Boundary layer water vapour statistics from high-spatial-resolution spaceborne imaging spectroscopy

Mark Richardson[1,2], David R. Thompson[1], Marcin J. Kurowski[1], Matthew D. Lebsock[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
[2]Department of Atmospheric Science, Colorado State University, Fort Collins, CO 90095, USA

Correspondence to: Mark Richardson (markr@jpl.nasa.gov)

Abstract. Daytime clear-sky total column water vapour (TCWV) is commonly retrieved from visible and shortwave infrared reflectance (VSWIR) measurements, and modern missions such as the upcoming Earth Surface Mineral Dust Source Investigation (EMIT) offer unprecedented horizontal resolution of order 30—80 m. We provide evidence that for convective planetary boundary layers (PBLs), spatial variability in TCWV corresponds to variability in PBL water vapour. Using an observing system simulation experiment (OSSE) applied to Large Eddy Simulation (LES) output, we show that EMIT can retrieve horizontal variability in PBL water vapour, provided that the domain surface is uniformly composed of either vegetated surfaces or mineral surfaces. Random retrieval errors are easily quantified and removed, but biases from -7 % to +34 % remain in retrieved spatial standard deviation and are primarily related to the retrieval's assumed atmospheric profiles. Future retrieval development could greatly mitigate these errors. Finally, we account for changing solar zenith angle (SZA) from 15—60° and show that the non-vertical solar path destroys the correspondence between footprint retrieved TCWV and the true TCWV directly above that footprint. Even at the 250 m horizontal resolution regularly obtained by current sensors, the derived maps correspond poorly to true TCWV at the pixel-scale, with $r^2<0.6$ at SZA=30°. However, the derived histograms of TCWV in an area are closely related to the true histograms of TCWV at the nominal footprint resolution. Upcoming VSWIR instruments, primarily targeting surface properties, can therefore offer new information on PBL water vapour spatial statistics to the atmospheric community.

1 Introduction

Thermodynamic information about the planetary boundary layer (PBL), including information about water vapour (q_v), is a targeted observable recommended by NASA's Decadal Survey (National Academies of Sciences, Engineering, 2018). PBL q_v estimates would go beyond the current TCWV and free-tropopause products to provide new information about the vertical moisture structure for weather and climate applications. The Decadal Survey explicitly recognised the PBL's importance since it "literally couples the surface of the Earth to the atmosphere above", and among other important factors, gradients of moisture between the surface and PBL, and between the PBL and free troposphere are strong controls on vertical atmospheric heat and

moisture transport. The formation of boundary layer clouds was also highlighted due to their importance for Earth's energy balance. A critical measurement gap in the current observations of PBL thermodynamics is the inability to quantify mesoscale variations in PBL q_v. Mesoscale aggregation in PBL water vapour appear to play an important role in determining the timing of deep convective events (Stirling and Petch, 2004; Wulfmeyer et al., 2006). Furthermore, in-situ observations suggest that

5   the majority of the variation in the TCWV prior to convective initiation can be explained by variability within the PBL (Couvreux et al., 2009). The mesoscale spatial variability of q_v is not resolved by current global weather or climate models but instead it must be parameterised. Modern approaches to parameterize PBL variability include Eddy-Diffusivity/Mass-Flux approaches (Suselj et al., 2019) and higher order closure approaches that include prognostic equations for higher order moments such as the variance (Golaz et al., 2002; Larson et al., 2002). However, we lack observations at a global scale to

10   evaluate the small-scale variability produced by these models. This paper will address the feasibility of addressing this measurement gap using upcoming observations from very high spatial resolution VSWIR observations from space.

This study is primarily motivated by the ongoing development of spaceborne hyperspectral VSWIR measurement capacity at fine horizontal resolution. We focus on the EMIT mission, planned to launch to the International Space Station (ISS) in 2022 with average footprint size (Δx) of 60 m (Green and Thompson, 2020). However, similar or improved capacity is anticipated

15   in response to NASA's Surface Biology and Geology (SBG) designated observable, with the Hyperspectral Infrared Imager (HysPIRI, Lee et al. (2015)) concept offering Δx of 30—60 m; and for ESA's Copernicus Hyperspectral Imaging Mission for the Environment (CHIME), also known as Sentinel 10, for which the prime contractor was selected in July 2020 and whose Mission Requirements Document refers repeatedly to Δx <30 m (Rast et al., 2019).

Of present missions, this analysis may be applicable to the Italian PRecursore IperSpettrale della Missione Applicativa

20   (PRISMA, Candela et al. (2016)), which provides similar spectral range and sampling to EMIT at Δx=30 m. Some of the conclusions will also apply to other recent instruments, such as Sentinel-2's Multi-Spectral Imager (MSI, Drusch et al. (2012)) which offers Δx =20 m, albeit with far fewer channels, or the DLR Earth Sensing Imaging Spectrometer (DESIS, Krutz et al. (2019)) which provides hyperspectral measurements over a smaller wavelength range. These modern and upcoming instruments offer Δx that are substantially smaller than past VSWIR instruments that retrieve TCWV, such as ESA's Medium

25   Resolution Imaging Spectrometer (MERIS) on Envisat, whose smallest provided Δx is approximately 0.25 km $\times$ 0.30 km, which allowed the identification of horizontal convective rolls during a high pressure event over Germany (Carbajal Henken et al., 2015), but cannot resolve the smaller scales of variability. Recently, Thompson et al. (2021) used VSWIR measurements from the Airborne Visible Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) to capture information about PBL q_v variability at spatial scales <1 km, which cannot be determined with footprint sizes similar to MERIS.

30   EMIT-like instruments could allow retrieval of bulk PBL q_v, which we henceforth refer to as the partial column water vapour in the PBL (PCWV$_{PBL}$) via two demonstrated approaches. The first approach uses VSWIR measurements alone, and the second combines separate above-PBL water vapour (PCWV$_{upper}$) and TCWV to obtain PCWV$_{PBL}$ = TCWV – PCWV$_{upper}$. A third

approach, that has not been demonstrated operationally to our knowledge, is to perform joint retrievals using both VSWIR and vertically resolved sounding measurements.

The direct VSWIR-only method can be seen in Trent et al. (2018), who estimated $PCWV_{PBL}$ from the Greenhouse Gases Observing Satellite (GOSAT, Kuze et al. (2009)), while the second is explored in Millán et al. (2016), who paired TCWV from passive microwave measurements with $PCWV_{upper}$ above horizontally-uniform clouds from Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared retrievals. The resultant $PCWV_{PBL}$ values showed good agreement with radiosondes and ERA-Interim reanalysis and a promising candidate approach is to use VSWIR TCWV in place of the microwave measurements.

The physical principle of VSWIR TCWV retrievals is differential optical absorption spectroscopy (DOAS). More TCWV leads to increasing depth of H_2O absorption features relative to other wavelengths. This applies to $TCWV_{VSWIR}$ from missions including MERIS (Bennartz and Fischer, 2001; Guanter et al., 2008), MODIS (Diedrich et al., 2015; Gao and Kaufman, 2003), TROPOMI (Borger et al., 2020; Schneider et al., 2020), SCIAMACHY (Noël et al., 2004), GOME (Noël et al., 1999), GOME-2 (Grossi et al., 2015) and OCO-2 (Nelson et al., 2016).

These instruments vary in spectral range and sampling, but all must contend with the measured spectra responding to properties other than TCWV. The retrievals only operate for daytime cloud-free scenes, and commonly only over land since water surfaces are dark such that insufficient light reaches the sensor to allow for a TCWV retrieval, with exceptions for sun glint as exploited in the aforementioned AVIRIS-NG study (Thompson et al., 2021). Thompson et al. selected these AVIRIS-NG flights because DOAS techniques respond to the total light path absorption including the slanted sunlight path from the top of atmosphere (TOA) to the surface. This horizontally smears the effective footprint size, with larger smearing for larger solar zenith angle (SZA). As footprints become smaller, the proportional effect of this smearing may become more important, and so here we apply solar ray-tracing to determine whether observations with a nominal Δx of 20—50 m obtain useful information about the spatial statistics of $PCWV_{PBL}$ at that spatial resolution. We use two performance metrics: (i) the correlation between retrieved TCWV and true TCWV, which was used as input for our forward simulations, and (ii) the spatial standard deviation, σ_x of retrieved TCWV within a snapshot relative to the LES output $PCWV_{PBL}$ σ_x, which we refer to as the true σ_x.

We employ a new type of Observing System Simulation Experiment framework and performs simulated VSWIR retrievals of TCWV from high-spatial-resolution large eddy simulation (LES) output to determine whether horizontal spatial variability in PBL q_v can be obtained from retrieved TCWV, and conclusions are limited to daytime non-cloudy conditions. The purpose of this is a detailed sensitivity study using retrieval code and tools already developed for EMIT. We consider $\Delta x \geq 40$ m since this is appropriate for EMIT and several LES cases in our archive that were run at that resolution.

Here we test the use of the iterative optimal estimation code Imaging Spectrometer Optical Fitting (Isofit) for a spaceborne application, specifically target TCWV and address the following questions:

1) In LES, how does horizontal variability in TCWV relate to $PCWV_{PBL}$?
2) What uncertainties are introduced into the retrieval by EMIT instrumental error, non-uniform AOD and different surface types, and can these errors be anticipated and quantified from observations alone?

3) What is the correlation coefficient between retrieved and true TCWV, and can the spatial standard deviation be estimated? How does this depend on LES of different convective PBL types?

4) How does the solar path across different SZAs affect these conclusions?

This scope excludes important factors such as topography, inter-channel correlated errors in the instrument, imperfect cloud masking and cloud 3-D radiative effects, and our paper is structured to address these questions in turn, with each analysis section containing its own methodology, results and discussion. Section 2 explores the raw LES output to address question 1, Section 2 describes the synthetic retrievals and analysis methodology to address questions 2—3, Section 4 adds solar path analysis to address question 4, and Section 5 discusses and concludes.

2 Large Eddy Simulations

2.1 Model setup, scenarios and snapshot selection

We use output from five LES simulations named RICO, ARM, ARM_lsconv, BOMEX and DRY, which are summarised with references in Table 1. They all represent convective boundary layers characterised by either low-altitude or no cloud cover. The 23 separate snapshots are identified by timestamp, e.g. ARM_18000s is five hours into the ARM simulation. Simulation Δx sets the implied measurement horizontal resolution and varies from 20–50 m.

The simulations are performed with two different models: EULAG (Prusa et al. (2008); ARM and ARM_lsconv) and JPL-UCONN LES (Matheou and Chung (2014); RICO, BOMEX, DRY). Each simulation applies periodic lateral boundary conditions and a horizontally-homogeneous initial state. For the RICO case, interactive sensible and latent heat surface fluxes over constant-temperature ocean are used, while the other cases are driven by prescribed (either constant for DRY and BOMEX, or time-dependent for ARM and ARM_lsconv) surface fluxes. All other setup details are explained in the Table 1 references, these references also show how the ARM, BOMEX and RICO LES simulations, which were based on detailed field campaigns, accurately reproduce the main features observed during those campaigns. Each three-dimensional LES snapshot is merged with one-dimensional MERRA-2 reanalysis profiles aloft to produce a full-depth atmospheric column. Reanalysis data are chosen for the dates and locations of the field campaigns the LESs refer to. The DRY and ARM_lsconv cases share the same upper-atmospheric profiles as ARM. In all cases except for DRY, Table 1 rows (vii)—(ix) show that the LES capture >85 % of total TCWV. For retrieval purposes we ignore the LES surface type and apply an assumed surface reflectance spectrum below the LES profiles.

2.2 Profiles and PBL height

Definitions of PBL height, z_{PBL} vary widely. We found similar results from four standard thermodynamic calculations (von Engeln and Teixeira, 2013), so henceforth we define z_{PBL} to be the altitude of max(dθ/dz) where θ is the all-sky mean potential temperature. Mean all-sky profiles of T and q, horizontal standard deviation in q (σ_q), and cloud fraction are shown in Figure

1. Changes in σ_q are the largest differences between timesteps, but are small (<10 %) relative to the mean, so measuring this variability will require precise observations. Also, σ_q is negligible in the layers in the free-troposphere that lie above the PBL but are resolved by the LES, implying that the LES domains capture q_v variability. We later support this claim using real-world airborne LIDAR retrievals.

2.3 Water vapour spatial variability statistics and the relationship between TCWV and PCWV$_{PBL}$

Figure 1 displays all-sky conditions but our retrievals only target clear-sky thereby missing a moister tail to the distribution (Supplementary Figure 1). Within-cloud retrievals would require alternative measurement approaches, such as differential absorption radar (Roy et al., 2018, 2020), and the restriction to clear-sky scenes is a limitation that also applies to current thermal infrared and lidar retrievals.

We assess TCWV-PCWV$_{PBL}$ spatial variability by calculating clear-sky PCWV up to capping altitudes from 0.5—5 km and then correlating these with TCWV. Figure 2 confirms that >90 % of horizontal variance in LES TCWV at these scales is explained by PCWV$_{PBL}$. It is reasonable to ask whether this finding that the PBL variance dominates the TCWV variance is representative of the real atmosphere. Indeed the LES results are supported by the same statistics calculated from High Altitude Lidar Observatory (HALO) flights over the Pacific Ocean in April 2019 (Bedka et al., 2021), as presented in Thompson et al. (2021) and shown in Figure 2(f). In these calculations TCWV is only calculated up to 8 km due to flight altitude, but this real-world data includes free-tropospheric moisture variability and furthermore will have lower r values due to the presence of random retrieval error. The horizontal resolution is ~3 km versus the 20—50 m of LES, and the HALO sampling is sparse and often separated by hundreds of km due to clouds. Nevertheless, the HALO flights show that horizontal TCWV variability can be well captured within 3 km altitude in real scenes, and provide evidence that the LES domains capture horizontal variability in q_v.

