ABSTRACT: The assimilation of hyperspectral infrared sounders (HIS) observations aboard Earth-observing satellites has become vital to numerical weather prediction, yet this assimilation is predicated on the assumption of clear-sky observations. Using collocated assimilated observations from the Atmospheric Infrared Sounder (AIRS) and the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP), it is found that nearly 7.7% of HIS observations assimilated by the Naval Research Laboratory Variational Data Assimilation System–Accelerated Representer (NAVDAS-AR) are contaminated by cirrus clouds. These contaminating clouds primarily exhibit visible cloud optical depths at 532 nm (COD$_{532\text{nm}}$) below 0.10 and cloud-top temperatures between 240 and 185 K as expected for cirrus clouds. These contamination statistics are consistent with simulations from the Radiative Transfer for TOVS (RTTOV) model showing a cirrus cloud with a COD$_{532\text{nm}}$ of 0.10 imparts brightness temperature differences below typical innovation thresholds used by NAVDAS-AR. Using a one-dimensional variational (1DVar) assimilation system coupled with RTTOV for forward and gradient radiative transfer, the analysis temperature and moisture impact of assimilating cirrus-contaminated HIS observations is estimated. Large differences of 2.5 K in temperature and 11 K in dewpoint are possible for a cloud with COD$_{532\text{nm}}$ of 0.10 and cloud-top temperature of 210 K. When normalized by the contamination statistics, global differences of nearly 0.11 K in temperature and 0.34 K in dewpoint are possible, with temperature and dewpoint tropospheric root-mean-squared errors (RMSDs) as large as 0.06 and 0.11 K, respectively. While in isolation these global estimates are not particularly concerning, differences are likely much larger in regions with high cirrus frequency.

KEYWORDS: Remote sensing; Satellite observations; Data assimilation

1. Introduction

Since their first operational use for numerical weather prediction (NWP) in 2002, data from hyperspectral infrared sounder (HIS) sensors, such as the Infrared Atmospheric Sounding Interferometer (IASI; Hilton et al. 2012) and the Cross-track Infrared Sounder (CrIS; Bloom 2001), have greatly expanded to become some of the most valuable satellite observations in terms of forecast impact, particularly for assimilation of clear-sky observations (e.g., see Fig. 1 in Marquis et al. 2021, hereafter M21). Aware of the overwhelmingly positive impact, the World Meteorological Organization (WMO) has recommended the placement of at least five HIS sensors in geostationary for regular full global coverage particularly for NWP efforts (WMO 2019). Given these recommendations and current planned launches of HIS sensors, it is likely that NWP dependence on HIS radiance assimilation will continue to grow.

While the assimilation of infrared (IR) radiances from cloudy scenes is currently under development at major NWP centers (e.g., Geer et al. 2018), this assimilation is mostly limited to overcast scenes, rudimentary treatment of cloud properties, and the assimilation of few spectral bands (McNally 2009; Okamoto 2013). More recently, techniques of cloud clearing, that is, assimilating clear-sky equivalent radiances from cloudy skies, have been developed and implemented in some modeling systems, though largely in testing and not operational settings (e.g., Reale et al. 2018; Li et al. 2022). As such, current HIS radiance assimilation at operational centers is largely dominated by scenes which presume assimilated observations are cloud- and aerosol-free. To ensure that assimilated observations are from clear sky, several observation tests are typically performed. For instance, the Naval Research Laboratory Atmospheric Variational Data Assimilation System–Accelerated Representer (NAVDAS-AR; Xu et al. 2005) implements cloud detection similar to that described by McNally and Watts (2003) where observations with innovations (i.e., observed radiances minus radiiances simulated from a background atmospheric profile) above 3 times the observational error (typically 1–3 K) are
rejected. Yet, such a system, or even newer methods that may use imager data to assist in cloud detection (e.g., Eresmaa 2014), is likely to be imperfect as optically thin clouds and aerosols remain present even in products with rigorous cloud screening using both visible and IR bands (e.g., Toth et al. 2013). For satellite-based retrievals of sea surface temperatures (SSTs), Marquis et al. (2017) found approximately 25% of Moderate Resolution Imaging Spectroradiometer (MODIS) assimilation-quality retrievals in the tropics were contaminated by cirrus cloud resulting in a 0.3–0.5 K cold bias. Contamination by aerosol plumes with visible optical depths above 0.30 has been reported in over 16% of assimilated HIS observations (M21).

