Supplement of Atmos. Meas. Tech., 16, 1923–1949, 2023
https://doi.org/10.5194/amt-16-1923-2023-supplement

Atmospheric Measurement Techniques
Open Access

Supplement of

Differences in MOPITT surface level CO retrievals and trends from Level 2 and Level 3 products in coastal grid boxes

Ian Ashpole and Aldona Wiacek

Correspondence to: Ian Ashpole (ian.ashpole@smu.ca)

S1. Verification that results obtained using MOPITT V8 data hold using recently released V9

Certain parts of the analysis presented in the paper have been re-done using Version 9 ("V9") data to verify that the main conclusions presented based on Version 8 ("V8") data hold for V9 too:

1. The global land-water sensitivity contrast shown in Fig. 3 (V8; included here as Fig. S1) is also present if Fig. 3 is re-plotted using V9 (Fig, S2). This confirms that the land-water sensitivity contrast remains.
2. The land-water sensitivity differences within coastal L3 grid boxes (demonstrated by comparing L3L and L3W for these grid boxes), shown by the scatterplots in Fig. 4 (V8; included here as Fig. S3), is also present if the analysis is reproduced using V9 and Fig. 4 replotted with these data (Fig. S4).
3. The comparison of a) mean surface level retrieved VMR, and b) temporal trends therein, was repeated for selected grid boxes containing large cities analysed in Sect. 3.4 (based on V8 data), using V9 data. Results for both V8 and V9 are shown in Table S1, below, and similar differences to V8 exist in V9. Although this analysis is restricted to L3L and L3W only, given that the L2 → L3 processing method is unchanged reason to expect that similar differences would emerge for L3O V9 subsets too.

(Note that due to time and data storage limitations (owing to the fact that V9 was released after this study was completed), this V9 analysis has been restricted to the data years 2010-2015 inclusive for results 1 and 2 above. The clarity of the results gives confidence that they would remain valid if the whole study period considered for V8 had been reanalysed using V9. L3L and L3W time series for the full period studied using V8 were able to be obtained for V9 for the analysis leading to result 3 above).

Table S1. Mean retrieved surface level VMR, and its temporal trend, in L3L and L3W, for L3 grid boxes containing selected cities from Sect. 3.4 of the submitted manuscript, using V8 and V9.

		V8			*V9*			*V9 conclusion same as V8?*
		L3L	*L3W*	*Δ (L–W)*	*L3L*	*L3W*	*Δ (L–W)*	
Bangkok	*Mean*	*314.4*	*261.3*	*53.1*	*286.4*	*268.3*	*18.1*	*Y*
	Trend	*-3.03*	*-2.00*	*-1.03*	*-2.93*	*-2.21*	*-0.72*	*Y*
Dubai	*Mean*	*180.0*	*163.3*	*16.7*	*186.6*	*168.8*	*17.8*	*Y*
	Trend	*-2.90*	*-0.90*	*-2.00*	*-2.97*	*-1.06*	*-1.91*	*Y*
Hong Kong	*Mean*	*336.1*	*260.1*	*76.0*	*307.2*	*270.5*	*36.7*	*Y*
	Trend	*-8.06*	*-3.55*	*-4.51*	*-7.50*	*-4.37*	*-3.13*	*Y*
Miami	*Mean*	*160.7*	*143.5*	*17.2*	*158.7*	*149.8*	*8.9*	*Y*
	Trend	*-1.52*	*-0.75*	*-0.77*	*-1.44*	*-1.08*	*-0.36*	*Y*
Sydney	*Mean*	*94.0*	*86.8*	*7.2*	*89.2*	*88.6*	*0.6*	*Y*
	Trend	*-0.74*	*-0.24*	*-0.50*	*-0.58*	*-0.42*	*-0.16*	*Y*
Toronto	*Mean*	*238.4*	*254.5*	*-16.1*	*240.7*	*255.9*	*-15.2*	*Y*
	Trend	*-1.09*	*-1.99*	*0.9*	*-1.71*	*-1.87*	*0.16*	*Y*