The TCWV-PCWV$_{PBL}$ fit coefficients for ARM, ARM_lsconv, BOMEX and RICO range from 0.99—1.04 mm mm^{-1}, i.e. a 1 mm change in PCWV$_{PBL}$ means a 0.99—1.04 mm change in TCWV. This confirms that almost all horizontal q_v variability occurs within the mean PBL height. For the DRY case, coefficients range from 1.06—1.12 mm mm^{-1}. These coefficients mean that PCWV$_{upper}$ spatially correlates with PCWV$_{PBL}$, which could be explained by moister plumes rising and having higher local z_{PBL} than the domain-mean value used in the calculation. In summary, we have answered question 1 from Section 1 and can expect spatial variability in retrieved TCWV for these cases to represent real variability in PCWV$_{PBL}$, and so use TCWV and PCWV$_{PBL}$ interchangeably from now on.

3 Simulated EMIT retrievals of TCWV in LES

This experiment requires a large number of inversions over a wide spatial field. Simulating synthetic spectra and performing a retrieval for every grid point proved to be prohibitively computationally expensive. Consequently, we develop an emulator to statistically reproduce the result of the full inversion but with dramatically better efficiency. Retrievals will include a range

of surfaces in a subset of the snapshots (to identify sensitivity to surface type) and then a fixed surface type across all snapshots (to identify sensitivity to atmospheric conditions). Sensitivity tests will be performed on individual subsets of snapshots as required, and a correction method for identifying the random component of retrieval error will be introduced. Section 3.1 describes the relevant methods, Section 3.2 the results and 3.3 discusses limitations.

5   3.1 Retrieval methodology

3.1.1 MODTRAN6.0 forward model, EMIT instrument characteristics and Isofit retrievals

We use the same retrieval code as in Thompson et al. (2021), Imaging Spectrometer Optimal Fitting (Isofit) for our synthetic retrievals (https://github.com/isofit/isofit). This iterative optimal estimation code simultaneously retrieves surface reflectance, aerosol optical depth (AOD) and TCWV, differing from older techniques that retrieve properties sequentially (e.g. Guanter et

10 al. (2008) for MERIS). Isofit is described and shown to have a closed error budget in Thompson et al. (2018), and has been applied to observations from several airborne campaigns (Thompson et al., 2019, 2020, 2021). Conceptually it targets surface reflectance ρ_s and the estimation of TCWV is seen as part of an atmospheric correction.

Forward simulations use MODTRAN6.0 (Berk et al., 2014, 2015), which provides a plane-parallel solution to the radiative transfer equation. Atmospheric reflectance and transmittance vectors ρ_a, t and spherical sky albedo s are calculated at

15 wavenumber separation Δk=0.1 cm^{-1} ($\Delta\lambda\approx$0.002 nm) before being convolved with the EMIT spectral response function), and the instrument is assumed to be nadir viewing from 100 km altitude. With no substantial atmosphere above 100 km, this gives the same results as the ISS altitude near 400 km, where EMIT will be hosted. A correlated-k method and the HITRAN database (Rothman et al., 2009) are used for gaseous absorption while scattering is handled by DISORT (Laszlo et al., 2016; Stamnes et al., 1988). The EMIT instrument properties are derived from the current mission instrument model, which accounts for all

20 signal-independent noise terms like electronic noise, and photon shot noise calculated using predicted efficiencies of the instrument mirrors, lens, grating, and focal plane array. Spectral range is 380—2500 nm with $\Delta\lambda_{channel}$=10 nm and full-width at half maximum averaging $\Delta\lambda_{FWHM}\approx$11 nm.

For forward simulations, merged LES-reanalysis T and q profiles are interpolated onto a profile with 8 points from 0—6 km then vertical resolution slowly degrades over 6—100 km. Interpolated TCWV differs from the LES-reanalysis but we assume

25 that conclusions regarding derived sensitivities and errors will not be strongly affected.

The forward radiance vector I is calculated using a standard Lambertian approximation (e.g. as in Vermote et al. (1997)):

$$I = \frac{I_0\mu_0}{\pi}\left[\rho_a + \frac{t\circ\rho_s}{1-s\circ\rho_s}\right] \tag{1}$$

Where I_0 is the downward top of atmosphere (TOA) solar radiance, μ_0 the cosine of the solar zenith angle, ρ_s the surface reflectance and $\circ$ represents channel-by-channel multiplication. The ρ_s elements represent the hemispheric-directional

30 distribution function (Schaepman-Strub et al., 2006). The atmospheric coefficient vectors t, ρ_a and s represent the transmittance of the solar reflected optical path, the path reflectance, and the spherical sky albedo, respectively. These coefficients are

obtained from simulations over a black surface. Using Eq. (1) in forward simulations results in negligible differences to retrieved TCWV compared with inserting the surface directly into MODTRAN forward simulations (Supplementary Figure 2). Use of Eq. (1) means that just one MODTRAN simulation is needed per column, rather than one for each combination of column and surface type. The pseudo-observation, I_{obs}, is I with random uncorrelated noise added, generated using the EMIT noise model.

The I_{obs} are input as observations to Isofit, while its state vector x contains surface reflectance in each channel, TCWV and aerosol optical depth at λ=550 nm (AOD), i.e. $x = [\ \rho_s\ \ AOD\ \ TCWV\ ]$. We mask the most strongly-absorbing channels due to lack of any surface information, so the retrieval uses 176 EMIT channels and therefore x has 176+2=178 elements.

The ρ_s elements are constrained via a covariance matrix whose mean is derived from a library of real surfaces, thereby capturing realistic spectral shapes. We retrieve absolute ρ_s, rather than the normalised value discussed in Thompson et al. (2018), and the prior is loosely constrained, however, ensuring that most information comes from the measurements.

Isofit uses Eq. (1) with a lookup-table (LUT) for its forward model, populating ρ_a, t, s for selected AOD and TCWV and then linearly interpolating in TCWV, AOD space to estimate I_{obs} given x. The LUT uses the default midlatitude summer profile and scales its $q(z)$ and aerosol extinction(z) to match desired AOD (from 0.05—0.30) and TCWV (from 5—53.5 mm). The Isofit default configuration uses the U.S. Standard Atmosphere 1976 (Sissenwine et al., 1976), but MODTRAN applies a relative humidity limit and the U.S. 1976 atmosphere is cool enough that MODTRAN automatically restricts its moisture content, such that the TCWV cannot reach the values seen in any LES case except for DRY. The midlatitude summer TCWV limit is just over 53.5 mm, so that defines our LUT maximum.

Our prior and first guess TCWV is 40 mm with a standard deviation of 7.5 mm although observationally a heuristic band ratio is commonly used to provide a first guess and a locally appropriate prior would be selected. However, this choice of prior doesn't change our derived spatial statistics (Supplementary Figure 3), although it results in a small mean shift of mean retrieved TCWV and reflectance (e.g. posterior TCWV shifts by 0.15 mm when prior is shifted by 32.5 mm).

3.1.2 Profile subsets, emulator development and sensitivity tests

All retrievals use radiances simulated at SZA=45°, using the profiles associated with an individual footprint and assuming a plane-parallel atmosphere. We define "clear sky" as where cloud water path<1×10^{-3} mm, approximately τ<0.3 in a typical subadiabatic cloud (e.g. Szczodrak et al. (2001)). Clear sky columns are ranked by TCWV and 101 columns equally spaced in terms of this ranking are taken (Supplementary Figure 4 justifies N=101).

All snapshots in a given LES case are combined and Isofit-retrieved TCWV$_{ret}$ is used to fit an emulator in combination with the forward-model TCWV via:

$$TCWV_{ret} = a_1 TCWV + a_2 + \epsilon \qquad , \qquad (2)$$

Where a_1 and a_2 are the trend and intercept parameters from an optimised-least-squares fit and ϵ is random zero-centred Gaussian noise with standard deviation from the emulator fit residuals. Tests with SZA from 14—60° show no significant

differences in a_1 with SZA, while the standard deviation of ϵ increases by up to 25 % at SZA=60° relative to SZA=45° (Supplementary Figure 5, Supplementary Table 1). Section 3.1.4 shows how we are able to identify and remove the effect of ϵ on derived statistics, so given that a_1 did not change with SZA in these tests we anticipate that our conclusions will largely apply to SZA up to and including 60°.

5 Forward simulation AOD varied from 0.1—0.2 and most footprints were assigned AOD=0.2. Supplementary Figures 6—7 show weak sensitivity of retrieved TCWV to AOD. The analysis is separated into two parts: Section 3.2.1 shows results for sensitivity of TCWV_{ret} to changes in surface spectrum within selected ARM snapshots and Section 3.2.2 shows changes in retrieved TCWV over a single surface type for all snapshots.

3.1.3 Development and fitting of retrieval emulator

10 For each emulator we use all snapshots within an LES run to fit Eq. 2 (separate snapshots fits do not affect conclusions, Supplementary Figure 8) and full-snapshot fields of TCWV_{ret} are then emulated using Eq. (2) with LES TCWV as input. The surface analysis uses the first 3 ARM snapshots and seven surface spectra from the Isofit surface model clusters, three of which are typical of vegetation, and the others of mineral surfaces. The database used to generate the surface model includes artificial surfaces, which are largely captured by the "mineral" spectra. An additional test was run with ARM_18000s profiles over

15 uniform Lambertian surfaces with ρ_s=0.1–0.5 in increments of 0.1. The atmospheric analysis uses the MODTRAN cropland and ocean ρ_s spectra for all 23 snapshots, although poor performance over dark surfaces means that the main emulator results are reported only for the land-surface retrievals.

Figure 3(a) shows typical spectra simulated over several surfaces, notably the MODTRAN ρ_s spectra have sharp changes that are not included in the Isofit surface model and therefore provide a challenging test of the retrieval code's ability to retrieve

20 TCWV outside of surface conditions for which it was developed.

With regards to the emulator parameters, non-unity a_1 represents biases in the local retrieval sensitivity $d\text{TCWV}_{\text{ret}}/d\text{TCWV}$. Possible causes will be discussed in Section 3.2.3, but this is the main concern for retrieval of local variability statistics because the retrieved standard deviation will be scaled by a_1, and this scaling will be undetectable in the absence of independent validation data. Changes in a_1 also change the derived spatial r^2, since $a_1>1$ increases retrieved σ_x variance and will increase

25 r^2. The parameter a_2 is related to a combination of the mean bias and the magnitude of a_1 within a snapshot, and may depend on factors such as surface type or biases in the LUT-assumed T and q profiles as seen for MERIS retrievals in Lindstrot et al. (2012). For our spatial statistics, a_2 has no effect since it is subtracted during calculation. The parameter ε represents non-systematic errors within a scene.

Importantly, σ_ε is not the typical error seen in validation or inter-comparison exercises (Diedrich et al., 2015; Nelson et al.,

30 2016; Pérez-Ramírez et al., 2014), since in these studies the varying biases between products in different conditions will add to the reported errors and make them larger than the σ_ε appropriate for our retrievals.

3.1.4 Estimating random error from retrieved fields

Random retrieval error ε with standard deviation σ_ε adds variance and therefore reduces r^2 while adding a high-bias term to estimated σ_x. Knowing σ_ε would allow removal of its bias contribution to σ_x, and clearly interpretation of spatial variability at a footprint level requires that σ_ε is small relative to σ_x. TCWV variability between columns separated by 50 m in the horizontal is far smaller than at larger separations. We will exploit this to estimate the spatially constant σ_ε using an approach based around the second order structure function S_2. Here we describe the recipe and mathematical justification, see Supplementary Figures 9—10 for a step-by-step illustration. For a TCWV field:

$$S_2(\Delta r) = E\left[\left(TCWV(x + \Delta r) - TCWV(x)\right)^2\right] \tag{3}$$

This is the variance between pairwise footprints separated by the distance Δr, and retrieved S_2 includes contributions from the spatial variance characteristic at that separation, $\sigma_x^2(\Delta r)$, and the observational uncertainty σ_ϵ^2. The subtraction removes the field mean $\overline{TCWV}$, and each of the terms TCWV(x) and TCWV($x+\Delta r$) will contribute $\sigma_x^2(\Delta r) + \sigma_\epsilon^2$ to the variance. We treat these as independent so their variances add to give the retrieved $S_{2,\text{ret}}$:

$$S_{2,ret}(\Delta r) = 2\sigma_x^2(\Delta r) + 2\sigma_\epsilon^2 \tag{4}$$

For ARM_18000s , $\sigma_x(\Delta r$=50 m) is 0.03 mm, compared with the full-snapshot σ_x of 0.29 mm. We exploit the smallness of σ_x at small Δr by smoothing the field in one direction with no overlap between smoothed footprints, and then calculating the structure function at Δr = 1 footprint (20—50 m, depending on the LES) perpendicular to the smoothing direction. For n-footprint smoothing, the independent component of variance shrinks by $1/n$, which we attribute to σ_ϵ^2. The steps are:

(i) Select a direction and evaluate $S_2(\Delta r)$ in that direction for Δr = 1 footprint separation

(ii) Smooth the field in the direction perpendicular to Δr by averaging over n_{foot}=2 footprints

(iii) Recalculate $S_2(\Delta r, n_{foot}$=2), treat the calculated value (i) as $S_2(\Delta r, n_{foot}$=1), regress $S_2(\Delta r, n_{foot})$ against $1/n_{foot}$, and take the best-fit trend to be equivalent to $2\sigma_\epsilon^2$.

By smoothing in one direction and then calculating orthogonally, the separation distance Δr does not grow with smoothing and so we maintain the advantages of the small $\sigma_x(\Delta r$=20—50 m). For estimating TCWV σ_ε with EMIT-like Δx, this method outperforms a standard spatial smoothing filter approach (Supplementary Figure 11).

3.1.5 Calculating spatial statistics and relationship to spatial smoothing

We calculate the spatial standard deviation σ_x of clear-sky TCWV and TCWV$_{\text{ret}}$ for each snapshot. The random error σ_ε is then estimated following Section 3.2.3 and subtracted in quadrature,

$$\sigma_{x,ret,corr} = \sqrt{\sigma_{x,ret}^2 - \sigma_{\epsilon,ret}^2} \tag{5}$$

Where the subscript "ret" means retrieved and "corr" means corrected.

