” clear-sky radiances) were then assimilated by applying a clear-sky assimilation method. However, since MODIS clear-sky information is extended to the cloudy area, the constructed hyperspectral infrared radiances tend to be biased toward the clear-sky values.

In the scheme introduced by McNally and Watts (2003), hyperspectral infrared radiances over cloudy areas are used, but only for the channels not contaminated with the presence of clouds. In this case, channels are deemed to be clear if observed minus simulated channel radiances are smaller than a preset criterion, and then those selected clear channels are used for the assimilation as clear-sky radiances.

Pavelin et al. (2008) introduced a method to use explicit cloud properties for cloudy-sky data assimilation, which was referred to as the “Cloudy 1D-Var” method. In this technique, for a given field of view, cloud parameters (i.e., cloud top and fraction) are determined under the single-layer graybody cloud assumption. Those retrieved cloud parameters are then taken into account for using cloud-affected radiances in the subsequent one-dimensional variational (1D-Var) analysis. They showed that 1D-Var analysis with the retrieved cloud parameters results in better performance, compared to result from clear-sky only method or the McNally and Watts (2003) method.

However, it is difficult to evaluate the retrieved cloud properties based on the assumption of opaque and single-layer cloud because such assumed cloud properties are mostly far from the observed cloud properties. Furthermore, although the retrieved cloud parameters in the Cloudy 1D-Var method play an important role not only in simulating radiances but also in selecting channels used for the following 1D-Var analysis, it has not been examined whether these retrieved cloud parameters are the best solutions for the 1D-Var analysis. In other words, there might be different solutions that lead to better 1D-Var analysis even if the single-layer graybody cloud assumption is made.

This study attempts to improve 1D-Var analysis over cloudy areas by introducing the best solutions that yield the optimal 1D-Var analysis of temperature and humidity under the single-layer graybody cloud assumption. We will examine how the obtained cloud parameters from the Cloudy 1D-Var method are different from those best solutions. Then we will devise a way to retrieve cloud parameters as close as possible to those that yield the best 1D-Var analysis results. The newly developed retrieval method will be first tested within a simulated dataset, to evaluate whether it provides a positive

FIG. 1. Flowchart of IASI 1D-Var assimilation process. Two different cloud retrieval methods: (a) MR method and (b) ANN method are compared in this study.

impact on the 1D-Var analysis. Furthermore, the impact of the new method will be examined for the possibility of implementation in the operational preprocessor of the Unified Model (UM) NWP system.

2. IASI 1D-Var data assimilation

In this section, we briefly introduce the Cloudy 1D-Var method (Pavelin et al. 2008) in the Infrared Atmospheric Sounding Interferometer (IASI) 1D-Var assimilation. The flowchart in Fig. 1 summarizes the steps of the IASI 1D-Var assimilation.

a. Retrieval of cloud top pressure and cloud fraction

In the Cloudy 1D-Var method, the assumed single-layer graybody cloud is expressed with cloud top pressure (CTP) and cloud fraction (CF). For the calculation of the cloud's radiative effect within the given IASI field of view, a first guess pair of CTP and CF is provided. They are simultaneously retrieved by the minimum residual method (Eyre and Menzel 1989) (hereafter referred to as MR method), as expressed as the method (Fig. 1a) in the dashed rectangle of Fig. 1. The first guess pair is determined from the observed and simulated radiances at 10 IASI channels given in Table 1. These 10 channels were selected to locate the channel's weighting functions evenly throughout the troposphere so that the cloud's influences on all the radiances are represented.

Figure 2 shows an example of the residual between observed and simulated radiances, cumulated over the 10 IASI channels, for all given CTP and CF pairs. For a CTP and CF pair, the residual is defined as follows:

TABLE 1. List of 10 IASI channels used in the MR method.

Channel number	Wavenumber (cm ⁻¹)	Channel number	Wavenumber (cm ⁻¹)
179	689.50	280	714.75
205	696.00	327	726.50
212	697.75	407	746.50
242	705.25	662	810.25
246	706.25	1786	1091.25

$$\text{residual} = \sum_{i=1}^{10} [R_{\text{obs}}^i - H(\mathbf{x}_b, p_c, N)]^2$$

$$= \sum_{i=1}^{10} \{R_{\text{obs}}^i - [(1-N)R_{\text{clr}}^i + NR_{\text{cld}}^i(p_c)]\}^2, \quad (1)$$

where R_{obs} is the observed radiance; H is a forward operator; $\mathbf{x}_b$ is the background state; p_c and N are the cloud top pressure and cloud fraction, respectively; R_{clr} is the simulated clear-sky radiance; $R_{\text{cld}}(p_c)$ is the simulated radiance for the overcast sky whose cloud top pressure is p_c ; and i is an index for the 10 IASI channels. When the MR method finds one pair of CTP and CF having the smallest residual, for practical purposes to reduce the computational burden, each RTTOV-pressure level (Table 2) is assigned as a CTP and a corresponding CF is solved:

$$N = \frac{\sum_{i=1}^{10} [R_{\text{cld}}^i(p_c) - R_{\text{clr}}^i](R_{\text{obs}}^i - R_{\text{clr}}^i)}{\sum_{i=1}^{10} [R_{\text{cld}}^i(p_c) - R_{\text{clr}}^i]^2}. \quad (2)$$

Thus, at each RTTOV pressure level, a CTP and CF pair satisfying the single-layer gray cloud assumption is found (marked by “×” in Fig. 2). After finding pairs of CTP and CF at all 27 RTTOV pressure levels, a pair showing the overall minimum residual (marked by “Δ” in Fig. 2) is found.