Cirrus clouds are of particular concern for HIS radiance assimilation due to frequent global occurrence—particularly over regions of tropical cyclogenesis (e.g., Ramanathan and Collins 1993; Clement et al. 2005; Benestad 2009). For instance, Mace et al. (2009) report cirrus clouds in 40%–60% of global observations from satellite-based lidar. Of these clouds, roughly half can be classified as optically thin, exhibiting visible cloud optical depths at 532 nm (hereafter COD532nm) below 0.30 (Sassen and Cho 1992; Campbell et al. 2015). Using observations of cirrus clouds from the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP; Winker et al. 2010) instrument aboard NASA’s CALIPSO satellite, the rate of cirrus contamination in observations assimilated by NAVDAS-AR from the Atmospheric Infrared Sounder (AIRS; Aumann and Pagano 1994) bands 253 and 318 (corresponding to wavelengths of 13.85 and 13.49 μm, respectively) are estimated. Since these bands peak in the troposphere, any cirrus cloud present would indicate a contaminated observation that should not have been assimilated. The contamination statistics for these two bands for the period of June through August 2018 are shown in Fig. 1. Both bands experience residual cirrus in 7.7% of assimilated observations with the majority of the contaminating observations exhibiting COD532nm below 0.20. Since other studies report cirrus contamination of passive radiometric products near 25% (e.g., Ackerman et al. 2008; Marquis et al. 2017) these findings suggest HIS screening of optically thin cirrus clouds performs better. That is, tolerances in data thinning and cloud screening (McNally and Watts 2003; Eresmaa 2014) appear more rigorous for HIS assimilation. Consistent with the SST contamination statistics reported by Marquis et al. (2017), and the propensity for cirrus cloud occurrence rates globally (e.g., Campbell et al. 2015), relative sample contamination frequency shown in Fig. 1 increases exponentially at thinner CODs. Nearly all residual clouds exhibit cloud tops of 245 K or cooler, with 86% exhibiting tops between 190 and 220 K, consistent with the properties of cirrus clouds (Campbell et al. 2015).

Given the contamination of assimilated HIS radiances by cirrus clouds demonstrated above, it is important to examine the impacts associated with assimilating cirrus-contaminated observations. Here we follow a methodology similar to that reported by M21, in which the impact of assimilating dust-contaminated HIS radiances on analyzed temperature and humidity profiles was quantified through including/excluding dust aerosol in the forward model using synthetic HIS observations. Note that using lidar observations of dust, M21 found 16.2% of NAVDAS-AR assimilated HIS observations were collected from scenes exhibiting visible aerosol optical depths above 0.3, resulting in analysis differences above the expected background uncertainty. While the contamination rate of aerosols is much higher than for cirrus clouds, cirrus clouds will always absorb and emit in the IR spectrum unlike all types of observed aerosols. Additionally, cirrus clouds occur at much higher altitudes and are therefore much colder, presenting much stronger thermal contrast with the lower and middle troposphere than that of most tropospheric aerosols that are of concern for HIS assimilation.

In this study, the impact of assimilating cirrus cloud-contaminated HIS radiances is estimated. Specifically, the analysis difference associated with assimilating cirrus-contaminated HIS observations is estimated using a one-dimensional variational assimilation (1DVar) system. Finally, the estimated analysis difference is compared to typical background uncertainties and methods for mitigation are suggested.

2. Models and methodology

Cirrus-induced differences in temperature and humidity are estimated using the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) 1DVar, version 1.2 (Havemann 2020), system coupled with the Radiative Transfer for TOVS (RTTOV) model, version 12.3 (Saunders et al. 2018), for forward and Jacobian radiative transfer. Specifically, temperature and humidity differences are estimated by performing assimilation of HIS observations contaminated by cloud. To ensure correct system configuration and no biases exist in the system, an experiment including cloud properties in the RTTOV forward and Jacobian models is performed using modified 1DVar source code.

Ideally, cirrus-induced differences would be estimated using observed cloud and atmospheric profiles and collocated HIS observations. However, several factors limit such an analysis. For instance, and at a minimum, collocated radiosonde, HIS observations, and preferably lidar observations of clouds are required. Yet, assimilation of many HIS channels, particularly those with lower-tropospheric sensitivity and those most likely to experience impacts due to residual cirrus, are typically operationally limited to over-ocean observations where radiosonde observations are largely unavailable. As such, collocations of meteorological, cloud, and HIS observations are mostly unavailable. Further, even if all observations

\[^{1}\text{In particular, NAVDAS-AR screens for clouds by separating observed radiances into those most sensitive to carbon dioxide (CO}_2\text{ region) and those most sensitive to water vapor (H}_2\text{O region). HIS channels in the CO}_2\text{ region are ranked according to the channel’s sensitivity to cloud and altitude, with those sensitive to high clouds ranked highest and those channels sensitive to low clouds ranked lowest. If a CO}_2\text{ channel fails the innovation check, that channel and all lower channels are rejected. However, for channels in the H}_2\text{O region, only the innovation of a channel near the atmospheric window (10.4 μm) is checked. If that window region band fails, the innovation checks for all H}_2\text{O channels are rejected.}
\]
required can be collocated, representativeness errors will exist due to the HIS observational footprint being on the order of 10 km², whereas lidar and meteorological profile observations are generally point observations. As such, uncertainty associated with non-homogeneity of the surface, cloud, and meteorology would be present in an analysis based solely upon observations.