The other target statistic is r^2 between TCWV and TCWV$_{\text{ret}}$, we calculate this directly and also estimate it via:

$$r_{est}^2 = \frac{\sigma_{x,ret}^2 - \sigma_{\epsilon,ret}^2}{\sigma_{x,ret}^2} \tag{6}$$

Where emulator trend $a_1=1$, estimated error from Section 3.2.3 is accurate, and there are no spatially-varying errors, Eq. (6) should reproduce retrieval r^2. However, $a_1 \neq 1$ means each $\sigma_{x,ret}^2$ term will be multiplied by a_1^2, resulting in an erroneous r^2 estimate. User requirements for r^2 will depend on application, we arbitrarily select $r^2=0.9$ as a target and compare Eq. (6) estimates with the true field values. True r^2 is unknowable without perfect knowledge of the TCWV field, but operational estimation using Eq. (6) would allow users to determine whether their requirements are likely to be met.

If r^2 is too low for the desired application then averaging over footprints may address this, although it results in loss of fine spatial information, it may be necessary to suppress errors or may be enforced by effective horizontal smearing where SZA>0°.

We show the results of sequentially smoothing the TCWV and TCWV$_{ret}$ field on both σ_x and r^2, and smooth in both horizontal directions, for example going from 50 m × 50 m to 100 m × 100 m. Smoothed footprints do not overlap, so are independent, and the smoothing is done on TCWV$_{ret}$ rather than on the radiance field. This avoids the requirement for additional forward model runs, and furthermore allows predictions of how r^2 changes with resolution by applying Eq. (6) with a minor modification:

$$r_{est}^2 = \frac{\sigma_{x,ret}^2 - \frac{\sigma_{\epsilon,ret}^2}{n}}{\sigma_{x,ret}^2} \tag{7}$$

Where n is the number of footprints over which TCWV$_{ret}$ has been smoothed, e.g. for the 50 m × 50 m to 100 m × 100 m transition $n=4$. In this case, $\sigma_{\epsilon,ret}$ must be calculated at the native resolution, and therefore exploits the smaller TCWV variance at $\Delta r \sim 50$ m, rather than the higher variance in a smoothed field with larger Δr.

3.2 Simulated retrieval results

3.2.1 TCWV retrievals over different surfaces

We first remind readers that "retrieval error" here only includes errors present in these synthetic retrievals, and excludes several real-world sources, such as how the true atmosphere is not plane-parallel as assume in our radiative transfer. Retrieved surface ρ_s spectra and TCWV$_{ret}$ versus forward model TCWV are shown in Figure 4. Surface ρ_s are retrieved well, with mean bias magnitude equivalent to 0.2—1.6 % of true ρ_s (e.g. for Lambertian $\rho_s=0.1$, the mean bias is 0.00021) and standard deviation across all channels is 2—4 % of true ρ_s. The largest contribution to errors is from spikes near $\lambda \sim 2.06$ μm. Inspection found that the MODTRAN CO_2 concentration changes between default profiles versus prescribed T and q profiles. In future an up-to-date CO_2 mixing ratio will be assigned, but the higher LUT value (361 ppmv) versus the forward model value (323 ppmv) results in the retrieval overly brightening the surface in the strong CO_2 band near $\lambda \sim 2.06$ μm.

Comparing Figure 4(d—f), TCWV$_{ret}$ over mineral surfaces is a mean 0.44 mm higher than over vegetation. From panel (f), some of this difference is likely related to mean surface brightness: darker surfaces give higher TCWV$_{ret}$. The other differences

in TCWV$_{ret}$ between surfaces must be due to spectral shape, but it appears that surface-induced errors are small when considering only mixed vegetation or mixed mineral surfaces. Regardless of the surface, a bias of order ~1 mm remains, which is similar to the largest difference introduced by surface type and may be related to other retrieval errors such as inappropriate atmospheric profile shapes assumed in the LUT. However, the derived spatial statistics we are interested in here are not affected by any mean bias.

Figure 5 shows example scenes with different surface types. The true TCWV standard deviation σ_x is 0.28 mm (panel a), while over the uniform surfaces the retrieval gives 0.33 mm (panels b,c), with the larger value due mainly to the σ_ε contribution. Over the striped surfaces it is 0.40 mm (panel d) due to the additional variance from combining surface types. However, if the top or bottom half of panel d is selected, then both return σ_x of 0.33 mm, i.e. the same as over a fixed vegetation or mineral surface. Statistics should not be taken over scenes with both vegetation and mineral surfaces but the Isofit surface classification, which is output by the retrieval, should be used to identify areas of sufficiently similar surface type for calculation of TCWV spatial statistics. The rest of the analysis assumes the MODTRAN cropland default surface.

3.2.2 TCWV retrievals over vegetation surfaces in all LES snapshots

Figure 6 shows TCWV retrievals over the MODTRAN cropland and ocean surfaces. The poor performance over ocean absent sun glint justifies our land-only investigation. Over land the mean bias ranges from -3.0 % (DRY) to +1.8 % (BOMEX) while the within-scene σ_ε is from 0.52 % (ARM_lsconv) to 0.67 % (BOMEX). As discussed in Section 3.1.3, VSWIR TCWV validation studies typically report error metrics larger than our σ_ε, but their values include inter-product differences in bias, which are potentially far larger. Bias is indeed sensitive to the assumed meteorological profiles, since by re-running the ARM_18000s retrievals using a LUT developed with the MODTRAN default "tropical" atmospheric profile shifts the mean bias from 0.33±0.04 mm to 0.14±0.04 mm (mean±2σ).

For the purpose of spatial variability in TCWV at scales of tens of km, the distinction between large-area and small-area retrieval errors is important. Generally speaking, the error in an individual column TCWV retrieval is of order 2—3 % since that includes the bias term, but for estimates of sub-10-km spatial variability, the within-LES 0.5—0.7 % is the error of interest.

3.2.3 Emulator parameters

Emulator parameters with ±2σ confidence intervals are listed in Table 2 and significant ($p<0.05$) non-unity trends can be seen most clearly for BOMEX (green) and RICO (purple) in Figure 6(a); the retrieved properties are more variable than reality, with trends of 1.34 and 1.22 mm mm^{-1} respectively. Meanwhile, the ARM and ARM_lsconv trends are both <1 mm mm^{-1}. Therefore σ_x calculated for BOMEX will be 34 % too high, and for ARM 6 % too low.

We argue that the most-likely causes of emulated trend bias is related to the vertical T and q profile. Firstly, dI/dq is nonlinear and varies with atmospheric conditions due to line broadening and interaction with aerosol layers. The a_1 fit parameter may

therefore be sensitive to differences between true profiles and those assumed in the retrieval LUT. Secondly, the LUT uniformly scales $q(z)$ profiles whereas the horizontal variability in q tends to peak at specific altitudes (Figure 1).

Two tests provide some evidence for this. Firstly, when using different standard atmospheres to generate lookup tables for the DRY case, retrieval gradient changes by 5 % and secondly, when re-running all BOMEX retrievals with forward radiances generated using the same q profile shape that has been scaled to match the original range of TCWV, the retrieval gradient changes by 9 % (Supplementary Figure 12). These results suggest that retrievals could be improved by using more-accurate meteorological profiles in the LUT development, and by using a more-appropriate scaling for q as a function of z in the LUT.

3.2.4 Snapshot statistics and estimation of random error

Figure 7(a) shows how σ_x of $TCWV_{ret}$ is overestimated in every snapshot (circles). Figure 7(b) shows that the estimated retrieval error σ_ε agrees excellently with the truth, and after removing σ_ε, the triangles in Figure 7(a) show the consistent overestimate is removed. Random error, such as that introduced by some instrumental uncertainties, can be precisely identified and removed from the spatial variance calculations. For estimating σ_x, the largest error source we consider is due to emulator slope. Other potential sources would be due to surface variation, which can be mitigated by selecting regions of similar surface classification as suggested in Section 4.1, and due to spatially-varying errors, such as inter-pixel calibration biases or those induced by unmodelled temperature gradients across the sensors. Separate approaches are required to account for these issues. Next, we consider the r^2 coefficient between TCWV and $TCWV_{ret}$, with an illustration in Figure 8, where the RICO_14400s $TCWV_{ret}$ fields are shown at the native resolution, and after smoothing down to 80 m. The random retrieval error is visible as speckling Figure 8(a) and clearly reduces following smoothing. The 2d histograms Figure 8(c,d) demonstrate the increase in r^2 from 0.82 to 0.95 following a coarsening of the Δx resolution from 40 m to 80 m.

Figure 9 summarises the true and estimated statistical values as horizontal resolution is sequentially degraded. Comparison of Figure 9(a,b) reveals that there is only a small decrease in σ_x as resolution coarsens up to hundreds of metres, and the biases between estimated and true values follow emulator a_1 trends as expected, with ARM and ARM_lsconv too low and DRY, RICO and BOMEX too high.

Regarding r^2 in Figure 9(c,d), Eq. (7) reliably predicts true r^2, so a user could determine the spatial resolution required to achieve a desired r^2. In all snapshots $r^2>0.90$ at 150 m resolution, and in 21/23 cases this is achieved at 100 m. Therefore, with the errors accounted for here, the EMIT instrument could capture 90 % of spatial variability in $PCWV_{PBL}$ at 100 m resolution in the PBL conditions examined here, a factor of 7.5 improvement on the MERIS full-resolution retrievals. However, this conclusion does not account for the spatial smearing caused by SZA.

3.3 Discussion of retrieval results and limitations

This section has addressed questions 2) and 3) from Section 1, and shown that random errors introduced by EMIT's instrumental error can be accurately identified and removed. Provided that an observed domain consists of mixed vegetation

or mixed mineral surfaces, then our derived error in σ_x using EMIT is from -7 % to +34 %. Isofit returns surface type, meaning that such domains can be identified from retrievals.

Computational limitations forced adoption of an emulator approach, which provides a useful framework to assess error sources. Firstly, this framework shows that the errors of interest for retrieval of spatial statistics of $PCWV_{PBL}$ are the gradient a_1,

5 equivalent to $dTCWV_{ret}/dTCWV$, and random error σ_ε. We show that σ_ε can be estimated and removed, and that the main error is that in a_1, most likely driven by the retrieval's atmospheric profile assumptions, which can be addressed in future development. Our method to derive σ_ε also allows users to predict spatial correlation, in particular we found that an $r^2>0.9$ requirement requires smoothing to 100—150 m resolution. This is a factor of 3—8 improvement in sampling relative to MERIS full resolution.

10 Limitations include the use of the same radiative transfer code for forward and inverse simulations, so spectroscopic errors were ignored, as were errors in cloud- and shadow masking, those caused by topography, or errors that correlate between footprints.

Spectroscopy errors can be estimated (Thompson et al., 2020) and should shrink in future with developments, with ongoing research in water vapour absorption spectroscopy (Elsey et al., 2020; Lechevallier et al., 2018; Menang et al., 2021) and a

15 history of targeted development of spectroscopy to improve retrievals, such as for OCO-2 (Drouin et al., 2016; O'Dell et al., 2018; Payne et al., 2020). The surface remote-sensing community has tools for addressing topography (Kobayashi and Sanga-Ngoie, 2008; Teillet et al., 1982), and there are also approaches to dealing with nearby clouds to minimise the effect of imperfect cloud edge identification, shadowing and 3D cloud radiative effects (Massie et al., 2021). Nevertheless, these are all topics that are worth evaluating for Isofit-like TCWV retrievals.

20 We note that our σ_ε is smaller than the errors reported in product intercomparison studies, but those studies implicitly capture variance due to differing mean biases, i.e. the a_2 term in our emulator, which is larger than the other terms. An evaluation of our retrieved σ_x or would require independent validated sources such as passive microwave or differential absorption lidar data with $\Delta x \leq 50$ m that are collocated with VSWIR $TCWV_{ret}$. Reported comparisons are typically of bias and root-mean-squared error (RMSE) of satellite VSWIR retrievals relative to surface-based or other satellite products, and are calculated from

25 datasets across a range of times and sometimes places. Furthermore, the comparison data generally has larger Δx and may not be perfectly collocated in time and space, introducing additional variance that contributes to reported RMSE. Therefore typical published analyses include within their RMSE uncertainties due to differences in space and time of measurements, and any differences between the a_2 terms between the VSWIR and validation dataset retrievals. Therefore, these reported errors cannot be compared with our values which are calculated within individual LES runs. We can, however, report that our errors are

30 similar to Thompson et al. (2021)'s airborne Isofit retrievals statistics against nearby AERONET surface stations, which reported RMSE of 2.8 mm. Flight C data from Figure 9 of that paper show a spatial standard deviation of 0.19 mm when smoothed to $\Delta x=48$ m, which is within the LES simulated range, and σ_ε is estimated at 0.18 mm, although that is not comparable to our values since it uses AVIRIS-NG rather than EMIT and is over ocean sun-glint, rather than land.

Reported RMSE over land for other VSWIR instruments include 0.9—1.3 mm for OCO-2 (Nelson et al., 2016), 1.4—3.7 mm for MERIS (Lindstrot et al., 2012), 0.9—2.0 mm for MODIS (Diedrich et al., 2015), 1.3—3.3 mm for OLCI (Preusker et al., 2021) and up to 2.4 mm for Sentinel-2 (Obregón et al., 2019). The range of $TCWV_{ret}$ simulated in Figure 4 is therefore consistent with typical errors reported for other instruments. Interestingly, Obregón et al. (2019) report a gradient of 0.9

5   between Sentinel-2 and AERONET $TCWV_{ret}$. This is derived from data across multiple sites and times so cannot be compared to our gradients derived from individual LES cases, but indicates that different retrievals may indeed have relationships between TCWV and $TCWV_{ret}$ which are not 1:1, and thus our non-unity a_1 parameters, which scale derived σ_x, are credible.

4 Effect of SZA variation on retrieved properties

4.1 Calculation of TCWV$_{ret}$ accounting for light path at different solar zenith angles

10 Along-path integrated water vapour (IWV) for SZA ranging from 0—60° inclusive in increments of 15° is calculated using ray tracing. The sunlight's horizontal component is in the negative y direction, viewing zenith angle is 0°, and the ray is traced from the top of atmosphere to the centre of each surface footprint. Each partial grid cell encountered has its q weighted by the pressure-corrected path through that cell, before obtaining IWV. The cloud mask is extended by a "shadow mask" where cloud LWP$>1\times10^{-3}$ mm along the solar direct ray path. This IWV is referred to as a TCWV for consistency with standard retrieval

15 terminology, even though it is not directly a measure of the column over the footprint. The Section 3 analysis is then repeated, using the same emulators developed using radiative transfer with SZA=45°, a plane-parallel assumption and footprint column profiles. Different SZAs may change the sensitivities somewhat but we do not expect results substantially outside the range of those considered here.