The obtained CTP and CF are then used as initial guess values for optimizing CTP and CF with background atmospheric variables in 1D-Var. This time, 182 operational IASI channels are used for producing optimized CTP and CF as outputs. In this optimization process, the UM operational 6-h forecast-error covariance matrix $\mathbf{B}$ and the clear-sky observation error covariance matrix $\mathbf{R}$ are used. Diagonal and off-diagonal components of temperature, humidity, and surface states exist in the background error covariance while only diagonal components are available for CTP and CF. Values for diagonal components of background error covariance for CTP and CF are 1000² and 1², respectively. For the observation error covariance, only

FIG. 2. One example showing the residual distribution for all possible pairs of CTP and CF, for the given set of 10 IASI channels. The × marks denote CTP and CF pairs obtained by applying Eq. (2) at 27 levels, and “Δ” represents a first guess pair of CTP and CF from the MR method.

the diagonal component is used for 182 channels. If the initial CTP and CF values are not reasonable, the cost function in the 1D-Var will become too large and the solution will not converge. If an IASI pixel results in such a failure, then the scene is considered to be nonconverged and the IASI radiances will not be assimilated.

b. 1D-Var analysis

Once the optimized CTP and CF are obtained at each IASI pixel, the assimilation type is determined based on the CF value. If the optimized CF is larger than 0.05, the field of view is considered to be cloudy. Otherwise, it is considered to be clear. If the pixel is determined to be cloudy, channels are selected among the 182 channels for further assimilation if the integrated temperature Jacobian value above the optimized CTP height is greater than 90% of the total Jacobian value. Thus, either 182 channels for clear conditions or selected channels under cloudy conditions are used in the 1D-Var assimilation. In this 1D-Var process, the UM operational 6-h forecast-error covariance matrix $\mathbf{B}$ and the clear-sky observation error covariance matrix $\mathbf{R}$ are used. If the case does not converge in the 1D-Var analysis process, the scene is then rejected in the assimilation. The procedures of retrieval of CTP and CF, cloudy (or clear) scene determination, and 1D-Var process given in Fig. 1 is referred to as the IASI 1D-Var data assimilation. This assimilation process is applied only over open ocean due to the uncertainty of the surface emissivity over land and sea ice.

TABLE 2. Pressure levels used in the RTTOV model.

Index	Pressure (hPa)	Index	Pressure (hPa)	Index	Pressure (hPa)
1	102.05	10	358.28	19	749.12
2	122.04	11	396.81	20	795.09
3	143.84	12	436.95	21	839.95
4	167.95	13	478.54	22	882.80
5	194.36	14	521.46	23	922.46
6	222.94	15	565.54	24	957.44
7	253.71	16	610.60	25	985.88
8	286.60	17	656.43	26	1005.43
9	321.50	18	702.73	27	1013.25

3. Development of an ANN method for cloud retrieval

Here we introduce a new method to retrieve the first guess pair of CTP and CF (expressed as the method in Fig. 1b in the dashed rectangle of Fig. 1). In this study, the developed method will replace the MR method in the Cloudy 1D-Var method, in order to examine the impact of the new retrieval method on the IASI 1D-Var data assimilation. Details about the development of the retrieval method follow.

a. Preparation of simulation dataset

In this study, for developing the retrieval method and its validation, datasets are obtained from 25 000 atmospheric profiles that were compiled from the European Centre for Medium-Range Weather Forecasts (ECMWF) global operational short-range forecasts spanning the time period of 1 September 2013–31 August 2014 (Eresmaa and McNally 2014). These profiles were compiled through a randomized selection to preserve global and seasonal statistical features that the original forecast data hold. Out of 25 000 profiles, 14 804 profiles were selected for cloudy cases over open ocean, after excluding land profiles. Atmospheric profiles of about a half of cloudy-sky profiles (i.e., 7498 out of 14 804) were duplicated for constructing the clear-sky profiles (by assuming zero cloud liquid water), after considering the global cloud occurrence of around 70% (Wylie and Menzel 1999). Thus, a total of 22 302 sample profiles are used. The original ECMWF 14 804 samples over open ocean are not enough because we need cloudy and clear conditions, as well as training and validation datasets. Thus, constructed clear-sky profiles are assumed to be from locations very close to those for cloudy-sky profiles. However, these two datasets may become independent once simulated IASI brightness temperatures are binned with profiles.

The data include temperature and humidity profiles, and vertical distributions of cloud fraction and cloud liquid/ice water content in both northern and southern

hemispheres. In this study, these data are considered to be true (referred to as “truth $\mathbf{x}_t$ ”) and used for simulating IASI-observed radiances and model background states, following the methodology used by Pavelin et al. (2008).

The simulation of IASI radiances at 182 channels ($\mathbf{y}$), which are used in the operational UM IASI data assimilation system, is done as follows:

$$\mathbf{y} = H(\mathbf{x}_t) + \mathbf{R}, \quad (3)$$

where $\mathbf{R}$ is the observation error covariance matrix. The radiative transfer for TOVS (RTTOV) version 9.3 model was used as a forward operator H in this study. In addition to the truth temperature and humidity profiles used as inputs to the RTTOV model, the cloud profiles from the ECMWF forecasts are taken for the cloudy case simulations. After conducting the radiative transfer calculations, expected observation errors are added to simulated IASI brightness temperatures by assuming that observation errors are random with an unbiased Gaussian distribution. Observation errors are obtained from the diagonal component of the clear-sky $\mathbf{R}$ matrix used in the operational UM Observation Processing System (OPS).

In addition to the IASI radiance simulations, the development procedures leading to temperature and humidity analysis through the 1D-Var process require the model background atmospheric state. Since ECMWF forecasts were treated as the truth, random forecast errors from the UM operational 6-h forecast-error covariance matrix $\mathbf{B}$ are added to the truth profiles:

$$\mathbf{x}_b = \mathbf{x}_t + \mathbf{B}, \quad (4)$$

where the background fields $\mathbf{x}_b$ comprise temperature and humidity profiles, surface temperature, surface humidity, and skin temperature.