V8

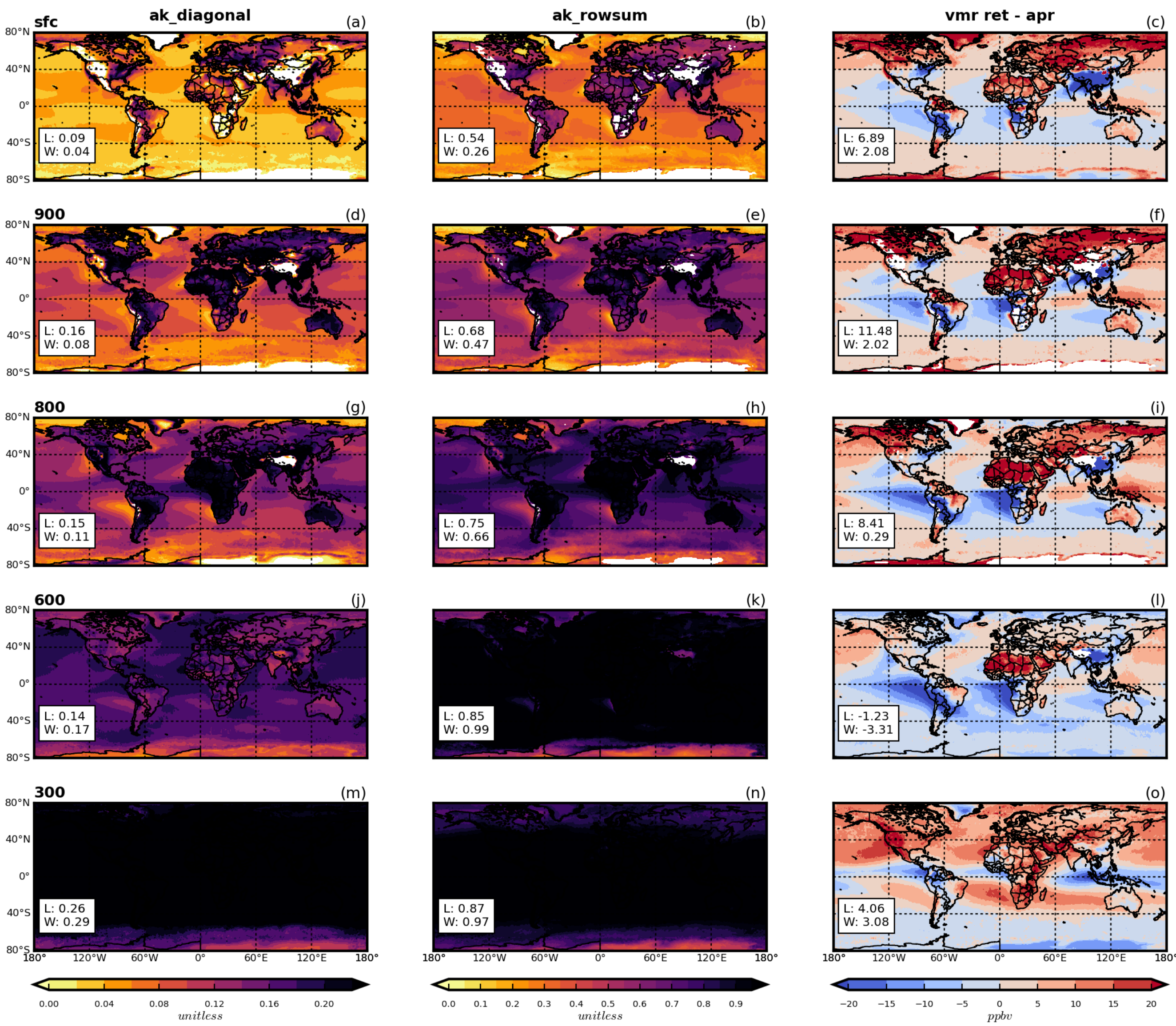


Figure S1. Mean sensitivity metrics from MOPITT L3 data (Version 8), averaged across the entire study period (September 2001 – February 2019, inclusive). Shown are AK diagonal values (left column), AK rowsums (center column) and VMR retrieved minus a priori values (right column) for the following levels of the retrieved profile: surface (top row), 900 hPa (second row), 800 hPa (third row), 600 hPa (fourth row), and 300 hPa (bottom row). Values in white boxes correspond to mean values across all land ("L") and water ("W") L3 grid boxes.