In this study, we instead choose to predefine the meteorological profile and simulate cirrus-contaminated radiances using RTTOV v12.3. Thus, not only do we know the true meteorological profiles, but, because observations are synthetic, we can examine the radiative impact of a wide range of cirrus clouds. The cirrus-induced difference is determined by performing assimilation using the 1DVar system with cloud-contaminated HIS observations and comparing the retrieved temperature and humidity analysis to the true profile used to create the synthetic observations. Note, in the case where a simulation that accounts for cloud is able to retrieve approximate truth, this difference is identical to a difference calculated between a cloud-permitting analysis and a clear-sky assumption analysis as performed for aerosol in M21. Note that testing revealed that the system is able to retrieve the true atmosphere if cloud properties are provided; thus, any differences are due to cloud contamination. Finally, bulk differences are estimated by integrating the contamination statistics presented in Fig. 1 with respect to the estimated differences.

a. Synthetic observations and cirrus impact to HIS radiances

Following M21, synthetic CrIS HIS observations are simulated. The CrIS sensor is chosen due to its current and future inclusion on Joint Polar Satellite System satellites (Goldberg et al. 2013). To ensure that a representative atmosphere...
associated with an observation is known, synthetic CrIS observations for the 1305 CrIS nominal spectral resolution (NSR) bands are created using RTTOV v12.3 and a standard tropical atmosphere from McClatchey et al. (1972) interpolated to the 54 pressure levels defined in the default background error covariance matrix used (explained below). For RTTOV simulations, the Discrete Ordinate Radiative Transfer model (DISORT) method (Stamnes et al. 1988) with eight streams and a cloud fraction of 1.0 within the cloudy layer is implemented. Clouds are simulated by providing cloud optical properties at each vertical level and wavelength. Otherwise, the RTTOV simulations use default options. This interpolated temperature and moisture profile is shown in Fig. 2.

Cirrus clouds are simulated from the first model level with temperature below 233 K up to the tropopause, indicated by horizontal dashed lines in Fig. 2. For each of the seven model layers between these boundaries, a cloud is placed within that single model level with the cloud’s COD532nm varying from 0.01 to 0.30 in 0.01 increments and from 0.35 to 1.0 in 0.05 increments. Thus, a total of 302 synthetic observations are created corresponding to 43 COD532nm for each of the 7 vertical layers, and 1 clear-sky observation. Note, due to the stretched vertical grid, the cloud physical thickness varies slightly with altitude (854 m at a cloud-top pressure of 236 hPa to 1021 m at a cloud-top pressure of 97.2 hPa). To simulate each cloud, we provide RTTOV with cloud optical properties of absorption coefficient, scattering coefficient, and a 30-term Legendre expansion of the phase function from the Yang et al. (2013) and Bi and Yang (2017) optical model (explained below) for each of the 1305 CrIS NSR bands at each model level. Since cloud optical properties are a function of ice effective radius, we determine the effective radius of our cirrus clouds using cloud-top temperature in the Dolinar et al. (2022) parameterization for cirrus clouds in the tropics.

Finally, we estimate the impact of cirrus clouds on the observed HIS radiances by comparing clear-sky synthetic radiances using the McClatchey et al. (1972) atmosphere with the synthetic radiances with cirrus clouds. The magnitude of these differences is examined for different cloud-top temperatures (i.e., effective radii; discussed below) and COD532nm. The magnitude of differences is then compared to the contamination statistics reported in Fig. 1 to examine thresholds where a cirrus cloud is most likely to pass screening.

b. Cloud optical properties

To simulate clouds, an eight-column aggregate with severely roughened surface ice optical model (hereafter 8CASM) from Yang et al. (2013) and Bi and Yang (2017) is used. This model represents what is characterized here as the current community standard for ice optical parameterizations presenting consistent results in active- and passive-based retrievals and polarized observations (Holz et al. 2016). The model was provided to the authors as geometric and optical properties of single ice crystals in radii ranging from 0.2 μm to 1 mm and wavelengths from 0.2 to 100 μm. Consistent with the methodology used to apply the 8CASM model for the Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 6 cloud property retrieval (Platnick et al. 2017), we apply the following gamma particle size distribution:

\[
n(r_{vp}) = n_{vp} \frac{1}{\left(3\sigma^2/vp\right)^{1.5}} \exp\left(-\frac{r_{vp}}{\sigma_{vp}}\right).
\]

where \(n(r_{vp})\) is the number of particles with ice crystal radius \(r_{vp}\), \(\sigma^2\) is the variance of the gamma distribution [assumed to be 0.10 consistent with Platnick et al. (2017)], and \(r_{eff}\) is the effective radius of the size distribution. Note that here we use the radius of an equivalent spherical ice crystal \(r_{eq}\), which is defined using the nonspherical ice crystal’s true volume \(V\) and projected area \(A\) due to ambiguities in the radius of nonspherical ice crystals (Fu et al. 1999):

\[
r_{eq} = \frac{3V}{4A}.
\]

Note, while the radius of nonspherical ice crystals may be ambiguous, this radius is not physically used in Eq. (2), only the volume and area of ice crystals.