4.2 Effect of SZA variation on retrieved properties

20 Figure 10(a,d) shows apparent TCWV in ARM_lsconv_36000s (i.e. when convection is most developed) at SZA=15 and 60°, and in Figure 10(c,d) the clear vertical pattern of positive followed by negative biases relative to true TCWV is clear, with greater magnitude and larger regions of continuous positive or negative bias at higher SZA.

Figure 11 shows that this spatial smearing destroys the correspondence between footprint and path TCWV, with r^2 around 0.1 with SZA as small as 30°. This can be compensated only somewhat by spatial smoothing, as shown in Figure 12 shows that

25 even footprints degraded to 300 m are affected by SZA. The calculated σ_x at SZA=0° match those from Figure 9, with biases from the emulator slope parameter in Table 2. Larger SZA in these cases increases the magnitude of this bias, but the difference in σ_x as SZA changes from 15—60° is smaller than the RICO or BOMEX emulator-trend-induced biases. The retrieved σ_x with footprint size tracks reality, suggesting that the horizontal distribution statistics might still be captured even at large SZA. Furthermore, the statistical error estimation from Section 3.1.4 has effectively identical performance regardless of SZA (not

30 shown).

However, Figure 12(b) shows that a VSWIR-retrieved map $TCWV_{ret}$ does not accurately represent the actual spatial variability in TCWV and by extension $PCWV_{PBL}$ even for SZA=15°, and this is a fundamental limitation caused by the solar path through the atmosphere. In fact, the $TCWV_{ret}$ map corresponds better to the TCWV map at the horizontal location where the downward solar path enters the PBL, but improvement in r^2 is limited (Supplementary Figure 13).

5 Figure 13 shows that while the retrieved TCWV distributions are biased, as previously discussed, SZA increases cause only minor visible changes in distribution shape. This indicates that important statistics of the TCWV (and therefore $PCWV_{PBL}$) field can be obtained at the native footprint resolution, despite the poor correspondence of any individual footprint to the column located at that position. The primary advantages of finer spatial resolution are (i) it allows better calculation of σ_ε than at coarser resolution using Section 3.1.4's method, due to the smaller Δr between footprints and (ii) when calculating statistics

10 such as standard deviation on local scales, statistical errors are reduced by the larger number of footprints. For example, Δx=50 m represents approximately 25 times more measurements than MODIS or MERIS. If standard deviation were desired for a 1 km × 1 km region, N=16 from 250 m footprints results in a sampling error of ±17.7 % versus ±3.5 % for N=400 from 50 m footprints.

5 Discussion and conclusions

15 Modern and upcoming VSWIR instruments promise unprecedented horizontal resolution for the study of surface properties, with emphases ranging from mineral regions that are the source of dust (EMIT) to routine observation of agriculture and biodiversity (CHIME) to the full spectrum of study under the NASA 2017 Decadal Survey's Surface Biology and Geology (SBG) designated observable.

This study suggests potential synergies with the Decadal Survey's PBL targeted observable by showing that $PCWV_{PBL}$

20 variability at high spatial resolution can be inferred using the $TCWV_{ret}$ that will be obtained from EMIT observations. While these measurements lack the vertical resolution that is necessary to advance PBL science they provide unique constraint on the mesoscale moisture variability and aggregation within the convective PBL. This analysis is restricted to daytime convective PBLs over land surfaces, which excludes deep convection but still represents a large fraction of meteorological conditions in the tropical to mid latitudes. Importantly these are the precise conditions in which it is suspected that PBL moisture aggregation

25 influences the timing of deep convective events. Furthermore, given the large number of scenes in which we expect to be able to derive these spatial statistics, these observations could prove useful for constraining the manner in which small scale variability is parameterized in shallow convection or unified parameterization schemes. The Isofit development team has curated additional spectra for a universal prior that includes Cryosphere surfaces, but additional work may be necessary to evaluate TCWV over snow, since there is a snow absorption feature near λ=1 μm whose depth depends on snow grain size

30 (Painter et al., 2007), and which overlaps q_v absorption. This may introduce surface-atmosphere covariance that affects the retrieval.

NASA's 2017 Decadal Survey encourages multi-instrument applications and the VSWIR retrievals discussed here could be combined with radio occultation or thermal infrared (TIR) or passive microwave sounders, which have far larger horizontal resolution but obtain vertical profiles. Early explorations of joint VSWIR-TIR retrievals are promising, suggesting that the sensors provide complementary information on both atmospheric and surface properties (Fahlen et al., n.d.). VSWIR could

5 provide a prior constraint on TCWV in a collocated TIR retrieval, or the TIR retrieved $PCWV_{upper}$ could be subtracted from VSWIR TCWV to estimate $PCWV_{PBL}$, with VSWIR also providing the horizontal statistics of clear-sky $PCWV_{PBL}$ variability within the TIR footprint. Another opportunity is to use coincident TIR retrieved profiles of T and q to either build a more appropriate LUT for the VSWIR retrieval, or to select from among pre-computed LUTs.

In Isofit, the atmospheric component contributes a bias to $dTCWV_{ret}/dTCWV$ and may be the largest source of our errors in

10 σ_x, which range from -7 % to +34 % of true σ_x. Development allowing the use prescribed profiles and the ability to assign variability in q to lower altitudes rather than uniform scaling at all altitudes should reduce these errors, as accounting for temperature reduced biases in MERIS $TCWV_{ret}$ (Lindstrot et al., 2012).

This study also showed how SZA as small as 15° significantly degrades the accuracy of retrieved spatial patterns in TCWV, even at coarser resolutions similar to current sensors such as MERIS. However, the TCWV distribution was far less sensitive

15 to SZA. While our results should strongly affect the interpretation of retrieved maps of TCWV from instruments like MERIS, they suggest that moments of the $PCWV_{PBL}$ distribution can be obtained at unprecedented horizontal resolution, which may be of use to developers of modern PBL schemes that use or assume such moments. We note that the LES TCWV distributions and their variation with spatial scale may not be realistic, since they tend to be overly dissipative on scales ≤6 grid cells (Bryan et al., 2003), but it is not clear that these biases affect our conclusion regarding the ability to obtain distributional statistics that

20 represent horizontal variability at scales as small as 50 m.

Future work could address uncertainties that are ignored here, such as topography or cloud 3-D radiative effects via 3-D radiative transfer simulations which avoid several of our assumptions, such as a plane-parallel atmosphere. A particular limitation is that this analysis did not consider vertical structure or PBL height beyond using that derived from the LES mean profiles. In reality there may be errors in locally estimated PBL height, or that obtained from other sensors may be inconsistent

25 with the $\max(d\theta/dz)$ value used here and targeted research on this topic would be worthwhile. Observational evaluation of these uncertainties could be performed using collocated airborne measurements of column water vapour from VSWIR and other instruments such as differential absorption lidar or passive microwave imagers, provided they can obtain sufficiently high spatial resolution. Finally, this work could be extended to other sensors, such as MSI on Sentinel-2, which is not hyperspectral but provides exceptionally fine Δx of approximately 20 m. Additional high-resolution analysis may be required

30 for this, since Figure 9(a,b) imply increases in retrieved σ_x at Δx=20 m for the two simulations that were run at that resolution. Despite these caveats, we have shown ways in which atmospheric correction outputs of surface property retrievals for EMIT can provide unique information on fine-scale PBL water vapour variability, and also identified specific development tasks to improve the quality of its atmospheric outputs. With current tools it therefore seems likely that missions such as EMIT and

CHIME, which are primarily designated as targeting surface observables, can provide unique information to the atmospheric science community.

Acknowledgments: The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The authors also thank Dr. Amin Nehrir and Brian Carroll for assistance with the HALO lidar data.

References

Bedka, K. M., Nehrir, A. R., Kavaya, M., Barton-Grimley, R., Beaubien, M., Carroll, B., Collins, J., Cooney, J., Emmitt, G. D., Greco, S., Kooi, S., Lee, T., Liu, Z., Rodier, S. and Skofronick-Jackson, G.: Airborne Lidar Observations of Wind, Water Vapor, and Aerosol Profiles During The NASA Aeolus Cal/Val Test Flight Campaign, Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-475, 2021.

Bennartz, R. and Fischer, J.: Retrieval of columnar water vapour over land from backscattered solar radiation using the Medium Resolution Imaging Spectrometer, Remote Sens. Environ., 78(3), 274–283, doi:10.1016/S0034-4257(01)00218-8, 2001.

Berk, A., Conforti, P., Kennett, R., Perkins, T., Hawes, F. and van den Bosch, J.: MODTRAN6: a major upgrade of the MODTRAN radiative transfer code, edited by M. Velez-Reyes and F. A. Kruse, p. 90880H., 2014.

Berk, A., Conforti, P. and Hawes, F.: An accelerated line-by-line option for MODTRAN combining on-the-fly generation of line center absorption within 0.1 cm-1 bins and pre-computed line tails, edited by M. Velez-Reyes and F. A. Kruse, p. 947217., 2015.

Borger, C., Beirle, S., Dörner, S., Sihler, H. and Wagner, T.: Total column water vapour retrieval from S-5P/TROPOMI in the visible blue spectral range, Atmos. Meas. Tech., 13(5), 2751–2783, doi:10.5194/amt-13-2751-2020, 2020.

Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J.-C., Khairoutdinov, M., Lewellen, D. C., Lock, A. P., MacVean, M. K., Moeng, C.-H., Neggers, R. A. J., Siebesma, A. P. and Stevens, B.: Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land, Q. J. R. Meteorol. Soc., 128(582), 1075–1093, doi:10.1256/003590002320373210, 2002.

Bryan, G. H., Wyngaard, J. C. and Fritsch, J. M.: Resolution Requirements for the Simulation of Deep Moist Convection, Mon. Weather Rev., 131(10), 2394–2416, doi:10.1175/1520-0493(2003)131<2394:RRFTSO>2.0.CO;2, 2003.

Candela, L., Formaro, R., Guarini, R., Loizzo, R., Longo, F. and Varacalli, G.: The PRISMA mission, in 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 253–256, IEEE., 2016.

Carbajal Henken, C. K., Diedrich, H., Preusker, R. and Fischer, J.: MERIS full-resolution total column water vapor: Observing

horizontal convective rolls, Geophys. Res. Lett., 42(22), 10,074-10,081, doi:10.1002/2015GL066650, 2015.

Couvreux, F., Guichard, F., Austin, P. H. and Chen, F.: Nature of the Mesoscale Boundary Layer Height and Water Vapor Variability Observed 14 June 2002 during the IHOP_2002 Campaign, Mon. Weather Rev., 137(1), 414–432, doi:10.1175/2008MWR2367.1, 2009.

5   Diedrich, H., Preusker, R., Lindstrot, R. and Fischer, J.: Retrieval of daytime total columnar water vapour from MODIS measurements over land surfaces, Atmos. Meas. Tech., 8(2), 823–836, doi:10.5194/amt-8-823-2015, 2015.

Drouin, B. J., Benner, D. C., Brown, L. R., Cich, M. J., Crawford, T. J., Devi, V. M., Guillaume, A., Hodges, J. T., Mlawer, E. J., Robichaud, D. J., Oyafuso, F., Payne, V. H., Sung, K., Wishnow, E. H. and Yu, S.: Multispectrum analysis of the oxygen A-band, J. Quant. Spectrosc. Radiat. Transf., 186, 118–138, doi:10.1016/j.jqsrt.2016.03.037, 2016.

10 Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F. and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, Remote Sens. Environ., 120, 25–36, doi:10.1016/j.rse.2011.11.026, 2012.

Elsey, J., Coleman, M. D., Gardiner, T. D., Menang, K. P. and Shine, K. P.: Atmospheric observations of the water vapour continuum in the near-infrared windows between 2500 and 6600 cm^{-1}, Atmos.

15 Meas. Tech., 13(5), 2335–2361, doi:10.5194/amt-13-2335-2020, 2020.

von Engeln, A. and Teixeira, J.: A Planetary Boundary Layer Height Climatology Derived from ECMWF Reanalysis Data, J. Clim., 26(17), 6575–6590, doi:10.1175/JCLI-D-12-00385.1, 2013.

Fahlen, J. E., Brodrick, P. G., Thompson, D. R., Herman, R. L., Hulley, G., Cawse-Nicholson, K., Green, R. O., Green, J. J., Hook, S. J. and Miller, C. E.: Remote Imaging Spectroscopy of Earth's Surface in UV to Thermal Wavelengths, Remote Sens.

20 Environ., n.d.

Gao, B.-C. and Kaufman, Y. J.: Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels, J. Geophys. Res. Atmos., 108(D13), n/a-n/a, doi:10.1029/2002JD003023, 2003.

Golaz, J.-C., Larson, V. E. and Cotton, W. R.: A PDF-Based Model for Boundary Layer Clouds. Part I: Method and Model Description, J. Atmos. Sci., 59(24), 3540–3551, doi:10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2, 2002.

25 Green, R. O. and Thompson, D. R.: An Earth Science Imaging Spectroscopy Mission: The Earth Surface Mineral Dust Source Investigation (EMIT), in IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, pp. 6262–6265, IEEE., 2020.

Grossi, M., Valks, P., Loyola, D., Aberle, B., Slijkhuis, S., Wagner, T., Beirle, S. and Lang, R.: Total column water vapour measurements from GOME-2 MetOp-A and MetOp-B, Atmos. Meas. Tech., 8(3), 1111–1133, doi:10.5194/amt-8-1111-2015,

30 2015.

Guanter, L., Gómez-Chova, L. and Moreno, J.: Coupled retrieval of aerosol optical thickness, columnar water vapor and surface reflectance maps from ENVISAT/MERIS data over land, Remote Sens. Environ., 112(6), 2898–2913, doi:10.1016/j.rse.2008.02.001, 2008.