In this study, since the $\mathbf{B}$ matrices in the UM OPS are only available at three latitudinal regions (90°–30°N, 30°N–30°S, and 30°–90°S), trainings of the new retrieval

method are taken at the same three latitudinal regions. However, because 22 302 samples are not enough for training as well as validation for three different regions, the same samples were used for developing the new method at three regions, but with regionally different background errors. It might be thought that extreme profiles at one region (such as profiles reflecting humid condition in 30°N–30°S region) may not fit in conditions for another region (such as 90°–30°N region). However, finding that the extreme ranges are quite similar to each other among three regions, due to the dataset covering the 1-yr period including the summer and winter, the current approach of using the same global data for three latitudinal regions is considered to be reasonable. Thus, the following process is equally applied with the same 22 302 profiles for three regions.

b. Optimum cloud parameters

This research was motivated from a theory that the first guess pair of CTP and CF, yielding a minimum residual, may not necessarily result in the best 1D-Var analysis. We define here a pair of CTP and CF yielding the best 1D-Var analysis result as “optimum” cloud parameters. To obtain the optimum cloud parameters, we first take 27 pairs of CTP and CF using Eq. (2) (marked by ×s in Fig. 2) for the 27 RTTOV pressure levels in Table 2. Note that the pair of CTP and CF at the “Δ” mark resulted in the overall minimum residual among 27 pairs. Each of those 27 pairs is now used as the first guess pair of CTP and CF for the IASI 1D-Var assimilation system outlined in Fig. 1, and resultant temperature and humidity analysis are obtained.

To examine how each pair gives rise to errors in the analysis, compared to the truth, the cumulative normalized root-mean-square error (RMSE) of temperature and humidity analysis is calculated by using the following equation:

$$\text{RMSE}' = \sqrt{\frac{1}{27} \sum_{k=1}^{27} \left[\frac{T_a^k(p_c, N) - T_t^k}{\sigma_T^k} \right]^2} + \sqrt{\frac{1}{27} \sum_{k=1}^{27} \left[\frac{q_a^k(p_c, N) - q_t^k}{\sigma_q^k} \right]^2}, \quad (5)$$

where T_a and q_a are the temperature and humidity from the analysis respectively; T_t and q_t are the truth temperature and humidity, respectively; and σ_T and σ_q are standard deviations of background error for temperature and humidity, respectively. In this RMSE' calculation, the UM bottom 27 pressure levels (from the surface to 102.05-hPa level) are used, and k is the index for the level.

FIG. 3. RMSE' at 27 pairs of CTP and CF denoted by × marks in Fig. 2. The “Δ” denotes a first guess pair of CTP and CF from the MR method whereas “*” indicates the optimum CTP and CF pair showing the smallest RMSE'.

The obtained RMSE' values for the given pairs of CTP and CF (marked by ×s in Fig. 2) are presented in Fig. 3. It is demonstrated that the smallest RMSE' is shown at the pair (CTP = 795.09 hPa, CF = 0.44) marked by “*,” instead of the pair (CTP = 436.95 hPa, CF = 0.16) marked by “Δ” that showed the minimum residual. From this demonstration, it suffices to conclude that the first guess pair of CTP and CF retrieved by the MR method may not necessarily produce the smallest RMSE' in the 1D-Var analysis, suggesting that there may be a different pair of CTP and CF producing the lowest RMSE' (or best analysis results). We call the CTP and CF pair yielding the best among 27 different analysis outputs as the optimum cloud parameters.

To construct the optimum pairs of CTP and CF from all 22 302 samples, we take the procedures summarized in Fig. 4. For any given IASI radiances and corresponding background state, all possible combination of CTP and CF pairs can be introduced to the assimilation system to find a CTP and CF pair yielding the smallest RMSE'. However, because of the computational burden, we follow the procedures employed in the MR method in the UM OPS; we first find each level's CTP and CF pair throughout 27 levels using Eq. (2), and then those 27 pairs of CTP and CF (i.e., output in Fig. 4a) are sequentially used as inputs for the following 1D-Var analysis.

For the given IASI radiances and the corresponding background atmospheric state, the 1D-Var analysis gives rise to 27 sets of profiles of temperature and humidity, which will then be compared against the truth by

FIG. 4. Schematic diagram showing the method to obtain the optimum cloud parameters. Red dashed box shows the process generating (a) 27 pairs of CTP and CF (as denoted by $\times$ marks in Fig. 2). (b) $RMSE'$ values at 27 RTTOV levels are calculated by comparing the analysis of temperature and moisture to the truth x_t , as exemplified in Fig. 3, and the optimum CTP and CF values are determined. The upper index n is the ordinal number of total N samples and the lower index j represents the RTTOV pressure level. Other notations are found in the text.

calculating their respective $RMSE'$ (i.e., process shown in Fig. 4b). Finally, the optical cloud parameters (CTP and CF pair) producing the smallest $RMSE'$ result are obtained for a particular profile. By repeating the procedures summarized in Fig. 4 for the entire truth data, 22 302 pairs of CTP and CF corresponding to ECMWF truth states are constructed as the optimum cloud parameters dataset, yielding the best analysis outputs. After eliminating cases either showing extreme background atmospheric states or not yielding optimum cloud parameters, 20 327, 20 085, and 20 206 samples of truth profile, simulated IASI radiances, background states, and the optimum CTP were constructed for 90° – 30° N, 30° N– 30° S, and 30° – 90° S latitudinal regions, respectively.

Before optimum cloud parameters data are used for developing the new cloud retrieval method, we examine the impact of the optimum cloud estimates on the 1D-Var analysis results. Results were divided into the three latitudinal regions to examine whether results significantly vary with different latitudes. The error profiles for

the 1D-Var analysis 1) using the optimum cloud parameters and 2) using optimized CTP and CF from the MR method are given in Fig. 5, along with the background error profiles. It is clear that the MR method results in larger biases and higher RMSEs for both temperature and humidity for all three regions, compared with the best analysis results. This analysis result suggests that the MR method can be improved when the optimum values of CTP and CF are used.

It is of interest to examine what factors contributed the most to the better analysis results when the optimum cloud parameters are used in the 1D-Var analysis. For this purpose, the optimum cloud parameters are compared with the optimized CTP and CF from the MR method in the three latitudinal datasets (Fig. 6). In the case of CTP (Figs. 6a,c,e), there is a significant number of cases, in which the optimized CTPs from the MR method are smaller than corresponding optimum CTPs. Corresponding CF distributions are presented in terms of mean difference between the optimum CF and the optimized CF with the MR method, in order to exemplify systematic biases (Figs. 6b,d,f). It is clearly shown that CF is overestimated by the MR method if cloud top height is underestimated, in comparison to the optimum values. Conversely, underestimation of CF is clear in case of overestimated cloud top height by the MR method.