NOTE: This is a reproduction of Fig. 3 from the submitted manuscript.

This figure should be compared to Fig. S2 shown on next page

V9

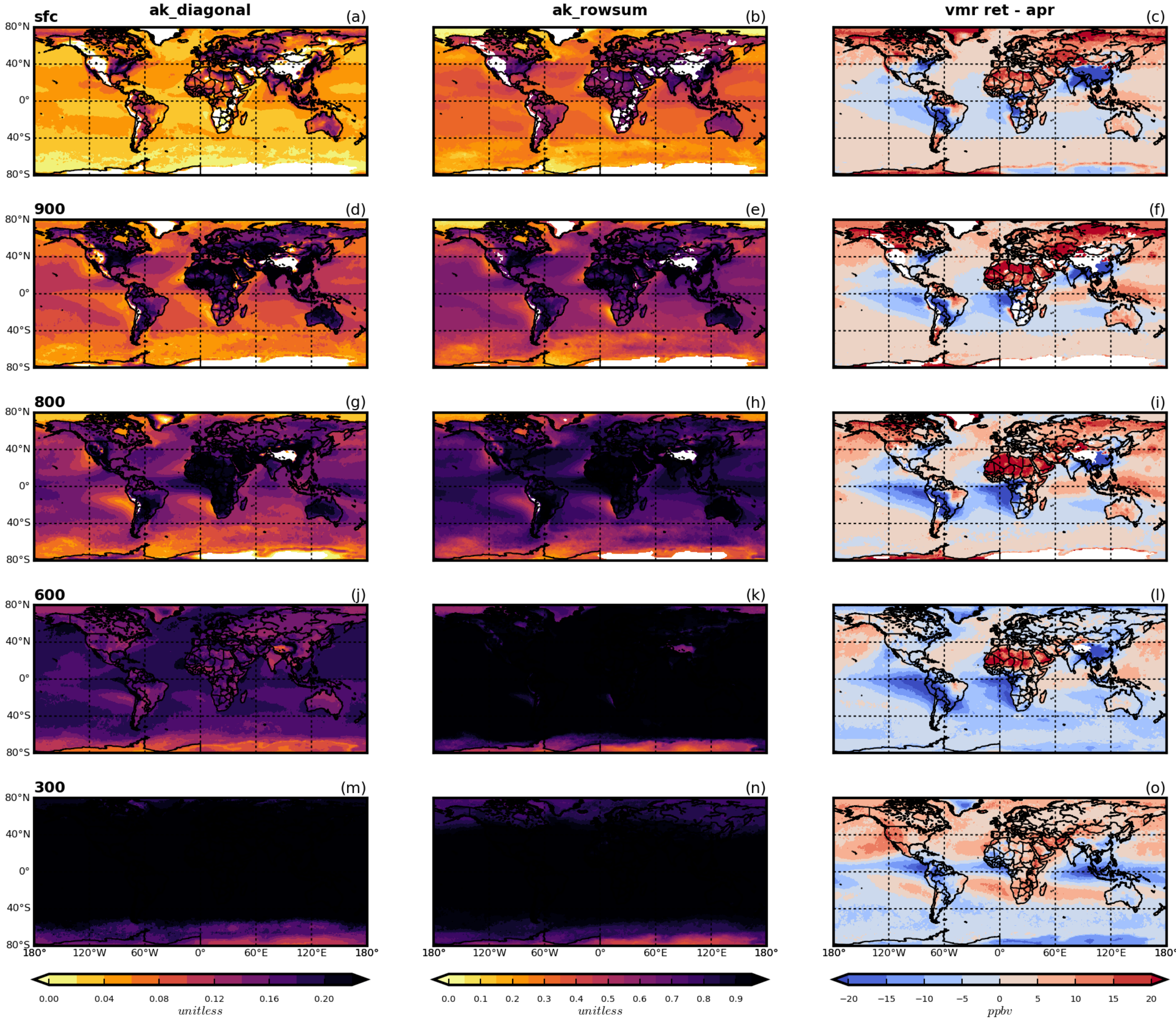


Figure S2. As Fig S1 except using MOPITT Version 9 data and for the sub-period 2010-2015 (inclusive), as explained in Sect. S1.