Kobayashi, S. and Sanga-Ngoie, K.: The integrated radiometric correction of optical remote sensing imageries, Int. J. Remote

Sens., 29(20), 5957–5985, doi:10.1080/01431160701881889, 2008.

Krutz, D., Müller, R., Knodt, U., Günther, B., Walter, I., Sebastian, I., Säuberlich, T., Reulke, R., Carmona, E., Eckardt, A., Venus, H., Fischer, C., Zender, B., Arloth, S., Lieder, M., Neidhardt, M., Grote, U., Schrandt, F., Gelmi, S. and Wojtkowiak, A.: The Instrument Design of the DLR Earth Sensing
 Imaging Spectrometer (DESIS), Sensors, 19(7), 1622, doi:10.3390/s19071622, 2019.

Kurowski, M. J., Grabowski, W. W., Suselj, K. and Teixeira, J.: The Strong Impact of Weak Horizontal Convergence on Continental Shallow Convection, J. Atmos. Sci., 77(9), 3119–3137, doi:10.1175/JAS-D-19-0351.1, 2020.

Kuze, A., Suto, H., Nakajima, M. and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Appl. Opt., 48(35), 6716, doi:10.1364/AO.48.006716, 2009.

Larson, V. E., Golaz, J.-C. and Cotton, W. R.: Small-Scale and Mesoscale Variability in Cloudy Boundary Layers: Joint Probability Density Functions, J. Atmos. Sci., 59(24), 3519–3539, doi:10.1175/1520-0469(2002)059<3519:SSAMVI>2.0.CO;2, 2002.

Laszlo, I., Stamnes, K., Wiscombe, W. J. and Tsay, S.-C.: The Discrete Ordinate Algorithm, DISORT for Radiative Transfer, in Light Scattering Reviews, Volume 11, pp. 3–65, Springer Berlin Heidelberg, Berlin, Heidelberg., 2016.

Lechevallier, L., Vasilchenko, S., Grilli, R., Mondelain, D., Romanini, D. and Campargue, A.: The water vapour self-continuum absorption in the infrared atmospheric windows: new laser measurements near 3.3 and 2.0 µm, Atmos. Meas. Tech., 11(4), 2159–2171, doi:10.5194/amt-11-2159-2018, 2018.

Lee, C. M., Cable, M. L., Hook, S. J., Green, R. O., Ustin, S. L., Mandl, D. J. and Middleton, E. M.: An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities, Remote Sens. Environ., 167, 6–19, doi:10.1016/j.rse.2015.06.012, 2015.

Lindstrot, R., Preusker, R., Diedrich, H., Doppler, L., Bennartz, R. and Fischer, J.: 1D-Var retrieval of daytime total columnar water vapour from MERIS measurements, Atmos. Meas. Tech., 5(3), 631–646, doi:10.5194/amt-5-631-2012, 2012.

Massie, S. T., Cronk, H., Merrelli, A., O'Dell, C., Schmidt, K. S., Chen, H. and Baker, D.: Analysis of 3D cloud effects in OCO-2 XCO2 retrievals, Atmos. Meas. Tech., 14(2), 1475–1499, doi:10.5194/amt-14-1475-2021, 2021.

Matheou, G. and Chung, D.: Large-Eddy Simulation of Stratified Turbulence. Part II: Application of the Stretched-Vortex Model to the Atmospheric Boundary Layer, J. Atmos. Sci., 71(12), 4439–4460, doi:10.1175/JAS-D-13-0306.1, 2014.

Menang, K. P., Gbode, I. E. and Adeyeri, O. E.: The effect of the differences in near-infrared water vapour continuum models on the absorption of solar radiation, Meteorol. Atmos. Phys., doi:10.1007/s00703-021-00781-6, 2021.

Millán, L., Lebsock, M., Fishbein, E., Kalmus, P. and Teixeira, J.: Quantifying Marine Boundary Layer Water Vapor beneath Low Clouds with Near-Infrared and Microwave Imagery, J. Appl. Meteorol. Climatol., 55(1), 213–225, doi:10.1175/JAMC-D-15-0143.1, 2016.

National Academies of Sciences, Engineering, and M.: Thriving on Our Changing Planet, National Academies Press, Washington, D.C., 2018.

Nelson, R. R., Crisp, D., Ott, L. E. and O'Dell, C. W.: High-accuracy measurements of total column water vapor from the Orbiting Carbon Observatory-2, Geophys. Res. Lett., 43(23), 12,261-12,269, doi:10.1002/2016GL071200, 2016.

Noël, S., Buchwitz, M., Bovensmann, H., Hoogen, R. and Burrows, J. P.: Atmospheric water vapor amounts retrieved from GOME satellite data, Geophys. Res. Lett., 26(13), 1841–1844, doi:10.1029/1999GL900437, 1999.

5   Noël, S., Buchwitz, M. and Burrows, J. P.: First retrieval of global water vapour column amounts from SCIAMACHY measurements, Atmos. Chem. Phys., 4(1), 111–125, doi:10.5194/acp-4-111-2004, 2004.

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer,

10   P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O. and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11(12), 6539–6576, doi:10.5194/amt-11-6539-2018, 2018.

Obregón, M. Á., Rodrigues, G., Costa, M. J., Potes, M. and Silva, A. M.: Validation of ESA Sentinel-2 L2A Aerosol Optical

15   Thickness and Columnar Water Vapour during 2017–2018, Remote Sens., 11(14), 1649, doi:10.3390/rs11141649, 2019.

Painter, T. H., Molotch, N. P., Cassidy, M., Flanner, M. and Steffen, K.: Contact spectroscopy for determination of stratigraphy of snow optical grain size, J. Glaciol., 53(180), 121–127, doi:10.3189/172756507781833947, 2007.

Payne, V. H., Drouin, B. J., Oyafuso, F., Kuai, L., Fisher, B. M., Sung, K., Nemchick, D., Crawford, T. J., Smyth, M., Crisp, D., Adkins, E., Hodges, J. T., Long, D. A., Mlawer, E. J., Merrelli, A., Lunny, E. and O'Dell, C. W.: Absorption coefficient

20   (ABSCO) tables for the Orbiting Carbon Observatories: Version 5.1, J. Quant. Spectrosc. Radiat. Transf., 255, 107217, doi:10.1016/j.jqsrt.2020.107217, 2020.

Pérez-Ramírez, D., Whiteman, D. N., Smirnov, A., Lyamani, H., Holben, B. N., Pinker, R., Andrade, M. and Alados-Arboledas, L.: Evaluation of AERONET precipitable water vapor versus microwave radiometry, GPS, and radiosondes at ARM sites, J. Geophys. Res. Atmos., 119(15), 9596–9613, doi:10.1002/2014JD021730, 2014.

25   Preusker, R., Carbajal Henken, C. and Fischer, J.: Retrieval of Daytime Total Column Water Vapour from OLCI Measurements over Land Surfaces, Remote Sens., 13(5), 932, doi:10.3390/rs13050932, 2021.

Prusa, J. M., Smolarkiewicz, P. K. and Wyszogrodzki, A. A.: EULAG, a computational model for multiscale flows, Comput. Fluids, 37(9), 1193–1207, doi:10.1016/j.compfluid.2007.12.001, 2008.

Rast, M., Ananasso, C., Bach, H., Ben-Dor, E., Chabrillat, S., Colombo, R., Del Bello, U., Feret, J., Giardino, C., Green, R.,

30   Guanter, L., Marsh, S., Nieke, J., CCH, O., Rum, G., Schaepman, M., Schlerf, M., Skidmore, A. and Strobl, P.: Copernicus hyperspectral imaging mission for the environment: Mission requirements document version 2.1, Frascati, Italy., 2019.

Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J.-P., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Fally, S., Flaud, J.-M., Gamache, R. R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W. J., Mandin, J.-Y., Massie, S. T., Mikhailenko, S. N., Miller, C. E.,

Moazzen-Ahmadi, N., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V. I., Perrin, A., Predoi-Cross, A., Rinsland, C. P., Rotger, M., Šimečková, M., Smith, M. A. H., Sung, K., Tashkun, S. A., Tennyson, J., Toth, R. A., Vandaele, A. C. and Vander Auwera, J.: The HITRAN 2008 molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transf., 110(9–10), 533–572, doi:10.1016/j.jqsrt.2009.02.013, 2009.

5 Roy, R. J., Lebsock, M., Millán, L., Dengler, R., Rodriguez Monje, R., Siles, J. V. and Cooper, K. B.: Boundary-layer water vapor profiling using differential absorption radar, Atmos. Meas. Tech., 11(12), 6511–6523, doi:10.5194/amt-11-6511-2018, 2018.

Roy, R. J., Lebsock, M., Millán, L. and Cooper, K. B.: Validation of a G-Band Differential Absorption Cloud Radar for Humidity Remote Sensing, J. Atmos. Ocean. Technol., 37(6), 1085–1102, doi:10.1175/JTECH-D-19-0122.1, 2020.

10 Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S. and Martonchik, J. V.: Reflectance quantities in optical remote sensing—definitions and case studies, Remote Sens. Environ., 103(1), 27–42, doi:10.1016/j.rse.2006.03.002, 2006.

Schneider, A., Borsdorff, T., aan de Brugh, J., Aemisegger, F., Feist, D. G., Kivi, R., Hase, F., Schneider, M. and Landgraf, J.: First data set of H<sub>2</sub>O/HDO columns from the Tropospheric Monitoring Instrument (TROPOMI), Atmos. Meas. Tech., 13(1), 85–100, doi:10.5194/amt-13-85-2020, 2020.

15 Siebesma, A. P., Bretherton, C. S., Brown, A., Chlond, A., Cuxart, J., Duynkerke, P. G., Jiang, H., Khairoutdinov, M., Lewellen, D., Moeng, C.-H., Sanchez, E., Stevens, B. and Stevens, D. E.: A Large Eddy Simulation Intercomparison Study of Shallow Cumulus Convection, J. Atmos. Sci., 60(10), 1201–1219, doi:10.1175/1520-0469(2003)60<1201:ALESIS>2.0.CO;2, 2003.

Sissenwine, N., Dubin, M. and Teweles, S.: US Standard Atmosphere, 1976, Washington, DC., 1976.

20 Stamnes, K., Tsay, S.-C., Wiscombe, W. and Jayaweera, K.: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, Appl. Opt., 27(12), 2502, doi:10.1364/AO.27.002502, 1988.

Stirling, A. J. and Petch, J. C.: The impacts of spatial variability on the development of convection, Q. J. R. Meteorol. Soc., 130(604), 3189–3206, doi:10.1256/qj.03.137, 2004.

25 Suselj, K., Kurowski, M. J. and Teixeira, J.: A Unified Eddy-Diffusivity/Mass-Flux Approach for Modeling Atmospheric Convection, J. Atmos. Sci., 76(8), 2505–2537, doi:10.1175/JAS-D-18-0239.1, 2019.

Szczodrak, M., Austin, P. H. and Krummel, P. B.: Variability of Optical Depth and Effective Radius in Marine Stratocumulus Clouds, J. Atmos. Sci., 58(19), 2912–2926, doi:10.1175/1520-0469(2001)058<2912:VOODAE>2.0.CO;2, 2001.

Teillet, P. M., Guindon, B. and Goodenough, D. G.: On the Slope-Aspect Correction of Multispectral Scanner Data, Can. J.

30 Remote Sens., 8(2), 84–106, doi:10.1080/07038992.1982.10855028, 1982.

Thompson, D. R., Boardman, J. W., Eastwood, M. L. and Green, R. O.: A large airborne survey of Earth's visible-infrared spectral dimensionality, Opt. Express, 25(8), 9186, doi:10.1364/OE.25.009186, 2017.

Thompson, D. R., Natraj, V., Green, R. O., Helmlinger, M. C., Gao, B.-C. and Eastwood, M. L.: Optimal estimation for imaging spectrometer atmospheric correction, Remote Sens. Environ., 216, 355–373, doi:10.1016/j.rse.2018.07.003, 2018.

Thompson, D. R., Cawse-Nicholson, K., Erickson, Z., Fichot, C. G., Frankenberg, C., Gao, B.-C., Gierach, M. M., Green, R. O., Jensen, D., Natraj, V. and Thompson, A.: A unified approach to estimate land and water reflectances with uncertainties for coastal imaging spectroscopy, Remote Sens. Environ., 231, 111198, doi:10.1016/j.rse.2019.05.017, 2019.

Thompson, D. R., Braverman, A., Brodrick, P. G., Candela, A., Carmon, N., Clark, R. N., Connelly, D., Green, R. O., Kokaly, R. F., Li, L., Mahowald, N., Miller, R. L., Okin, G. S., Painter, T. H., Swayze, G. A., Turmon, M., Susilouto, J. and Wettergreen, D. S.: Quantifying uncertainty for remote spectroscopy of surface composition, Remote Sens. Environ., 247, 111898, doi:10.1016/j.rse.2020.111898, 2020.

Thompson, D. R., Kahn, B. H., Brodrick, P. G., Lebsock, M. D., Richardson, M. and Green, R. O.: Spectroscopic Imaging of Sub-Kilometer Spatial Structure in Lower Tropospheric Water Vapor, Atmos. Meas. Tech., doi:10.5194/amt-2020-346, 2021.

Trent, T., Boesch, H., Somkuti, P. and Scott, N.: Observing Water Vapour in the Planetary Boundary Layer from the Short-Wave Infrared, Remote Sens., 10(9), 1469, doi:10.3390/rs10091469, 2018.

vanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., Cheng, A., Couvreux, F., Jiang, H., Khairoutdinov, M., Kogan, Y., Lewellen, D. C., Mechem, D., Nakamura, K., Noda, A., Shipway, B. J., Slawinska, J., Wang, S. and Wyszogrodzki, A.: Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO, J. Adv. Model. Earth Syst., 3(2), n/a-n/a, doi:10.1029/2011MS000056, 2011.