This result is consistent with the finding that the retrieved pair of CTP and CF lies along the line showing an inverse relation between CTP and CF as depicted in Figs. 2 and 3. Considering that overestimation of cloud top height by the MR method is much more frequent, as noted from CTP distributions (Figs. 6a,c,e), these results suggest that the MR method tends to produce higher cloud tops with smaller cloud fractions, compared to the optimum cloud parameters.

In the IASI 1D-Var assimilation, once optimized cloud parameters are obtained, channel selection procedures are performed to find IASI channels either not contaminated by clouds or partially contaminated by clouds. Here we examine how selected channels with the optimum results differ from channels selected with MR method results. The numbers of selected channels from the optimum cloud parameters are compared with those from the MR method (Figs. 7a,c,e). Both show a large number of selected channels in the upper-tropospheric CO_2 channels and less in the lower-tropospheric CO_2 and window channels. The difference in the number of the selected channels is presented in Figs. 7b, 7d, and 7f. It is noted that the optimum cloud parameters allowed more selected channels in all three latitudinal regions, compared with the MR method. The optimum CTPs tend to be at lower altitude, giving more selected

FIG. 5. Mean bias (solid lines) and RMSE (dashed lines) of (left) temperature and (right) humidity profiles from the 1D-Var analysis with the use of the optimized CTP and CF from the MR method (blue) and the optimum CTP and CF (red) at (a),(b) 90°–30°N, (c),(d) 30°N–30°S, and (e),(f) 30°S–90°S. Black dotted lines in the RMSE profiles represent the RMSE profiles of the background state.

FIG. 6. (left) Two-dimensional histograms of frequencies of optimum CTP vs optimized CTP from the MR method. Color represents data count. (right) Mean CF differences between optimum and the MR method are given in the same CTP pressure coordinates as in the left panels. Shown are the (a),(b) 90°–30°N, (c),(d) 30°N–30°S, and (e),(f) 30°–90°S regions.

channels whose Jacobians area above the cloud top, which are over 90% of the total. Moreover, because we expect not much discernible difference in the high peaking channels between two methods simply by lowering the

cloud top, the increased use of high peaking channels (i.e., from the 1st to 30th channels) represents more converged cases by the optimum cloud parameters. These results suggest that there may be more discarded

FIG. 7. (left) Number of used measurements at each IASI channel from the MR method (blue) and from the optimum CTP and CF (red). (right) Difference in the number of used measurements (optimum minus MR method). Latitudinal regions of (a),(b) 90°–30°N, (c),(d) 30°N–30°S, and (e),(f) 30°N–90°S. Black arrows in (e) and (f) represent IASI channels in CO₂, window, and water vapor absorption bands.

IASI channel radiances in the MR method due to the tendency to overestimate the cloud top height, and due to the reduced convergence in the assimilation process. Compared to the MR method, those discarded channel radiances might be used in the 1D-Var analysis when the optimum cloud parameters are used.

The mean difference between observed and background radiance ($O - B$) is calculated for each of 182 IASI channels, and results are given for the MR method and optimum parameters in Fig. 8. It is shown that $O - B$ means in three regions are nearly the same, except that $O - B$ means in the lower-tropospheric CO₂ and window channels (i.e., from 120 to 150 in the channel number) are slightly larger for the optimum cloud parameters. Taking this result together with Fig. 7, we conclude that the optimum cloud parameters led to use of more channels in the 1D-Var analysis and gave $O - B$ mean values similar to the MR method.

c. ANN-based cloud parameter retrieval

It was shown that the optimum cloud parameters produce the best 1D-Var analysis results that can be obtained from given 182 IASI channel radiances and background atmospheric states, compared to the truth states. However, the optimum cloud parameters can only be found when the truth atmospheric states are

known. Thus, it is desirable to retrieve cloud parameters as close as possible to the optimum cloud parameters. In this study, we attempt to retrieve the cloud top pressure close to the optimum value by linking variables of IASI radiances and background state to the corresponding optimum cloud parameters.

In doing so, we adopt an artificial neural network (ANN) approach that is well known for resolving any form of the nonlinear relationship (Desai et al. 2008). We tested results with various hidden layers and neuron numbers, and empirically found that the ANN method including one hidden layer with five neurons shows the best performance of capturing the nonlinearity between inputs and output used in this study.

As shown in Fig. 9, the ANN model consists of three layers: input layer, hidden layer, and output layer. As inputs, 182 IASI channel radiances, background temperature and humidity profiles, and surface variables (surface temperature, skin temperature, and surface humidity) are used. The 254 neurons in the input layer are scaled between -1 and 1 , and transported to the hidden layer via synaptic weights. Five neurons in the hidden layer sum up the weighted inputs with a bias value and those five-neuron values are passed through an activation function, which in this case is a tangent sigmoidal function. The outputs from the hidden layer

FIG. 8. (a)–(c) Mean $O - B$ (observed minus background brightness temperature) for the MR method (blue) and the optimum CTP and CF (red). Black arrows in (c) represent IASI channels in CO_2 , window, and water vapor absorption bands.

are used as inputs to the output layer, and the five values produce the final output through a linear function. When the coefficients for the weights and the linear function are determined, the iterative process minimizes the RMSE of CTPs up to 145 hPa by using a back-propagation algorithm to adjust the weights appropriately (Rumelhart et al. 1986).

Among 20 327, 20 085, and 20 206 samples for three different latitudinal regions ($90^\circ\text{--}30^\circ\text{N}$, $30^\circ\text{N--}30^\circ\text{S}$, and $30^\circ\text{--}90^\circ\text{S}$), randomly chosen 16 293, 16 108, and 16 198 samples (about 80% of the total samples) are used for the training of the ANN model. The chosen data are further divided into 70% for training and 30% for test. The remained 4034, 3977, and 4008 samples (about 20% of the total samples) are used to validate the ANN model at three latitudinal regions, respectively. It is noted that datasets allocated for training, test, and validation all show spatially and temporally homogeneous distributions (not shown). Considering that those datasets are from the randomized selection of ECMWF forecasts and show spatial and temporal homogeneity, they are considered to be independent of each other.