This figure should be compared to Fig. S2 shown on the previous page

V8

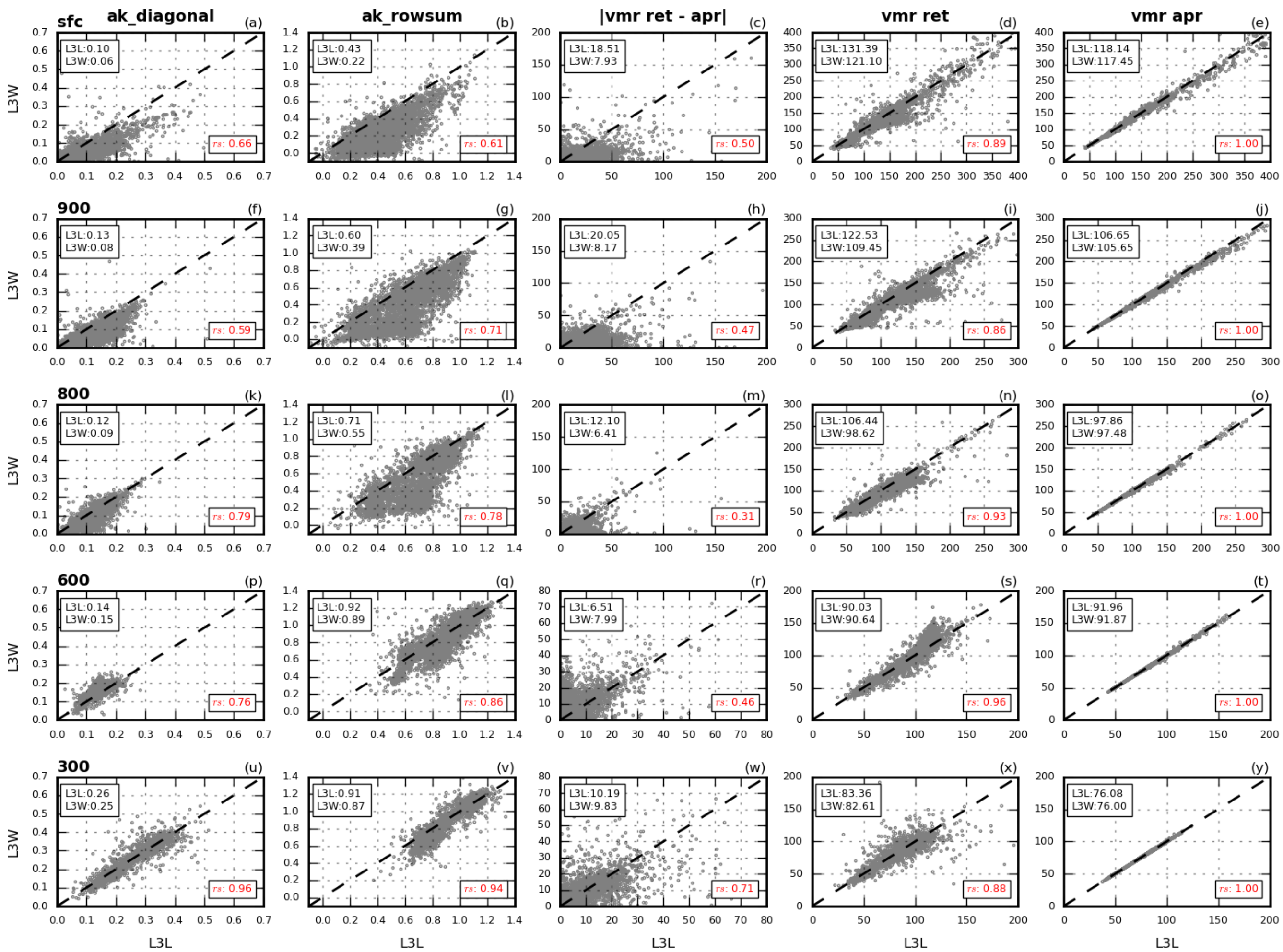


Figure S3. Mean sensitivity metrics and VMRs (retrieved and a priori) from coastal L3 grid boxes, from MOPITT V8 data. Values compared in the scatterplots are mean values from matched L3L and L3W retrievals within these grid boxes. "Matched" means that only days when both L3L and L3W are present, and the L3O surface index is mixed, are used to create the mean values analysed. Shown are AK diagonal values (left column), AK rowsums (second column), absolute VMR retrieved minus a priori values (third column), retrieved (fourth column) and a priori (fifth column) VMRs, for the following levels of the retrieved profile: surface (top row), 900 hPa (second row), 800 hPa (third row), 600 hPa (fourth row), and 300 hPa (bottom row). Values in boxes in the top-left corner of each panel correspond to mean values across all L3L and L3W grid boxes.

NOTE: This is a reproduction of Fig. 4 from the submitted manuscript.