With the ice crystal equivalent radius \(r_{eq}\) for any ice crystal defined, Eq. (1) can be applied to find the relative number of ice crystals for any user defined ice crystal effective radius \(r_{eff}\). Using the size distribution effective radius as defined in Dolinar et al. (2022), the relative number of ice crystals at any given size can be determined. Finally, this relative number curve can be applied to find the extinction, scattering, and absorption efficiencies.
The extinction, scattering, and absorption efficiencies for the gamma size distribution are calculated for each wavelength following Baum et al. (2011). Specifically,

\[
Q_{\text{ext,eff}} = \frac{\int_{r_{\text{min}}}^{r_{\text{max}}} Q_{\text{ext}}(r_{\text{vp}}) A(r_{\text{vp}}) n(r_{\text{vp}}) dr_{\text{vp}}}{\int_{r_{\text{min}}}^{r_{\text{max}}} A(r_{\text{vp}}) n(r_{\text{vp}}) dr_{\text{vp}}},
\]

where \(Q_{\text{ext,eff}}\) is the effective extinction efficiency of the size distribution and \(Q_{\text{ext}}(r_{\text{vp}})\) is the extinction efficiency of particles with equivalent radius \(r_{\text{vp}}\). Because it is desirable to know properties without knowing total number concentrations, the particle size distribution is normalized by the sum of the particle size distribution to get a relative number concentration and converted to summation form:

\[
Q_{\text{ext,eff}} = \frac{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} Q_{\text{ext}} A_i \Delta r_{\text{vp},i} n_i}{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} A_i \Delta r_{\text{vp},i} \sum_{j=r_{\text{min}}}^{r_{\text{max}}} n_j},
\]

where \(\Delta r_{\text{vp},i}\) is the step size for radius \(r_{\text{vp}}\) and \(r_{\text{min}}\) and \(r_{\text{max}}\) are the effective bounds of the size distribution for integration. Here we use minimum and maximum bounds for 1 μm and 1.5 mm, respectively. While these bounds can cause inaccurate optical properties at very small or very large ice crystals, those examined in this study range from approximately 10 to 80 μm and are unlikely to be impacted significantly by these bounds.

Likewise, the scattering and absorption efficiencies are given by

\[
Q_{\text{sca,eff}} = \frac{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} Q_{\text{sca}} A_i \Delta r_{\text{vp},i} n_i}{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} A_i \Delta r_{\text{vp},i} \sum_{j=r_{\text{min}}}^{r_{\text{max}}} n_j},
\]

and

\[
Q_{\text{abs,eff}} = \frac{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} Q_{\text{abs}} (1 - \text{ssa}) A_i \Delta r_{\text{vp},i} n_i}{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} A_i \Delta r_{\text{vp},i} \sum_{j=r_{\text{min}}}^{r_{\text{max}}} n_j},
\]

where ssa is the single scattering albedo for radius \(r_{\text{vp}}\) and index \(i\), and \(Q_{\text{sca,eff}}\) and \(Q_{\text{abs,eff}}\) are the bulk scattering and absorption efficiencies of the particle size distribution, respectively. The bulk single scattering albedo is simply

\[
\text{ssa}_{\text{eff}} = \frac{Q_{\text{sca,eff}}}{Q_{\text{sca,eff}} + Q_{\text{abs,eff}}} = \frac{Q_{\text{sca,eff}}}{Q_{\text{ext,eff}}}
\]

The bulk phase function is given by

\[
P(\theta)_{\text{eff}} = \frac{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} P(\theta)_{\text{i}} Q_{\text{sca,eff}} A_i \Delta r_{\text{vp},i} n_i}{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} Q_{\text{sca,eff}} A_i \Delta r_{\text{vp},i} \sum_{j=r_{\text{min}}}^{r_{\text{max}}} n_j},
\]

where \(P(\theta)_{\text{eff}}\) and \(P(\theta)\) are the bulk phase function magnitude and particle phase function magnitude for radius \(r_{\text{vp}}\) of index \(i\) at scattering angle \(\theta\), respectively. The bulk phase function is then converted to 30 Legendre expansion coefficients that are scaled using delta-M normalization (Wiscombe 1977).