In the ANN model development, only CTP is determined, and then CF is obtained by inserting CTP to Eq. (2). The ANN method only determines the CTP value because CTP is more closely related to the temperature information coming from CO_2 channels, which not only compose a major portion of the 182 IASI channels but also give the largest impact on the 1D-Var

analysis. Note that in the MR method, calculations of 27-level CFs [as in Eq. (2)], as well as corresponding residuals are required [as in Eq. (1)] before finding a pair of CTP and CF yielding the minimum residual. By contrast, because only one time calculation of CF is required in the ANN method [as in Eq. (2)], the computation time to find the first guess pair of CTP and CF can be reduced by the ANN method to roughly 1/54 (approximately 1.85%) of the MR method. This ANN method can replace the MR method in the flow of the IASI 1D-Var data assimilation, as depicted in Fig. 1.

4. Assessment of ANN retrieval method in the 1D-Var analysis

a. Simulation framework

Since ANN coefficients are dependent upon latitudinal regions, 4034, 3977, and 4008 validation samples are used to validate the ANN model at $90^\circ\text{--}30^\circ\text{N}$, $30^\circ\text{N--}30^\circ\text{S}$, and $30^\circ\text{--}90^\circ\text{S}$ regions, respectively. But, because region-dependent results are similar to each other, we provide individual results in the supplements (Figs. S1, S2, and S3 in the online supplemental material).

The optimized CTP and CF pairs from the MR method and the ANN method are compared with the optimum CTP and CF validation dataset of 12 019 samples (Fig. 10). The optimized CTPs from the MR method show a bias of -119 hPa, RMSE of 290 hPa, and a correlation coefficient of 0.66 with the optimum CTPs. Negative bias suggests that the optimized CTPs from the MR method are likely smaller (i.e., higher cloud top) than the optimum CTPs. The density histogram of CTP for the ANN method (Fig. 10c), on the other hand, shows better agreement with the optimum CTPs, with 0.87, -19 hPa, and 169 hPa for correlation coefficient, mean bias, and RMSE, respectively.

For the CF comparison, results are presented in the form of the mean difference from the optimum CF, in the same CTP coordinates as in Figs. 10a and 10c. The results showing generally higher cloud tops tend to have smaller cloud amounts. Opposite behaviors are found in the case of underestimates of the cloud top height. The general patterns shown in the difference between the ANN method and the optimum approach (Fig. 10d) are similar to results from the MR method (Fig. 10b). However, considering that the majority of frequencies of CTPs from the ANN method are located along a diagonal line (in Fig. 10c) where the CF difference appears to be small, CFs from the ANN method should be much more similar to the optimum CFs. In conclusion, biases in both cloud top height and cloud fraction shown in the MR method are substantially removed by the ANN

FIG. 9. Structure of the ANN model showing an input layer with 254 variables (182 IASI brightness temperatures, 43-level background profiles of temperature and humidity, surface temperature, skin temperature, and surface humidity), a single hidden layer with five neurons, and the target output (CTP).

method, resulting in CTPs as well as CFs that are closer to the optimum values.

It is noted that the use of optimum cloud parameters yields the best 1D-Var results under the assumption of the single-layer graybody cloud. Therefore, we expect that assignment of CTP and CF closer to the optimum parameters using the ANN method would give better analysis results (i.e., temperature and humidity profiles). To test this hypothesis, a 1D-Var analysis (shown in the second stage of the flow in Fig. 1) was performed for all 12 019 validation cases. The resultant statistics are from all cases regardless of the convergence state, including nonconverged cases that are replaced by corresponding background profiles. Convergence statistics are given in Table 3, in which all 4214 cases for clear 1D-Var and 7805 cases for cloudy 1D-Var analysis are found among 12 019 cases for the optimum cloud parameters. Nonconvergence cases are not shown in the optimum cases, because the optimum cloud parameters were chosen only if the convergence criteria were met in the 1D-Var analysis. From the MR method, 3211 and 7340 cases were found to be converged in clear and cloudy 1D-Var, respectively. The remaining 1468 cases were not converged and thus were replaced by their background state. Meanwhile, from the ANN method, 3516 and 7547 cases for clear and cloudy 1D-Var were converged, with 956 nonconverged cases replaced by their background state. These statistics suggest that the ANN method yields more converged cases in the 1D-Var analysis. As the ANN method gives CTPs closer to their

optimum values, retrieved CFs are expected to become closer to their optimum values as well, according to Eq. (2). Conversely, better cloud retrievals mean better clear scene determination, which may lead to more converged cases over clear skies.

Channel selection for the 1D-Var analysis was performed by using the optimized cloud parameters from the MR method and the ANN method, and by using the optimum cloud parameters for the 12 019 validation samples (Fig. 11). It is shown that more channels are used by the ANN approach over CO₂, water vapor, and window bands, compared to the channels selected by the MR method (Fig. 11a). The ANN method appears to allow the 1D-Var analysis to use more cloud-affected channels because of the larger number of converged cases in both clear and cloudy 1D-Var, compared to the MR method. In particular, lower cloud tops by the ANN method induce higher cloud amounts, which likely enable cloud-affected channels to be more available in the 1D-Var analysis. Considering that the cloudy-sky data assimilation is about to use channels not contaminated by cloud presence, more upper-level temperature sensitive channels can be available in the 1D-Var analysis because of the lower cloud tops.