Vermote, E. F., El Saleous, N., Justice, C. O., Kaufman, Y. J., Privette, J. L., Remer, L., Roger, J. C. and Tanré, D.: Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation, J. Geophys. Res. Atmos., 102(D14), 17131–17141, doi:10.1029/97JD00201, 1997.

Wulfmeyer, V., Bauer, H.-S., Grzeschik, M., Behrendt, A., Vandenberghe, F., Browell, E. V., Ismail, S. and Ferrare, R. A.: Four-Dimensional Variational Assimilation of Water Vapor Differential Absorption Lidar Data: The First Case Study within IHOP_2002, Mon. Weather Rev., 134(1), 209–230, doi:10.1175/MWR3070.1, 2006.

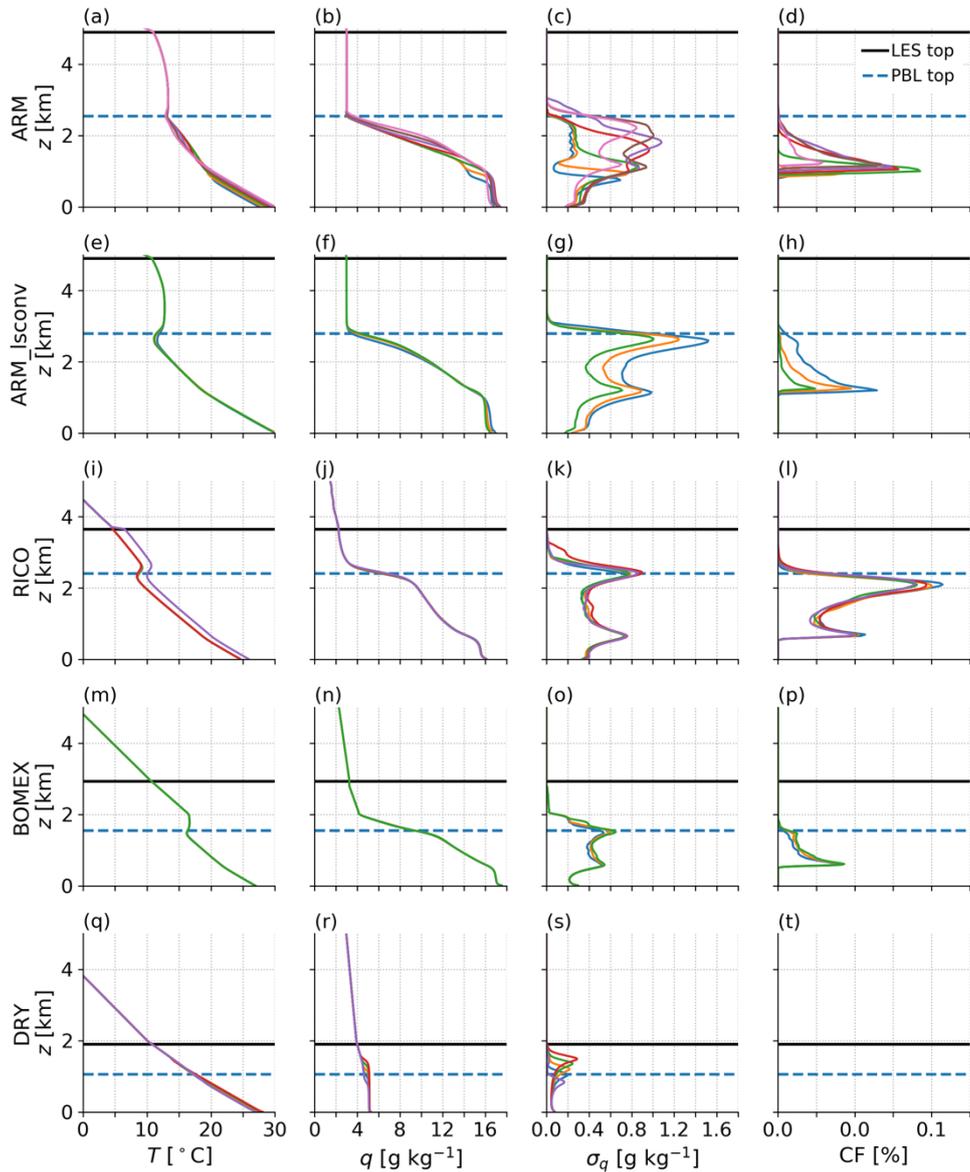



Figure 1. Output all-sky profiles for the (a—d) ARM, (e—h) ARM_lsconv, (i—l) RICO, (m—p) BOMEX and (q—t) DRY LES. On each panel the separate coloured lines represent different timesteps, the black horizontal line is the top of the LES and the dashed blue horizontal line is the PBL height calculated from the first shown timestep, whose lines are in the same blue. The column beginning with (a) is the mean T profile, that with (b) the mean q profile, that with (c) the profile of the spatial standard deviation in q and that with (d) the cloud fraction. Note that due to overlap, the fraction of cloudy columns listed in Table 1 is higher than the peak mean profile cloud fraction.

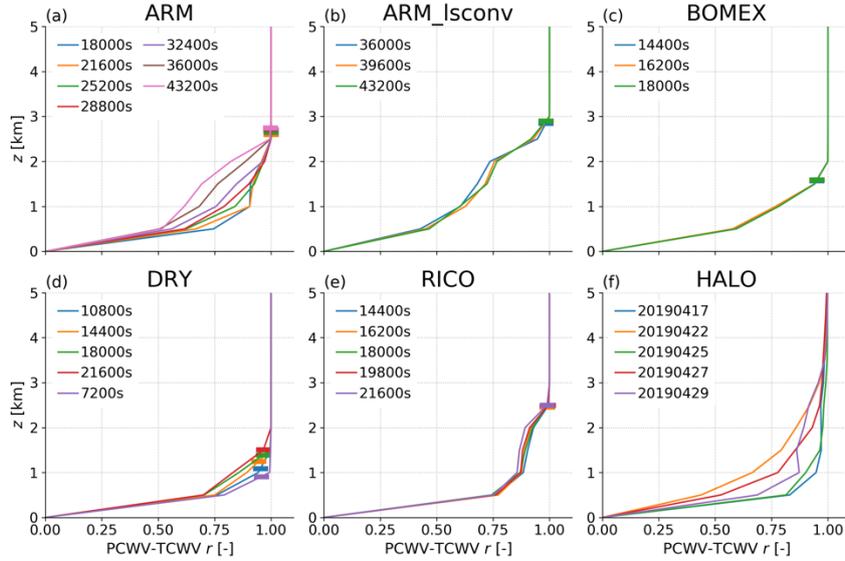



Figure 2. Correlation coefficient between clear-sky partial column water vapour (PCWV) integrated up to given capping altitudes, and the TCWV. (a—e) contain the snapshots of each individual LES run and (f) reproduces the values calculated from High Altitude Lidar Observatory flights over the Pacific (Bedka et al., 2021) as presented in Thompson et al. (2021). The LES profiles also have a horizontal bar appended at the derived PBL top height. The flight data differ from the LES outputs in that horizontal resolution is approximately 3 km along-track, they are dispersed over thousands of km, and the TCWV is only up to 8 km due to the flight altitudes.

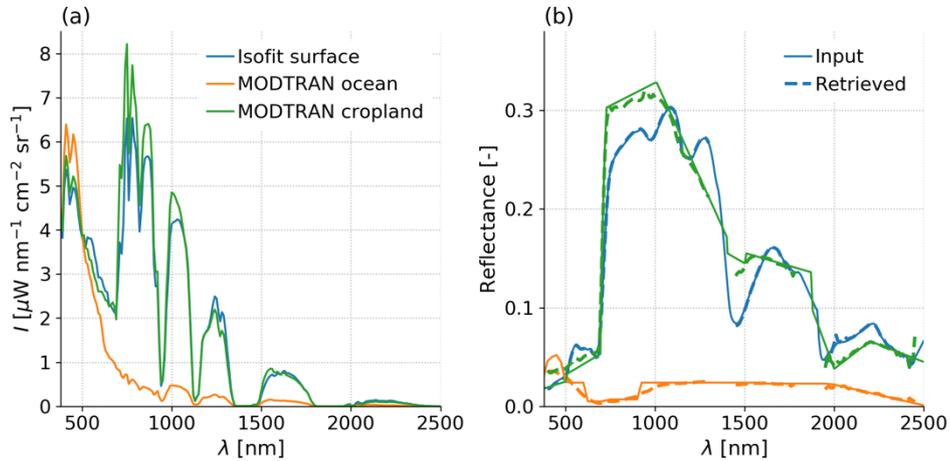



Figure 3. Examples of (a) simulated spectra and (b) used surface reflectances in the forward model (solid lines) and those retrieved by Isofit using EMIT instrument characteristics (dashed lines). Each colour refers to a surface type as listed in the panel (a) legend.

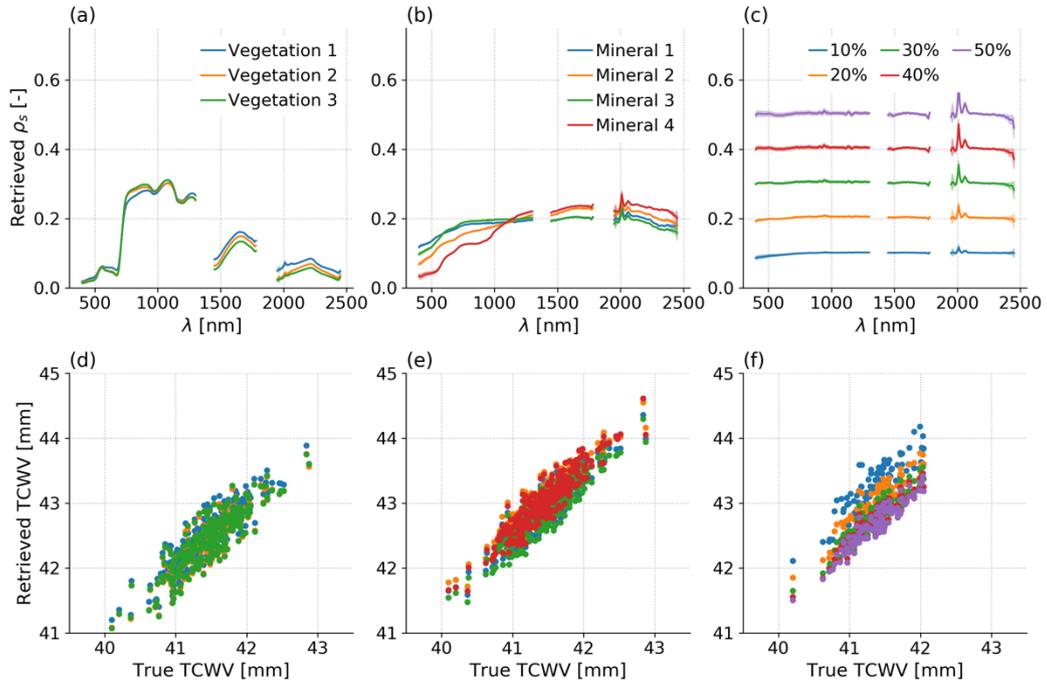



Figure 4. (a—c) retrieved reflectance spectra for (a) vegetation, (b) mineral and (c) spectrally uniform surfaces. Lines show the mean of all simulated retrievals and shading shows ±1σ. (d—f) retrieved TCWV as a function of true TCWV for the same. The Vegetation and Mineral cases use three snapshots (*N*=303) and the Lambertian surfaces just use ARM_18000s (*N*=101).

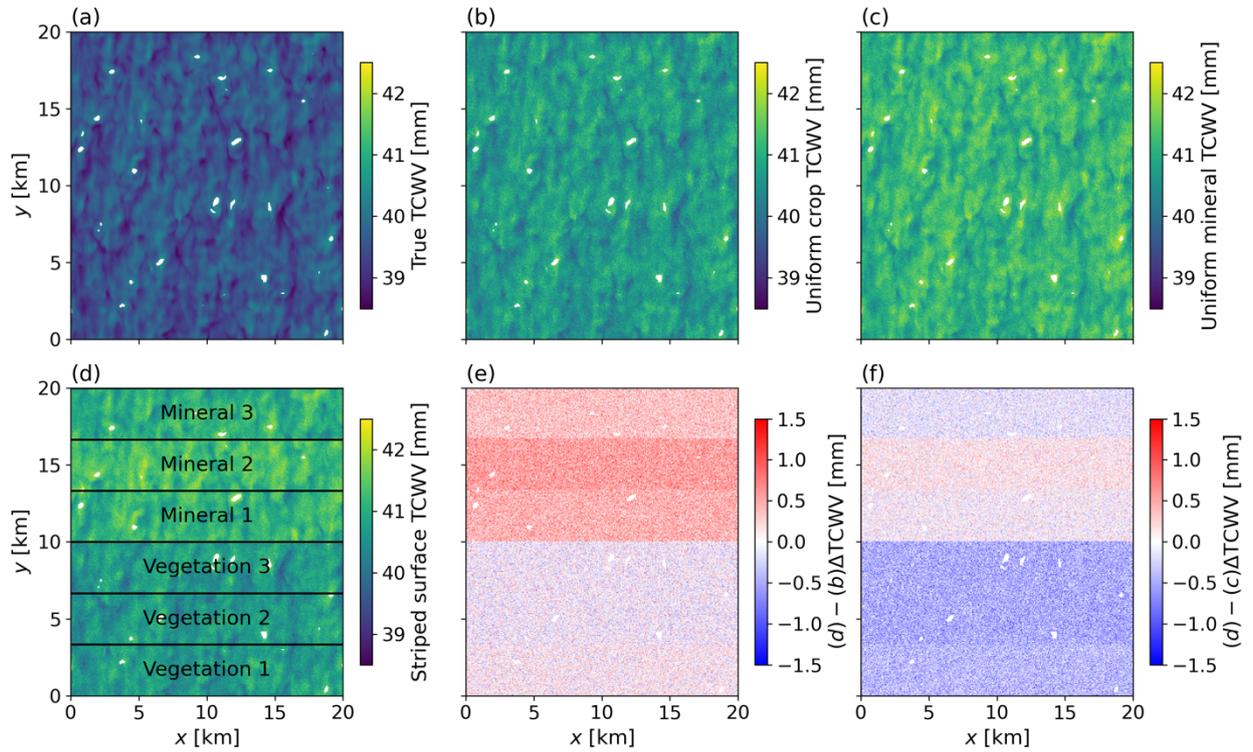



Figure 5. ARM_18000s (a) true TCWV, (b) retrieved TCWV over a uniform vegetated surface, (c) retrieved TCWV over a uniform mineral surface, (d) retrieved TCWV over stripes of uniform surface types as labelled in the figure, (e) difference induced in retrieved TCWV by surface type relative to mixed vegetation as (d) minus (b), (f) difference. Relative to mixed mineral vegetation as (d)minus (c). Clouds are masked in all cases.