While COD\(_{532\text{nm}}\) is defined at 532 nm for the simulations, COD\(_{532\text{nm}}\)—or more specifically extinction coefficient, \(\beta_{\text{ext,\_532nm}}\)—at all other wavelengths is determined using the ratio of extinction efficiencies between the desired wavelength and 532 nm:

\[
\beta_{\text{ext,\_532nm}} = \beta_{\text{ext,\_532nm}} \frac{Q_{\text{ext,i}}}{Q_{\text{ext,532nm}}},
\]

where COD\(_{532\text{nm}}\) is simply the vertical integration of extinction coefficient. Extinction coefficient is given by

\[
\beta_{\text{ext,\_532nm}} = \sum_{i=r_{\text{min}}}^{r_{\text{max}}} n_i \sigma_{\text{ext,i}}(\lambda),
\]

where \(\sigma_{\text{ext,i}}(\lambda)\) is the extinction cross section of ice crystals with equivalent radius \(r_{\text{vp}}\) at wavelength \(\lambda\). Where the extinction cross section is related to the extinction efficiency by

\[
Q_{\text{ext,eff}} = \frac{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} \sigma_{\text{ext,i}}(\lambda) A_i \Delta r_{\text{vp},i} n_i}{\sum_{i=r_{\text{min}}}^{r_{\text{max}}} A_i \Delta r_{\text{vp},i} \sum_{j=r_{\text{min}}}^{r_{\text{max}}} n_j},
\]

Absorption and scattering efficiencies are similarly related to the absorption and scattering cross sections following Eq. (11), respectively.

The extinction efficiencies \(Q_{\text{ext}}\) for each wavelength are determined using Eq. (4). With extinction coefficient (predefined at 532 nm) determined at all CrIS wavelengths using Eq. (9) and single scattering albedo known at each wavelength, the full optical properties required by RTTOV (absorption coefficient, scattering coefficient, and phase function expansion coefficients) can be provided to the 1DVar system for each wavelength. In other words, the cloud is physically consistent at all wavelengths—with the same ice water content (i.e., total mass), particle effective radius, particle number, and physical thickness.

\[c. \text{ The 1DVar system}\]

The 1DVar system requires the user to input observed radiances, a background temperature and moisture profile, a
background error covariance matrix, and an observational error variance matrix. Here, the observational error variance matrix for CrIS and the default background error covariance matrix included within the 1DVar package is used (Havemann 2020). The 54-level background error covariance matrix included in the 1DVar package defines temperature at 54 levels plus the surface, humidity in the lowest 26 levels plus the surface, and skin temperature. Similarly, as defined in the retrieval namelist, temperature in 54 levels plus the surface, humidity in the lowest 26 levels plus the surface, and skin temperature are retrieved. Note, the CrIS error variances include the RTTOV forward operator error and are provided in the 1DVar package (see section 3c in Havemann 2020). These observation error variances are on the order of 0.1–1.0 K, depending upon wavelength. The tropical standard atmosphere from McClatchey et al. (1972) is interpolated to the 54 pressure levels defined in the background error covariance matrix linearly by the natural logarithm of pressure for the background and true temperature and moisture profiles. This interpolated tropical standard atmosphere is used for the background/first-guess atmosphere provided to the 1DVar system. Additionally, this same interpolated atmosphere is used in RTTOV with the cloud properties discussed below to create the synthetic observations with RTTOV configured identically to that used in the 1DVar forward model call.

The 8CASM optical model described above has been implemented in the forward RTM (RTTOV) and Jacobian RTM (RTTOV-K) by providing the system with cloud optical properties of absorption coefficient, scattering coefficient, and Legendre expansion coefficients of the phase function for each wavelength and vertical layer. While the cloud optical properties are provided to the Jacobian model, cloud optical property Jacobian arguments are not set, causing the cloud properties to be treated as static variables within the Jacobian model (e.g., chapter 7.9 discussion on RTTOV-K in Hocking et al. 2019). That is, the temperature and humidity Jacobians will be influenced by the presence of cloud, but no cloud Jacobians exist, and the cloud optical properties are assumed constant between the background and analysis. In other words, the cloud field is error-free in this study. RTTOV options between those invoked by the 1DVar system and those used to create the synthetic observations are consistent. To limit the effects of solar radiation on the shorter wavelength observations, all simulations are performed with solar radiation turned off (i.e., only IR radiation). For Jacobian minimization, the Marquardt–Levenberg minimization option is used with a maximum of 20 iterations used for the Marquardt–Levenberg inner loop and a maximum of 10 outer loop iterations, as default for the 1DVar system. Cloud detection in the 1DVar system is set to false, and all other options are set to default as defined in the 1DVar User Guide (Havemann 2020). Finally, the surface is assumed to be seawater.

3. Results

To recognize the ability for certain clouds to pass cloud screening, the impact of different cirrus clouds on simulated CrIS brightness temperatures is presented in Fig. 3. For the clear-sky brightness temperatures (Fig. 3a), cooler brightness temperatures indicate wavelengths with higher-altitude sensitivity peaks, whereas warmer brightness temperatures indicate lower-tropospheric or surface-peaking wavelengths. For instance, wavelengths between 10 and 12 μm indicate lower-tropospheric and surface peaking wavelengths, and wavelengths greater than 14.5 μm peak above the troposphere. Note, the gaps in brightness temperature near 5 and 9 μm are the result of spectral gaps in the CrIS normal spectral resolution.