The $O - B$ means of selected channels by three methods are given in Fig. 11b. It shows that all three methods gave nearly the same $O - B$ means for CO₂ channels except lower-level picking CO₂ channels (i.e., from 120 to 135 in the channel number). The $O - B$ mean values in the window channels show weak positive

FIG. 10. (left) Two-dimensional histograms of frequencies of (a) optimum CTP vs optimized CTP from the MR method (MR CTP), and (c) optimum CTP vs optimized CTP from the ANN method (ANN CTP). Color represents data count. (right) Mean CF difference (b) between optimum and the MR method, and (d) between optimum and the ANN method. The mean differences are given in the same CTP pressure coordinates as in the left panels.

biases for the ANN method and the optimum cloud parameters. Overall, it is noted that the ANN method accommodates more cloud-affected channels than the MR method while results in quite similar $O - B$ means.

The error statistics of the 1D-Var analysis from the MR method, the ANN method, and the optimum cloud parameters are shown in Fig. 12. It is clear that the ANN method resulted in reduced biases in both temperature and humidity profiles, compared to results from the MR method. The temperature RMSE was also improved by the ANN method throughout the entire layer, particularly in the midtroposphere between 400 and 700 hPa. The RMSE profiles by the ANN method tends to be closer to those from the use of the optimum cloud parameters. However, the error profiles of humidity show that the ANN method gave RMSE profiles hardly discernible from the results from the MR method, although reduced bias by the ANN method is noted. The smaller

improvement in the water vapor field by the ANN method may be due to the smaller portion of the used water vapor channels in the 1D-Var analysis. Relatively larger $O - B$ over the water vapor band (as shown in Fig. 8) might also contribute to such smaller improvement. This seems to be characteristics of the current

TABLE 3. Number (percentage) of converged and nonconverged cases from three methods in the 1D-Var analysis using the simulation dataset.

Method	Converged cases (cloudy 1D-Var)	Converged cases (clear 1D-Var)	Nonconverged cases
Optimum CTP&CF	7805 (65%)	4214 (35%)	—
MR method	7340 (61%)	3211 (27%)	1468 (12%)
ANN method	7547 (63%)	3516 (29%)	956 (8%)

FIG. 11. (a) Difference in the number of used measurements (ANN method minus MR method) at each of the 182 IASI channels. (b) Mean $O - B$ for the optimum (green), the MR method (blue), and ANN method (red). Black arrows in (b) represent IASI channels in CO_2 , window, and water vapor absorption bands.

IASI 1D-Var assimilation, in which the temperature signal from the CO_2 channels is heavily weighted, due to the larger uncertainty of the background water vapor fields (Hilton et al. 2012). By the same reasons, significant differences are not shown in the humidity field even if the optimum cloud parameters are used. Nevertheless, it can be concluded that the ANN method tends to improve the error statistics in the 1D-Var analysis by selecting more channels, compared to the MR method.

b. Experiments with the UM NWP system

To examine the impact of the ANN method in the UM weather forecasts, we implemented the ANN method in the UM OPS. Two global model assimilation trials including control and experiment runs were conducted for a period from 15 July to 13 August 2017. The UM model version 10.2 at a resolution of 25 km was used, and the variational bias correction (VarBC) scheme described in Auligné et al. (2007) was applied in both trial experiments. Here, the trial run with the MR method is referred to as the control run while the trial run with the ANN method is referred to as the experiment run.

The optimum CTPs were obtained in the simulation framework because RMSEs of temperature and humidity profiles can be calculated against the given truth profiles. However, it is practically impossible to define optimum cloud parameters in the operational data

assimilation system, because the true atmospheric states are not known. Thus, we simply compare the optimized CTPs from the MR method and the ANN method, in order to confirm whether the CTP results are consistent with the results from the simulation framework. In Fig. 13a, the scattergram of CTPs from the UM OPS is shown for 0000 UTC 30 July 2017, which is from 109 155 samples over open ocean. It is shown that a majority of CTPs retrieved by the ANN method are lower than CTPs from the MR method. Although these UM OPS results appear to be consistent with results from the simulation dataset, the MR method produces cloud tops excessively higher than those from the ANN method. Because of the higher cloud tops, the MR method should have produced smaller CFs, as demonstrated in Fig. 13b.

In Table 4, among the total 109 155 samples at 0000 UTC 30 July 2017, converged cases for clear and cloudy cases are 24 626 and 54 794, respectively, in the control run, whereas 13 944 and 66 251 cases in the experiment run were converged. Therefore, a total of 79 420 and 80 195 scenes were used in the control run and the experiment run, respectively. It is interesting to note that converged clear-sky cases are more in the control run (i.e., MR method), in contrast to the result that both converged clear-sky and cloudy-sky cases are more for the ANN method in the simulation framework.

FIG. 12. Mean bias (solid lines) and RMSE (dashed lines) of (a) temperature and (b) humidity analysis profiles from the 1D-Var analysis with the use of the MR method (blue), and ANN method (red), and the optimum values (green). Black dotted lines in the RMSE profiles represent the RMSE profiles of the background state.

This may be due to the bias correction process in the UM OPS, which will affect the IASI radiances and background states used in the UM OPS. As noted in Fig. 13b, the MR method tends to yield substantially less cloud amounts when preprocessed IASI radiances and background states are used, likely leading to more clear-sky scenes employed in the assimilation. Overall, we have identified that more cases converged by the ANN method (corresponding to an increase of about 1% compared with the MR method), due to more cloudy scenes used in the data assimilation with the ANN method.

We also examine the difference in the number of selected IASI channels and in the $O - B$ means of used channels in the data assimilation between two trials, using the same 109 155 samples (Fig. 14). The use of the ANN method in the UM OPS resulted in more CO_2 channels for the data assimilation process is shown in Fig. 14a, compared to the MR method results. However, the $O - B$ means over the CO_2 band appear similar to each other between the two methods (Fig. 14b). This is consistent with the result from the simulation framework (Fig. 11) that showed a common feature that the ANN method used more cloud-affected channels by lowering the cloud top in spite of the same $O - B$ means. On the other hand, the ANN method resulted in less window channels used and larger negative $O - B$ means, compared to the MR method results. It may be due to

the tendency that the ANN method produces lower cloud tops and larger cloud fractions. Under the assumption of the graybody single layer cloud, the same radiance can be interpreted as lower cloud top and larger cloud fraction, instead of higher cloud top and smaller cloud fraction. Therefore, the ANN method probably assigns the clear-sky scenes that were determined by the MR method as cloudy scenes, which effectively prevents the use of window channels.