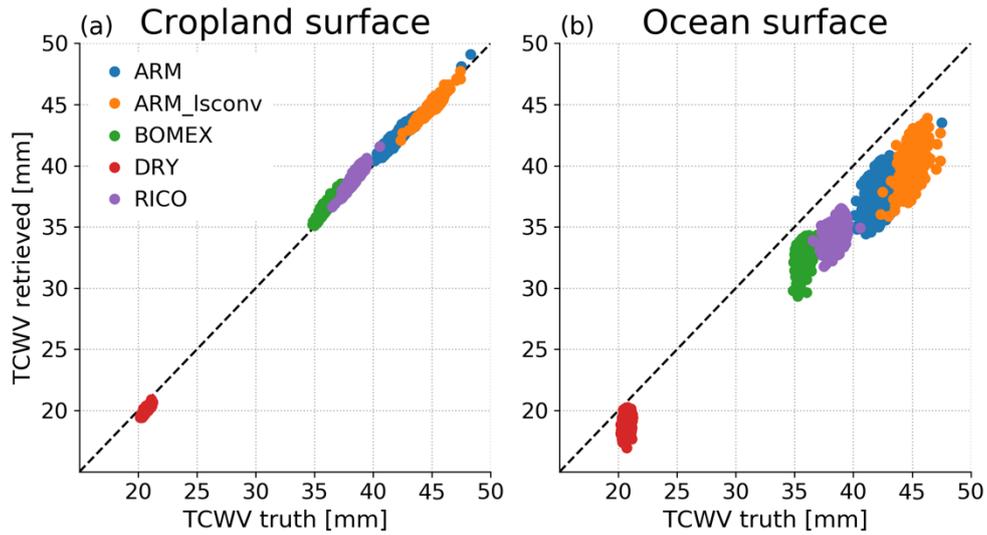



Figure 6. Retrieved TCWV as a function of the truth for all snapshots in each LES case over (a) cropland and (b) ocean. Note that the TCWV values differ from those derived from the LES due to differences in the MODTRAN layer interpolation and calculations.

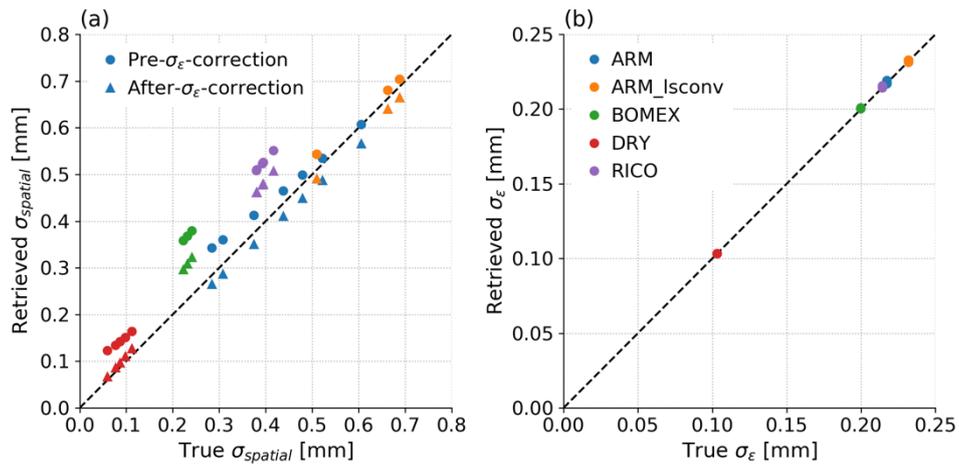



Figure 7. (a) Estimated clear-sky horizontal standard deviation as a function of the true value for each snapshot for raw retrievals (circles) and retrievals after removal of the random component of retrieval error, e.g. that induced by instrumental noise (triangles).
10 **(b) The estimate of retrieval error as in Section 3.1.4 as a function of the true error in each case.**

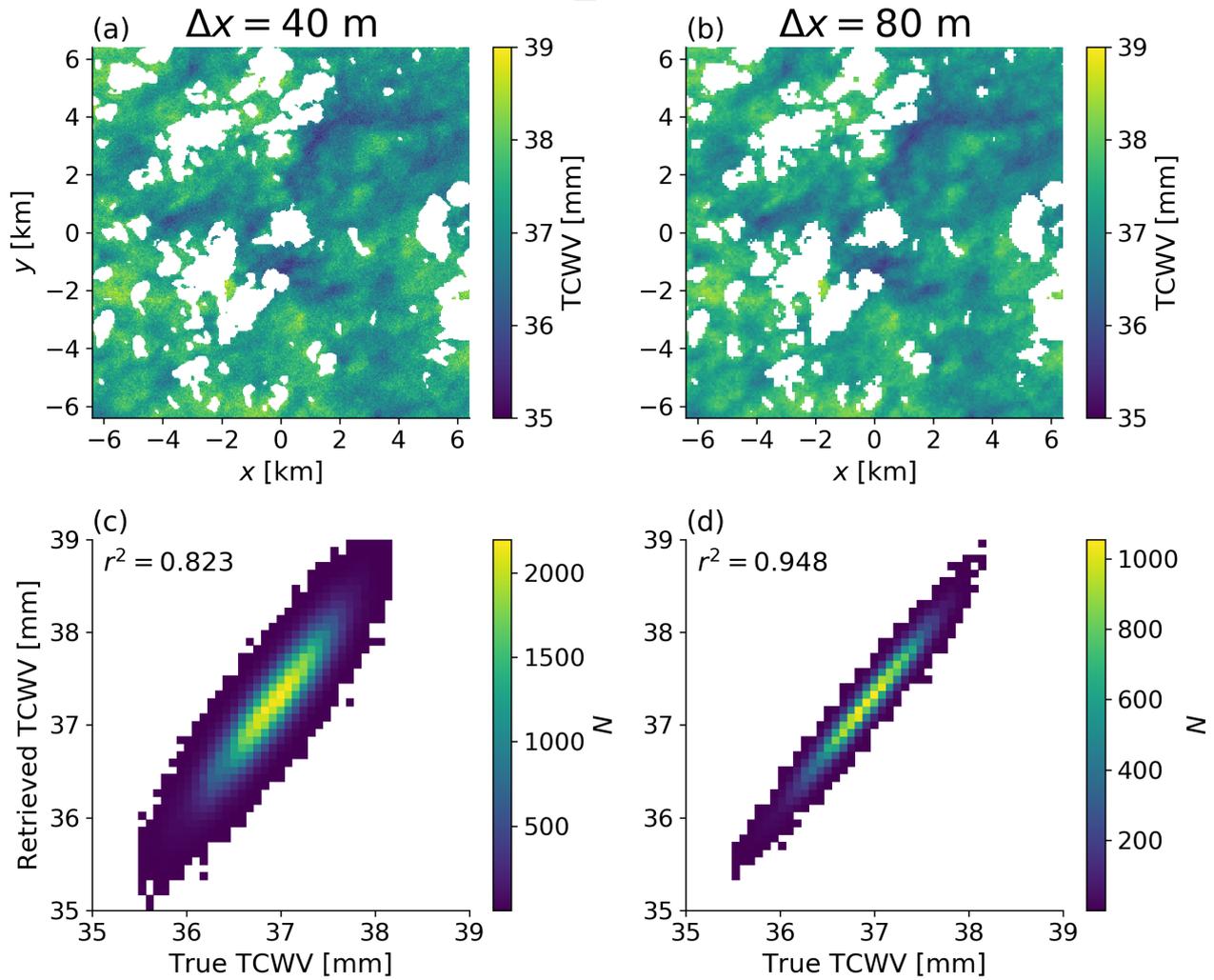



Figure 8. (a) retrieved TCWV at 40 m resolution, (b) retrieved TCWV at 80 m resolution, (c) 2d histogram of retrieved TCWV as a function of the truth at 40 m resolution, (d) 2d histogram of the same at 80 m resolution. The squared Pearson correlation coefficient, r^2, is written in the upper left corner of (c,d).

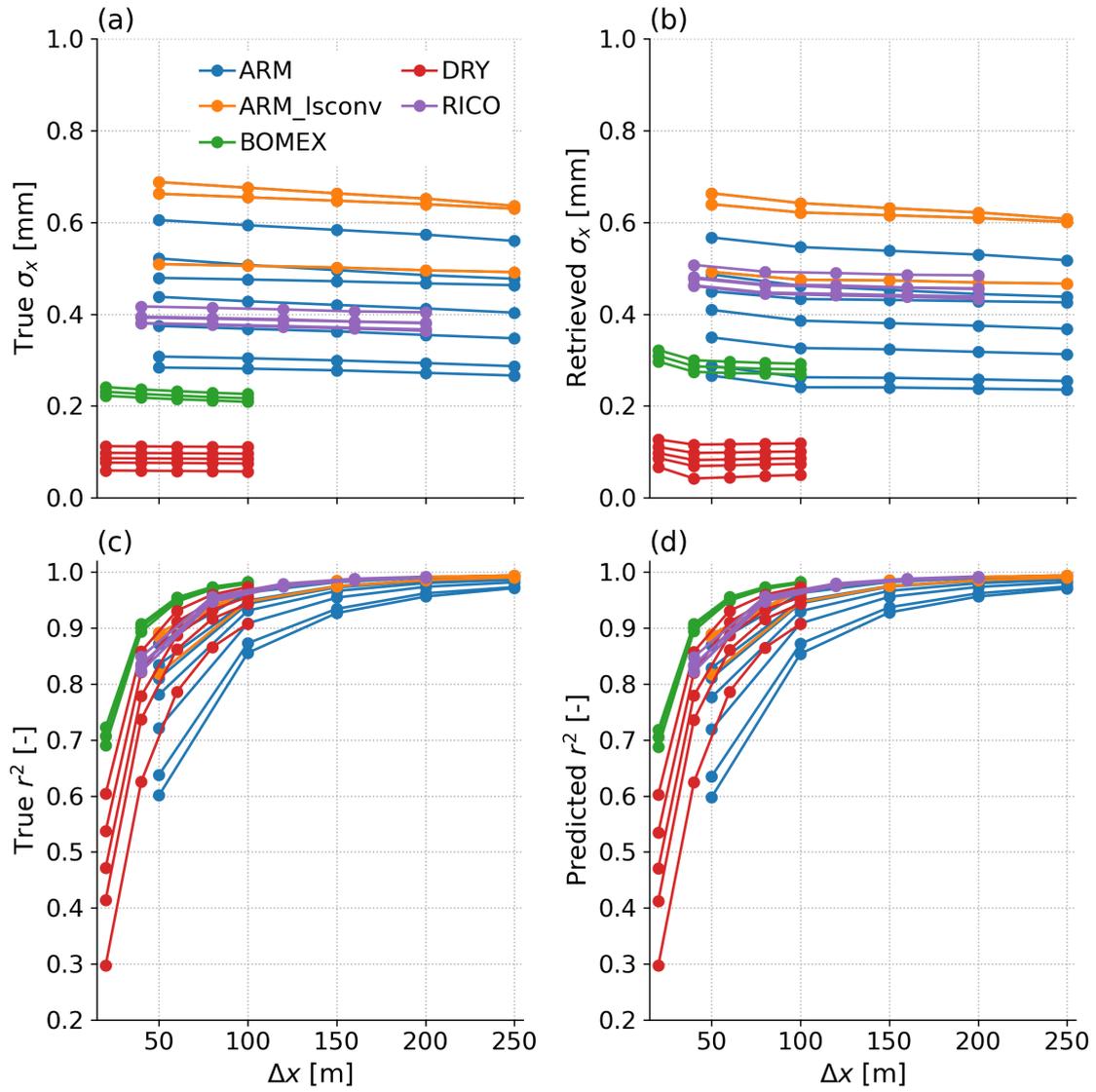



Figure 9. Changes in the true and retrieved statistical properties for LES as a function of spatial resolution Δx. (a) the true standard deviation calculated directly from the LES output, (b) retrieved standard deviation after removing the estimated retrieval error as in Section 3.1.4, (c) r^2 between true TCWV and TCWV$_{ret}$, (d) estimated r^2 using Eq. (7).