The impact of a cirrus cloud with cloud-top temperature of 210 K and COD532nm varying between 0.0 (no cloud) and 0.5 is shown in Fig. 3b. As expected, brightness temperature decreases as optical depth increases. With a COD532nm of 0.10 (shown in the dotted line), the brightness temperature differences (i.e., departure from the clear-sky value) peak greater than –2 K at the H2O screening wavelength of 10.4 μm. While the results in Fig. 1 are for bands between 13 and 14 μm, due to likely entire tropospheric channel rejection in the presence of cloud, it is not surprising a large number of the residual cirrus clouds shown in Fig. 1 exhibit COD532nm at or below this threshold with a similar temperature. Optically thicker clouds yield larger differences that are likely to be removed by screening with cold differences as large as 8 and 14 K for COD532nm of 0.30 and 1.00, respectively. The brightness temperature difference induced by a cirrus cloud of COD532nm = 0.30 and cloud-top temperature varying between 195 and 228 K is shown in Fig. 3c. Note that the combined effect of changes in cloud temperature and cloud properties due to the use of the temperature-dependent ice crystal effective diameter parameterization is evident. Here, colder clouds induce larger differences, most significantly around 12–13 μm albeit with less sensitivity than to COD532nm.

To ensure no implicit biases exist within the system, static cloud properties are provided to a modified 1DVar system for forward and Jacobian RTTOV calls. The analysis differences resulting from assimilating cirrus-contaminated radiances is presented as a function of different COD532nm and cloud-top temperatures in Fig. 4. As expected, as COD532nm increases, the differences correspondingly increase. Interestingly, even clouds with COD532nm as low as 0.05 (i.e., likely to pass screening) result in temperature and dewpoint differences near 2 and 5 K, respectively, with the temperature and dewpoint differences maximized much closer to the surface than the cloud layer. Finally, colder clouds also result in larger differences in both temperature and dewpoint as the thermal contrast with the lower atmospheric levels is maximized. Interestingly, dewpoint difference is as large as 23 K at a COD532nm = 0.30 and cloud-top temperature of 195 K. Note, as a sanity check, the analyses created when cloud properties were provided to the 1DVar system were compared to the true profile and with differences found to be approximately zero, with small deviations due to truncation error of the assimilated radiances on the order of 10−2 K. This finding helps to verify that cloud properties are being included as expected.

The temperature Jacobians from the final minimization iteration for the cloudy-sky analyses for CrIS band 745 (8.0 μm) and 1035 (6.2 μm) are shown in Fig. 5. For both
bands, when cloud properties are provided, Jacobian peaks are evident within the cloud layer with peak magnitude a function of COD$_{532\text{nm}}$ consistent with the results for different cloud fractions presented by Okamoto (2013). Note, even for a COD$_{532\text{nm}}$ up to and above 1.00, the cloud is still transparent enough in the IR to allow signal from the lower atmosphere to impact measured radiances. Thus, while current cloudy-sky assimilation using overcast observations exhibits increased analysis accuracy in the upper troposphere, the Jacobians here suggest that assimilation of radiances from semitransparent clouds could increase accuracy below the cloud as well. Note, the impact of the highest clouds at 195 K cause the temperature Jacobian to go slightly negative within the cloud layer. This feature may be implicit within RTTOV as finite-difference testing showed molecular—primarily water vapor—optical depth increases as temperature increases within this layer. As such, temperature Jacobian calculations within this layer are negative even with the cloud present as cloud properties are held constant.

While the temperature Jacobians presented in Fig. 5 exhibit the response of the analysis when cloud is considered, it is important to understand system response when the cloud impacted radiances are assumed clear. As such, the temperature Jacobians from the final minimization iteration for the clear-sky assimilation runs are shown in Fig. 6. That is, these Jacobians represent the atmosphere retrieved using cloud-contaminated observations assuming the radiances are clear sky, whereas the Jacobians in Fig. 5 represent the true atmosphere. Thus, while there is no cloud in the background, cloud exists within the assimilated observation and causes error in the analysis. Figure 6 allows us to examine if and where the cloud impact is being aliased in the retrieved atmosphere. Since there are no clouds
in the background or analysis, there is no Jacobian peak at the
cloud layer. Instead, large peaks are present in the lower tropo-
sphere, likely due to large increases in moisture (and thus
opacity at these wavelengths; shown in Fig. 4) in the lower tropsphere. Thus, when cloudy radiances are assimilated, the er-
ror will be aliased through the atmospheric column, often lower
in the troposphere where greater moisture and temperature values (and likely background error) can make compensatory
changes. Thus, assimilation of cloud-contaminated radiances is
likely to more impact on the lower troposphere than within the
cloud layer. But, since all channels are assimilated here, even
bands with Jacobian peaks at higher altitudes than the cloud
may have sensitivity to the cloud. Thus, differences are expected
both above and below the cloud.