To assess the impact of the ANN method on the initial analysis field of the UM operational assimilation system, we examined analysis fields collected during the two-week period from 30 July to 13 August 2017. More than 535 global daily radiosonde observations were used over the 14-day period and errors were calculated by comparing analysis fields with the radiosonde observations. In the operational UM OPS, instead of performing the 1D-Var analysis process, the selected IASI channels are passed to the 4D-Var assimilation. Thus, in this study, the analysis results at the $T + 0$ forecast time from the 4D-Var system were evaluated, to examine the 1D-Var analysis results from each scheme in the UM OPS. Figure 15 shows RMSEs of temperature and relative humidity of analysis at the $T + 0$ UM global forecast time. The experiment run shows RMSEs in temperature and relative humidity analysis profiles were slightly reduced, compared to the control run. The RMSE results of temperature and relative humidity for each latitudinal

FIG. 13. (a) Two-dimensional histograms of frequencies of ANN CTP vs MR CTP in the UM OPS system. Color represents data count. (b) Mean difference between MR CF and ANN CF in the UM OPS system. The mean differences are given in the same CTP pressure coordinates as in (a).

band are shown in Figs. S1 and S2, respectively. It can be seen that the temperature RMSEs in the Southern Hemisphere, and the relative humidity RMSEs in all latitude bands are reduced in the experiment run. However, improvement in the initial fields in the UM forecasts by the ANN method appears to be rather minor, leading to a conclusion that the overall impact on global weather forecasting of the ANN method employed in the assimilation system is found to be neutral.

5. Summary and discussion

In the IASI data assimilation, there have been efforts of directly using cloud-affected (or not clear) IASI radiances, instead of indirect use like removing cloud-contaminated channels or constructing “presumed” clear-sky radiances. In the Cloudy 1D-Var method employed by the IASI 1D-Var data assimilation system, the first guess pair of CTP and CF is assigned from infrared radiances of 10 IASI channels and model background field by minimizing the residual between observed and simulated channel radiances. In this approach, clouds are assumed to be a single layer behaving like a gray-body. Although the Cloudy 1D-Var method led to a better analysis performance, compared to the clear-sky only 1D-Var or McNally and Watts (2003) method, it is difficult to evaluate the accuracy of obtained cloud properties or examine how we improve the cloud retrieval capability in the Cloudy 1D-Var method.

In this study, we evaluated the cloud retrieval in the Cloudy 1D-Var method (i.e., MR method). Instead of comparing retrieved cloud properties with observations, retrievals were compared against the optimum cloud parameters that yield the best 1D-Var analysis results,

given IASI observations and model background field. For the comparison, atmospheric temperature and humidity profiles as well as cloud information from the ECMWF forecasts were considered to be the truth, from which IASI radiances and background fields were generated after taking their respective errors into account. The comparison indicated that cloud retrieval from the MR method tends to overestimate the cloud top height and thus tends to underestimate the cloud fraction. The overestimation brought in less active use of cloud-affected radiances, necessitating the improvement of the cloud retrieval method by correcting this overestimation of the cloud top height.

To improve the cloud retrieval, we developed a method of resembling the optimum cloud parameters as closely as possible. In doing so, we utilized an artificial neural network approach to train the inputs (IASI 182 channel radiances and model background fields) to produce cloud parameters similar to the optimum values (i.e., ANN method). It was shown that the ANN method produces cloud tops closer to the optimum values, compared to the results from the MR method. It was also found that corresponding CFs are in better agreement with optimum values. Moreover, it was noted that

TABLE 4. Number (percentage) of converged and nonconverged cases from the MR and ANN methods in the 1D-Var analysis of the UM OPS (0000 UTC 30 Jul 2017).

Method	Converged cases (cloudy 1D-Var)	Converged cases (clear 1D-Var)	Nonconverged cases
MR method	54 794 (68%)	24 626 (31%)	1173 (1%)
ANN method	66 251 (82%)	13 944 (17%)	398 (1%)

FIG. 14. (a) Difference in the number of used measurements (ANN method minus MR method) at each of the 182 IASI channels. (b) Mean $O - B$ for the optimum (green), MR method (blue), and for ANN method (red). Black arrows in (b) represent IASI channels in CO₂, window, and water vapor absorption bands.

the ANN method gave more converged cases in the 1D-Var analysis. It is believed that the ANN method allows the use of more cloud-affected channels, with similar $O - B$ values, and reduced RMSEs in the 1D-Var analysis.

The impact of the ANN method on the 1D-Var analysis was examined within the UM OPS by taking experiments over 30 days (from 15 July to 13 August

2017). In these experiments, retrieved cloud tops from the ANN method were generally lower than in the control run with the MR method. Additionally, more CO₂ channels were selected, with $O - B$ means nearly the same as found from the MR method. In spite of less channels used by the ANN method in the UM OPS, the overall convergence cases by the ANN method was 1% more than by the MR method. Furthermore, analysis

FIG. 15. RMSE profiles of (a) temperature and (b) relative humidity in the UM OPS analysis with the MR method (blue) and ANN method (red).

results for the ANN method at $T + 0$ forecast time is found to be neutral. Thus, the use of more channels (mostly window channels) by the MR method may not be always beneficial if more converged cases and the neutral analysis impact by the ANN method are considered. Given that the ANN method gives more converged cloudy cases, it tends to compensate less window channels with more CO_2 and water vapor channels.

Even if the impact is near neutral at the $T + 0$ analysis in the UM OPS, the implementation of the ANN method should be beneficial because the ANN method can substantially reduce the computational burden to search for an initial guess pair of CTP and CF. To find a CTP and CF pair showing a minimum residual, the MR method needs to calculate at least 27 CTP and CF pairs (from 100 hPa level to the surface level) and corresponding residuals. By contrast, in the ANN method, CTP is directly retrieved with ANN-derived coefficients, and then CF is determined using Eq. (2). Therefore, if the ANN method is introduced to the IASI 1D-Var data assimilation system, the computational time required for finding initial cloud parameters can be reduced to roughly 1/54 level (approximately 1.85%) of what required for the MR method. Furthermore, the ANN method may avoid problems of possible multiple solutions in the MR method because the CTP retrieval is independent of CF in the ANN method.