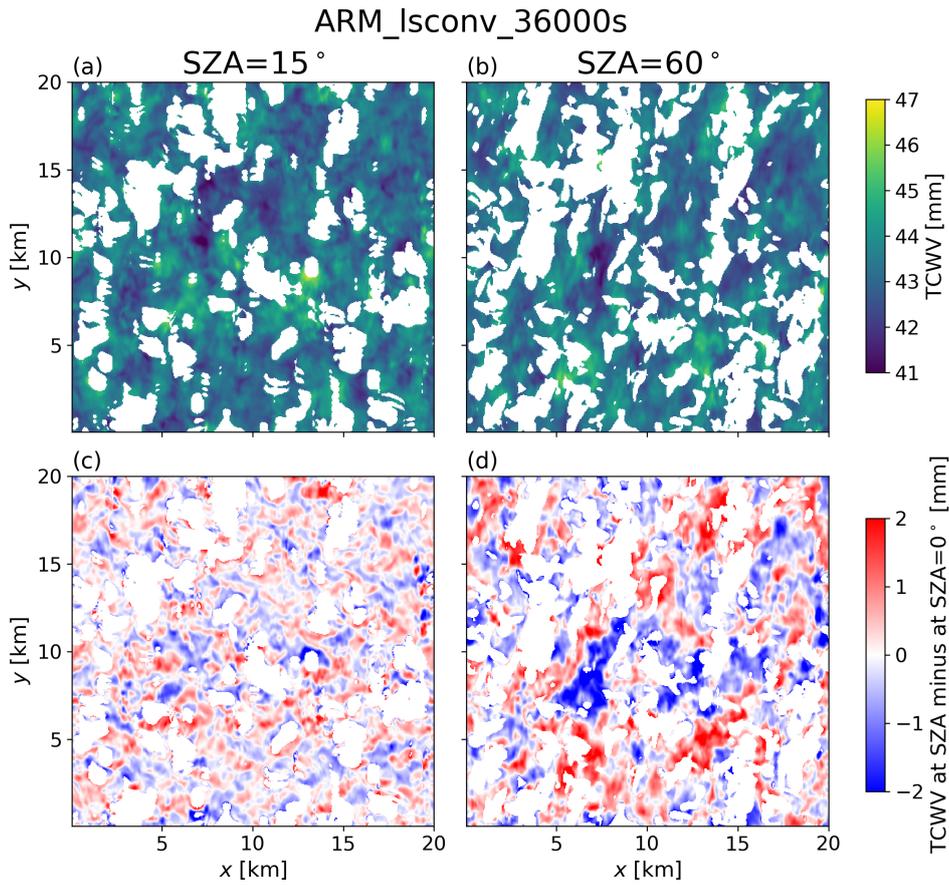



Figure 10. ARM_lsconv_36000s integrated water path calculated along (a) SZA=15° and (b) SZA=60° with the upward path directly up at zenith angle 0°, values labelled TCWV in colour bar for simplicity. (c) shows the difference for each footprint by subtracting the true TCWV at SZA=0° from panel (a), and (d) shows the same for subtracting the SZA=0° value from the SZA=60° value. The "cloud" mask in each case is now extended to include cloud shadows, and the illumination comes from the top of each panel, i.e. sunlight travelling down through the atmosphere has a component in the negative-*y* direction.

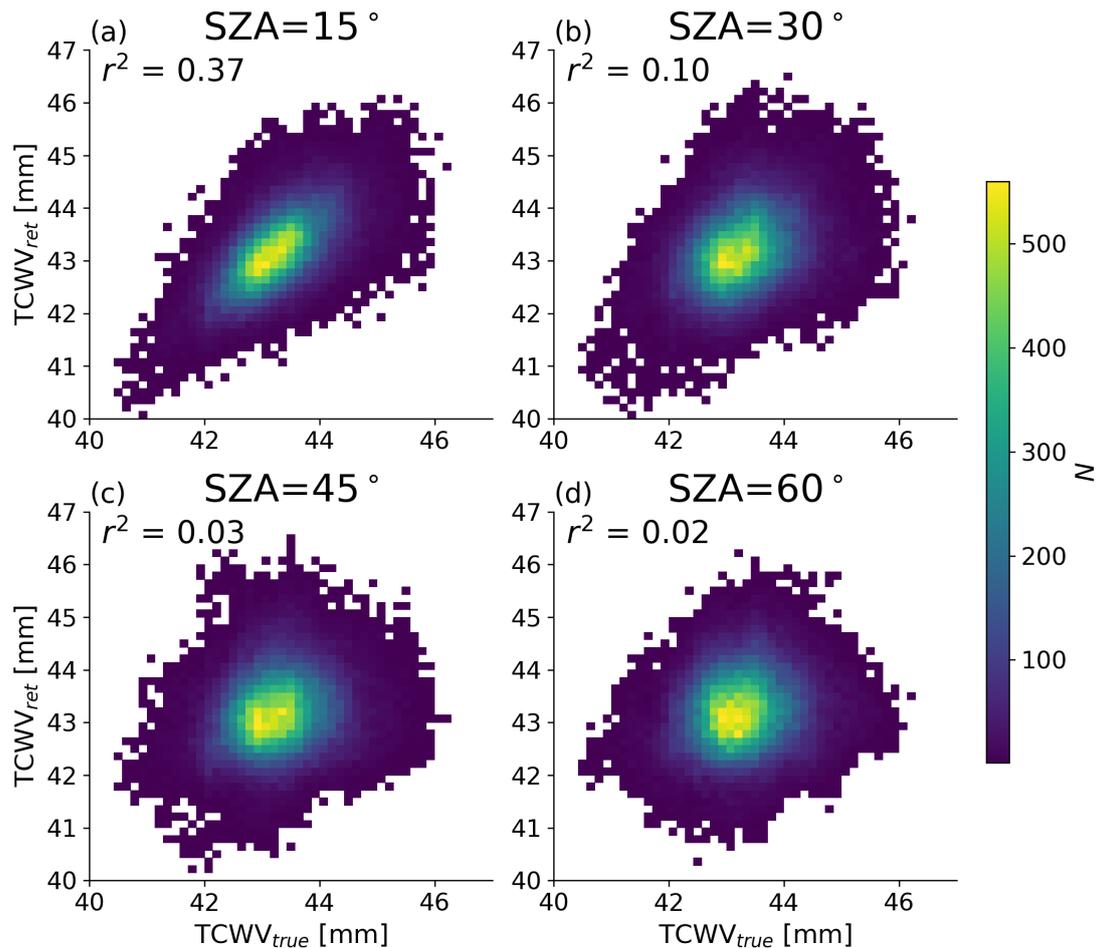
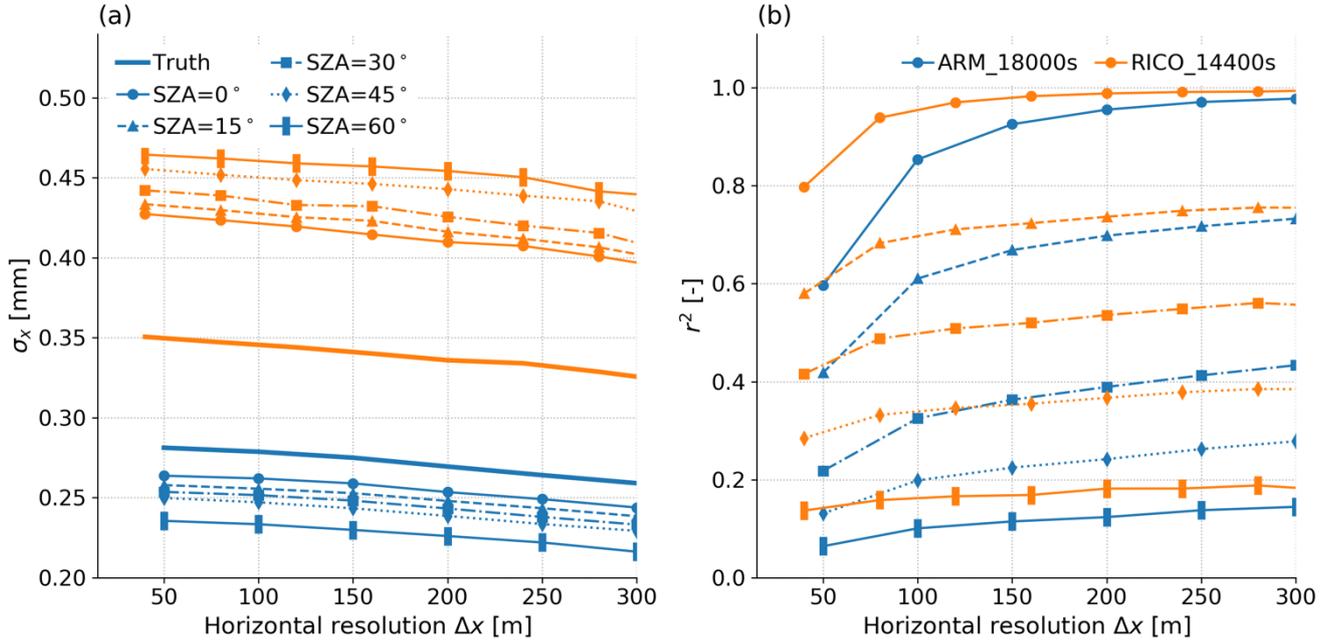



Figure 11. 2d histograms between clear-sky TCWV (true value integrated only in column over footprint) and the retrieved values at the corresponding footprint with SZA of (a) 15°, (b) 30°, (c) 45°, (d) 60°. The r^2 coefficient is on each panel, and the footprint resolution is the native output of Δx=50 m.

5   **Figure 12. Clear-sky TCWV horizontal spatial statistics calculated for ARM_18000s (blue) and RICO_14400s (orange) as a function of the horizontal footprint size. (a) standard deviation σ_x as in Figure 9 and including the random error correction from Section 3.1.4, (b) correlation coefficient between column true TCWV and that retrieved for the same footprint as SZA changes. Each line style represents a different SZA as labelled in the legend of (a).**

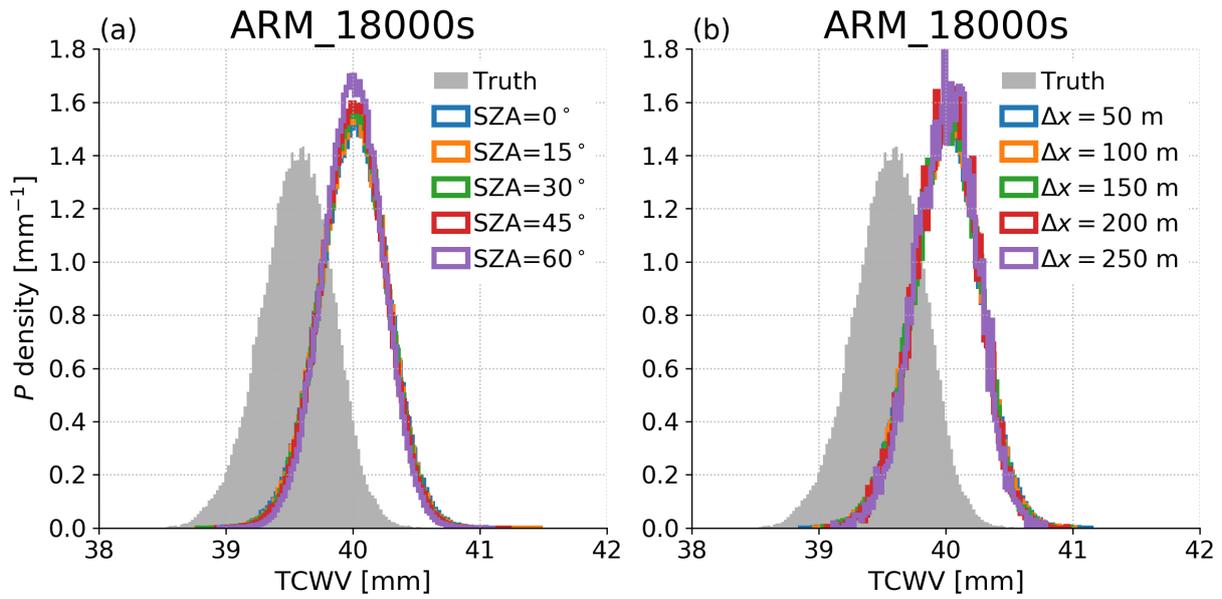



Figure 13. Histograms of footprint estimated clear-sky TCWV, with the truth shown in grey shading. The retrieval estimates are all scaled to remove the variance due to estimated random error. (a) variation with SZA calculated at footprint size Δx=50 m and (b) variation with footprint size at SZA=0°. In both panels the blue histograms are the same.

Table 1. Summary of LES properties. Where ranges are provided, these are the full range of clear-sky mean values from the snapshots used for each case. Row (vi) is the fraction of columns whose integrated liquid water path > 1.3×10^{-3} mm and differs from mean cloud fraction in Figure 1 due to overlap. The TCWV in row (vi) is derived from the combined LES and reanalysis profile, and separated into the LES and reanalysis partial column water vapour amounts in rows (vii) and (viii).

		ARM	ARM_lsconv	RICO	BOMEX	DRY
(i)	Snapshots used	18000s, 21600s, 25200s, 28800s, 32400s, 36000s 43200s	36000s, 39600s, 43200s	14400s, 16200s, 18000s, 19800s, 21600s	14400s, 16200s, 18800s	7200s, 10800s, 14400s, 18000s, 21600s,
(ii)	Domain size [km]	20.0	20.0	20.5	12.8	14.4
(iii)	Δx [m]	50	50	40	20	20
(iv)	LES top [km]	5	5	4	3	2
(v)	PBL top [km]	1—2.7	1—3.2	2.5—2.7	2.1	1.3
(vi)	Columns flagged cloudy [%]	1—21	5—20	24—28	16—19	0.0
(vii)	Clear-sky TCWV [mm]	39.6—42.2	43.3—43.8	36.9—37.0	35.6—35.7	19.8—20.2
(viii)	Clear-sky PCWV$_{LES}$ [mm]	36.2—38.9	40—40.5	33.1	30.6—30.7	9.7—10.2
(ix)	Clear-sky PCWV$_{reanalysis}$ [mm]	3.3	3.3	3.9	5.0	9.9
(x)	Description	Diurnal cycle of midlatitude shallow convection over land	As ARM, perturbed by large-scale convergence	Shallow precipitating trade-wind convection over ocean	Shallow non-precipitating trade-wind convection over ocean	Dry free convection
(xi)	Citation	Brown et al. (2002). REF case in Kurowski et al. (2020)	CON3 case in Kurowski et al. (2020)	vanZantern et al. (2011); Matheou and Chung (2014)	Siebesma et al. (2003); Matheou and Chung (2014)	Matheou and Chung (2014)

Table 2. Emulator parameters relating true TCWV to TCWV$_{ret}$. In Eq. (2) the trend is a_1, the intercept is a_2, and residual σ is the standard deviation used in generating the samples of ϵ. Values are shown $\pm 2\sigma$.

Case	Trend [mm mm^{-1}]	Intercept [mm]	Residual σ [mm]
ARM	0.94±0.02	0.29±0.07	0.22±0.03
ARM_lsconv	0.97±0.04	0.14±0.17	0.23±0.03
BOMEX	1.34±0.06	-1.15±0.22	0.20±0.03
DRY	1.13±0.03	-0.33±0.07	0.10±0.01
RICO	1.22±0.04	-0.77±0.15	0.21±0.03