This figure should be compared to Fig. S4 shown on the next page

V9

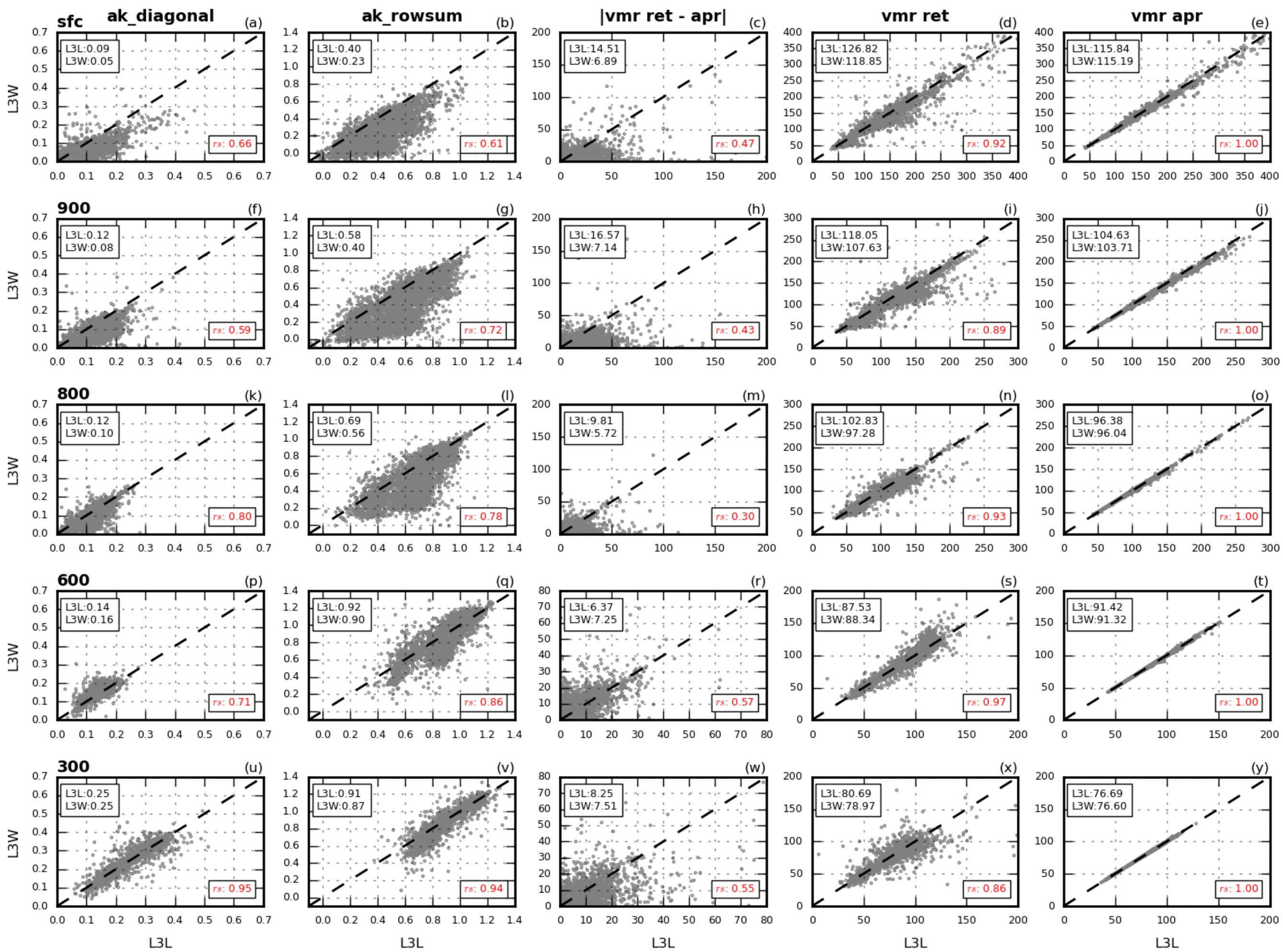


Figure S4. As Fig. S1c except using MOPITT Version 9 data and for the sub-period 2010-2015 (inclusive), as explained in Sect. S1.

This figure should be compared to Fig. S3 shown on the previous page

S2. Alternative version of Fig. 3

Figure S5 is an alternative version of Fig. 3, with colour scales modified to better show spatial patterns in AK diagonal values and rowsums at mid- and upper- troposphere levels (600 and 300 hPa, respectively). The point made in the main text (Sect. 3.1.1. "Land-water contrast in MOPITT sensitivity; global context"), that the land-water sensitivity contrast decreases with height through the profile, justifying focus on the surface level of the retrieved profile, holds with this version of the figure.