While Fig. 4 depicts the differences profile for individual
cases, it is desirable to examine a wider variety of cases. To do
so, we examine the differences using two metrics: maximum
tropospheric temperature and moisture analysis difference,
and temperature and moisture root-mean-squared deviation
(RMSD) throughout the troposphere. These metrics provide
an estimate of both the maximum differences possible, as well
as a total column-averaged difference. As such, a total of four
difference matrices (two for temperature and two for moisture)
are constructed as a function of COD$_{532}$nm and cloud-top
temperature. Note, we use cloud-top temperature as a
proxy for cloud altitude. Cloud-top temperature is preferred
due to the implementation of the Dolinar et al. (2022) parameterization and, thus, the varying of cloud optical properties
with cloud-top temperature. We then superimpose and inte-
grate according to the cloud contamination statistics to esti-
mate bulk differences.

Matrices of maximum tropospheric temperature and dewpoint
differences associated with assimilating cirrus-contaminated radi-
ances from cloud of differing COD$_{532}$nm and cloud-top tempera-
ture are shown in Fig. 7. Superimposed over the difference
estimates are the percent frequency of that cloud occurring in the
NAVDAS-AR assimilated observations presented in Fig. 1. As expected, maximum difference increases as COD$_{532\text{nm}}$ increases or cloud-top temperature decreases. Even for a hypothetical cloud with a cloud-top temperature of 210 K and COD$_{532\text{nm}}$ of 0.10 maximum temperature difference is near 2.5 K and dewpoint difference is over 11 K. These values are nearly 5 times the expected background uncertainty noted by NAVDAS-AR.

By integrating the frequencies of contamination with the maximum difference, a bulk difference estimate can be determined. The expected bulk maximum global difference in temperature is 0.11 and 0.35 K for dewpoint. While these differences are lower than the expected NAVDAS-AR uncertainty, they are also normalized by the 7.7% contamination rate. This contamination rate is not truly representative of what is possible regionally, however. Figure 8 shows the average frequency normalized by the total number of observations for observations containing cirrus exhibiting a COD$_{532\text{nm}}$ less than or equal to 0.10 for June 2006–March 2021 as determined using the CALIOP, version 4.2, 5-km cloud layer product (Young et al. 2018). In the tropics, cirrus of COD$_{532\text{nm}}$ less than or equal to 0.10 (i.e., clouds expected to pass through current screening measures) is regularly present at frequencies at or above 20%, with areas such as Southeast Asia exhibiting frequencies between 40% and 50%. While these frequencies include cirrus clouds that are unlikely to cause large differences if assimilated (e.g., those with COD$_{532\text{nm}}$ less than 0.01), these cirrus clouds only represent approximately 15% of cirrus with COD$_{532\text{nm}}$ less than or equal to 0.10. Additionally, since the integration of the difference matrices from Fig. 7 include these less impactful cirrus clouds, the bias estimates reported here inherently account for these optically thinner, yet less frequent, clouds. As such, while the differences are not particularly concerning globally, regionally the differences can be much larger than expected background uncertainties.

Fig. 5. Analysis temperature Jacobians as a function of pressure for (a),(c) a cloud with cloud-top temperature of 210 K and 532 nm cloud optical depth between 0 and 0.5 and (b),(d) a cloud with cloud-top temperature between 195 and 228 K and an optical depth of 0.30. These simulations are for CrIS bands (a),(b) 745 corresponding to 8.0 µm and (c),(d) 1035 corresponding to 6.2 µm.
Similar to Fig. 7, the tropospheric RMSD matrices are shown in Fig. 9. As with maximum difference, RMSD increases with increasing COD ($\text{COD}_{532\text{nm}}$) and decreasing cloud-top temperature. As expected, however, RMSDs are much lower than the maximum biases since they are integrated throughout the troposphere. Again, for the hypothetical cloud with COD$_{532\text{nm}}$ of 0.10 and cloud-top temperature of 210 K the tropospheric temperature and dewpoint RMSDs are near 2 and 5 K, respectively. Thus, despite being tropospheric integrated differences, these values are much larger than the expected background uncertainty. When integrated by the superimposed global contamination rates, the bulk RMSDs are 0.06 and 0.11 K for temperature and dewpoint, respectively. We reiterate that, while these values are relatively small on their own, when examined in regions with likely larger contamination rates, these values may be several times larger.

It is important to note the differences shown here are specific to both the tropical standard atmosphere and the background error covariance matrix used here. As such, differences particular to any given NWP center will vary from these values. Additionally, while the contamination rates from NAVDAS-AR are used for calculating bulk estimates, the estimated differences shown here likely represent a worst-case scenario for that system. Specifically, NAVDAS-AR assimilates observations from a range of observing platforms, whereas the simulations here only assimilate a single cirrus-contaminated observation from the CrIS sensor aboard the Suomi NPP satellite. Finally, the impact of including instrument noise based upon the observation error covariance matrix on the bulk statistics is investigated and results are within 0.2%–1.6% of those reported above.