On the other hand, the impact of the ANN method on the 1D-Var analysis appears neutral; analysis results are not much discernable from those by the MR method. It is likely caused by suboptimal observation and background error covariance matrices (i.e., **R** matrix and **B** matrix) employed in the UM OPS, which represent the clear-sky field of view. As shown in Figs. 2 and 3, both cloud retrieval methods find retrieved cloud parameters likely located along the line showing approximately an inverse relationship between CTP and CF. Concerning the MR method, which tends to overestimate the cloud height, retrieved CF should be smaller than optimum values. Thus, the MR method will likely identify the cloud-affected scenes more often to be clear because the criteria determining cloud presence is $\text{CF} > 0.05$. By contrast, the ANN method will more likely identify the same cloud-affected scenes to be cloudy, which are closer to optimum values. Since the ANN method allows more cloud-affected IASI measurements, as seen in the results, there will be more added CO_2 channels. However, low-level cloud-affected channels by the ANN method seem to be rejected during the data assimilation process. Note that observation and radiative simulation of channel radiances are subject to larger errors for cloud-sky cases, in comparison to the clear-sky cases. In particular, it is generally known that the

radiative transfer simulation for the cloud-sky scenes creates much larger error than the clear-sky simulation. These can cause less use of lower-level peaking channels and window channels, compared to MR method results. Thus, even if the ANN method employs more cloud-affected scenes, those may not be fully utilized in the assimilation, giving near-neutral results probably by adding CO_2 channels but losing window channels. Thus, one way to fully accommodate the ANN-retrieved CTP and CF in the UM OPS may be either to use observation and background error covariance matrices for the cloudy-sky scene or to allow larger uncertainty errors for the cloud-affected channels in the UM OPS. Further studies should be done along the line of those directions, to actively use cloud-affected IASI scenes in the data assimilation.

Acknowledgments. This study was carried out through a collaboration between the UKMO and the Korea Meteorological Administration (KMA). This work was funded by the Space Core Technology Development Program (NRF-2018M1A3A3A02065661) and by the Korea Meteorological Administration Research and Development Program under Grant KMIPA KMI2018-06910.

REFERENCES

- Auligné, T., A. P. McNally, and D. P. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **133**, 631–642, <https://doi.org/10.1002/qj.56>.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.
- Desai, K. M., S. A. Survase, P. S. Saudagar, S. S. Lele, and R. S. Singhal, 2008: Comparison of Artificial Neural Network (ANN) and Response Surface Methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan. *Biochem. Eng. J.*, **41**, 266–273, <https://doi.org/10.1016/j.bej.2008.05.009>.
- Eresmaa, R., and A. P. McNally, 2014: Diverse profile datasets from the ECMWF 137-level short-range forecasts, version 1.0. EUMETSAT Satellite Application Facility (NWP SAF) Tech Doc. NWPSAF-EC-TR-017, 14 pp., <https://nwpsaf.eu/oldsite/reports/nwpsaf-ec-tr-017.pdf>.
- Eyre, J. R., and W. P. Menzel, 1989: Retrieval of cloud parameters from satellite sounder data: A simulation study. *J. Appl. Meteor.*, **28**, 267–275, [https://doi.org/10.1175/1520-0450\(1989\)028<0267:ROCPFS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1989)028<0267:ROCPFS>2.0.CO;2).
- Hilton, F., and Coauthors, 2012: Hyperspectral Earth observation from IASI: Five years of accomplishments. *Bull. Amer. Meteor. Soc.*, **93**, 347–370, <https://doi.org/10.1175/BAMS-D-11-00027.1>.
- Joo, S., J. Eyre, and R. Marriott, 2013: The impact of MetOp and other satellite data within the Met Office global NWP system using an adjoint-based sensitivity method. *Mon. Wea. Rev.*, **141**, 3331–3342, <https://doi.org/10.1175/MWR-D-12-00232.1>.
- Li, J., C. Y. Liu, H. L. Huang, T. J. Schmit, X. Wu, W. P. Menzel, and J. J. Gurka, 2005: Optimal cloud-clearing for AIRS radiances using MODIS. *IEEE Trans. Geosci. Remote Sens.*, **43**, 1266–1278, <https://doi.org/10.1109/TGRS.2005.847795>.

- McNally, A. P., and P. D. Watts, 2003: A cloud detection algorithm for high-spectral-resolution infrared sounders. *Quart. J. Roy. Meteor. Soc.*, **129**, 3411–3423, <https://doi.org/10.1256/qj.02.208>.
- , —, J. A. Smith, R. Engelen, G. A. Kelly, J. N. Thépaut, and M. Matricardi, 2006: The assimilation of AIRS radiance data at ECMWF. *Quart. J. Roy. Meteor. Soc.*, **132**, 935–957, <https://doi.org/10.1256/qj.04.171>.
- McNally, T., M. Bonavita, and J. N. Thépaut, 2014: The role of satellite data in the forecasting of Hurricane Sandy. *Mon. Wea. Rev.*, **142**, 634–646, <https://doi.org/10.1175/MWR-D-13-00170.1>.
- Pavelin, E. G., S. J. English, and J. R. Eyre, 2008: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **134**, 737–749, <https://doi.org/10.1002/qj.243>.
- Rumelhart, D. E., J. L. McClelland, and PDP Research Group, 1986: *Parallel Distributed Processing*. Vol. 1, *Explorations in the Microstructure of Cognition: Foundations*, MIT Press, 567 pp.
- Wylie, D. P., and W. P. Menzel, 1999: Eight years of high cloud statistics using HIRS. *J. Climate*, **12**, 170–184, <https://doi.org/10.1175/1520-0442-12.1.170>.