4. Conclusions

While assimilation of observations from hyperspectral infrared sounders (HIS) have become vital for numerical
weather prediction, assimilation is still dependent on aerosol and cloud-free observations—a requirement often violated. For instance, 7.7% of the Naval Research Laboratory Variational Data Assimilation System—Accelerated Representer (NAVDAS-AR) assimilated HIS observations are contaminated by cirrus clouds. While this contamination rate is lower than that found for HIS aerosol contamination M21, 86% of residual cirrus exhibit cloud-top temperatures between 190 and 220 K. As such, these cloud exhibit much larger thermal contrast with the lower troposphere (and potential for analysis impact) than the tropospheric dust examined by M21.

The impact of assimilating cirrus-contaminated HIS radiances on temperature and moisture analyses for NWP is estimated using the Radiative Transfer for TOVS (RTTOV) model and a one-dimensional variational (1DVar) assimilation system. It is shown that a cirrus cloud with cloud optical depth at 532 nm (COD532nm) of 0.1 and a cloud-top temperature of 210 K will impart less than 3 K impact on the HIS observed brightness temperatures but exhibits temperature and dewpoint analysis differences of 2.5 and 11 K, respectively. Bulk difference estimates of 0.11 K for temperature and 0.35 K for dewpoint are found by integrating with respect to
the NAVDAS-AR cirrus contamination statistics. Finally, tropospheric root-mean-squared deviations (RMSDs) of near 2 K for temperature and 5 K for dewpoint are found for a cloud with COD_{32} of 0.1 and cloud-top temperature of 210 K, whereas bulk RMSDs of 0.06 for temperature and 0.11 K for dewpoint are shown.

This study suggests that thin cirrus clouds with visible optical depth less than 0.1 (visible) are unlikely to be detected by using HIS observations alone, introducing a nontrivial impact to HIS data assimilation. Even including traditional imager cloud masks (e.g., Eresmaa 2014) may be insufficient to screen these thin cirrus clouds given the evidence of cirrus contamination in other imager products (e.g., Toth et al. 2013; Marquis et al. 2017). This triggers the need for including a cirrus cloud sensitive instrument, either an active-based lidar or a passive-based thin cirrus cloud sensitive sensor, such as the 1.38 μm spectral channel in MODIS and VIIRS, to be companioned with future HIS missions for thin cirrus cloud detection, and especially for regions where both cirrus clouds and tropical cyclones are dominant, such as tropical regions of the Atlantic and Pacific Oceans. Cirrus detection during daytime at 1.38 μm has shown promise. McHardy et al. (2021) reports reliable cirrus detection at optical depths as low as 0.05 using the GOES-16 Advanced Baseline Imager 1.38 μm band. Since the VIIRS instrument (on board the same satellites as the CrIS sensor) contains a similar cirrus detection band near 1.38 μm, the VIIRS-CrIS fusion product (Baum et al. 2019) could likely be quickly adapted for such efforts though examination of the cirrus detection/contamination in that product is required. While methods similar to that presented by McHardy et al. (2021) would not remove all contaminating cirrus, it could be used to substantially decrease overall impact, even if only limited to daytime observations. Finally, given the sensitivity of HIS observations to layers beneath optically thin cirrus clouds, the methods of providing the forward and Jacobian radiative transfer model systems with static cloud optical properties ideally from state-of-the-art imager retrievals could be used to allow assimilation in cloudy scenes.

Acknowledgments. This work is supported under the Office of Naval Research Grants N0001420WX00481 and N0001421WX01241, and the NASA Grant 80HQTR18T0085. Authors J. Marquis and J. Zhang were also supported by the Office of Naval Research Grant N00014-16-1-2040 and the Naval Research Laboratory Grant BAA N00173-19-S-BA01. Author A. Garnier is supported by the Office of Naval Research Grant N00173-20-1-G002.

Data availability statement. All CALIOP data used in this study are openly available from NASA Earthdata’s Atmospheric Science Data Center. The ice optical model used is available from P. Yang (pyang@geos.tamu.edu) at Texas A&M University with bulk optical properties calculated as described above. The radiative transfer and assimilation code are available via EUMETSAT NWP SAF with edits to the original source code detailed above. Due to its proprietary nature, NAVDAS-AR data used to determine contamination statistics are typically not available. For further information about conditions for access, please contact J. Campbell (james.campbell@nrlmry.navy.mil).

REFERENCES


Baum, B. A., P. Yang, A. J. Heymsfield, C. G. Schmitt, Y. Xie, A. Bansemer, Y. Hu, and Z. Zhang, 2011: Improvements in


Toth, T. D., and Coauthors, 2013: Investigating enhanced Aqua MODIS aerosol optical depth retrievals over mid-to-high latitude Southern Oceans through intercomparison with