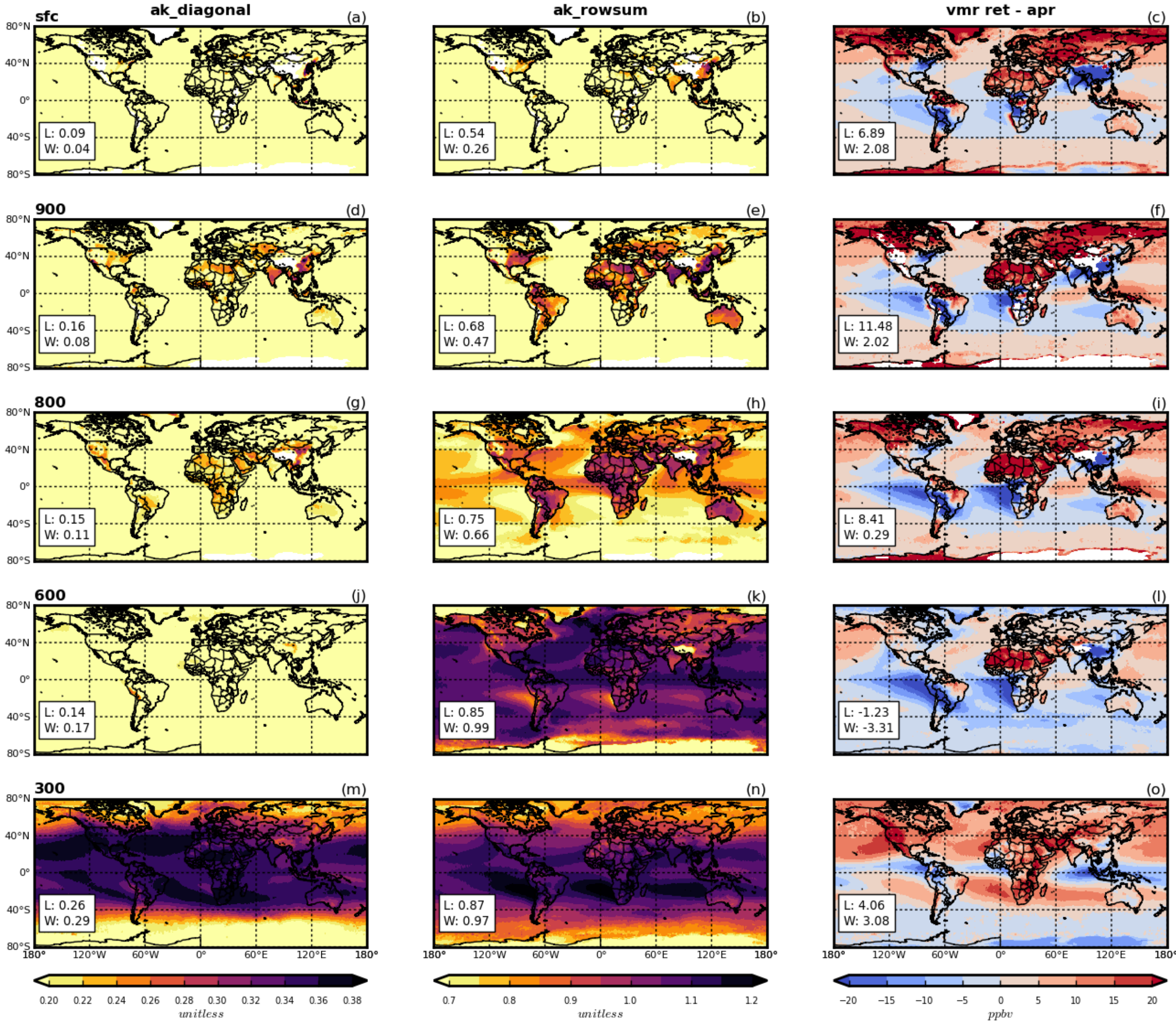


Figure S5. Mean sensitivity metrics from MOPITT L3 data, averaged across the entire study period (September 2001 – February 2019, inclusive). Shown are AK diagonal values (left column), AK rowsums (center column) and VMR retrieved minus a priori values (right column) for the following levels of the retrieved profile: surface (top row), 900 hPa (second row), 800 hPa (third row), 600 hPa (fourth row), and 300 hPa (bottom row). Values in white boxes correspond to mean values across all land ("L") and water ("W") L3 grid boxes.

S3. Analysis of latitudinal and seasonal variability in the land-water retrieval sensitivity contrast

The analysis presented in Section 3.1.1. ("Land-water contrast in MOPITT sensitivity; Global context") is based on data averaged across the whole study period, and is focused on the global scale. Figure S6 is a reproduction of Fig. 3, but also showing seasonal mean data (for surface-level sensitivity metrics only). This clearly shows that the global land-water contrast in retrieval sensitivity holds irrespective of season, with none of the seasonal plots looking very different from the plot for "all-seasons" (also referred to as "long-term mean" or "ltm"; top row of Fig. S6).

There are no clear and obvious patterns when breaking this analysis down by latitude (Figure S7), although there are some details to note. Averaged across the year, there is a general tendency for a greater land-water sensitivity contrast in the Northern Hemisphere than Southern Hemisphere, which is clearest in AK rowsums but also evident in absolute retrieved minus a priori VMR values. However, when breaking this analysis down by season, the pattern becomes less clear. Land-water AK rowsum differences tend to vary least by season in the tropical regions (between 30^{o} South and 30^{o} North) and show the greatest contrast in the midlatitudes (30^{o} – 60^{o}) in the respective hemisphere's spring and summer months, with smallest differences in the winter months. This might be explained by thermal contrast peaking over land in the summer months and reaching a minima in winter months, when the surface can reach freezing temperatures and inverted temperature gradients (i.e. the temperature close to the surface is cooler than the air above) may occur, limiting retrieval sensitivity at the lowest retrieval levels (as discussed in Ashpole and Wiacek, 2020). The land-water differences in absolute retrieved minus a priori VMR values do not necessarily reflect this pattern, however, although this could be due to retrieved minus a priori VMR values being affected by factors in addition to retrieval sensitivity, such as the accuracy of the a priori VMR with respect to the "true" VMR.

In conclusion, there is some seasonal and latitudinal variance in the magnitude of the land-water contrast in retrieval sensitivity, but sensitivity is greater over land than water in nearly all circumstances.

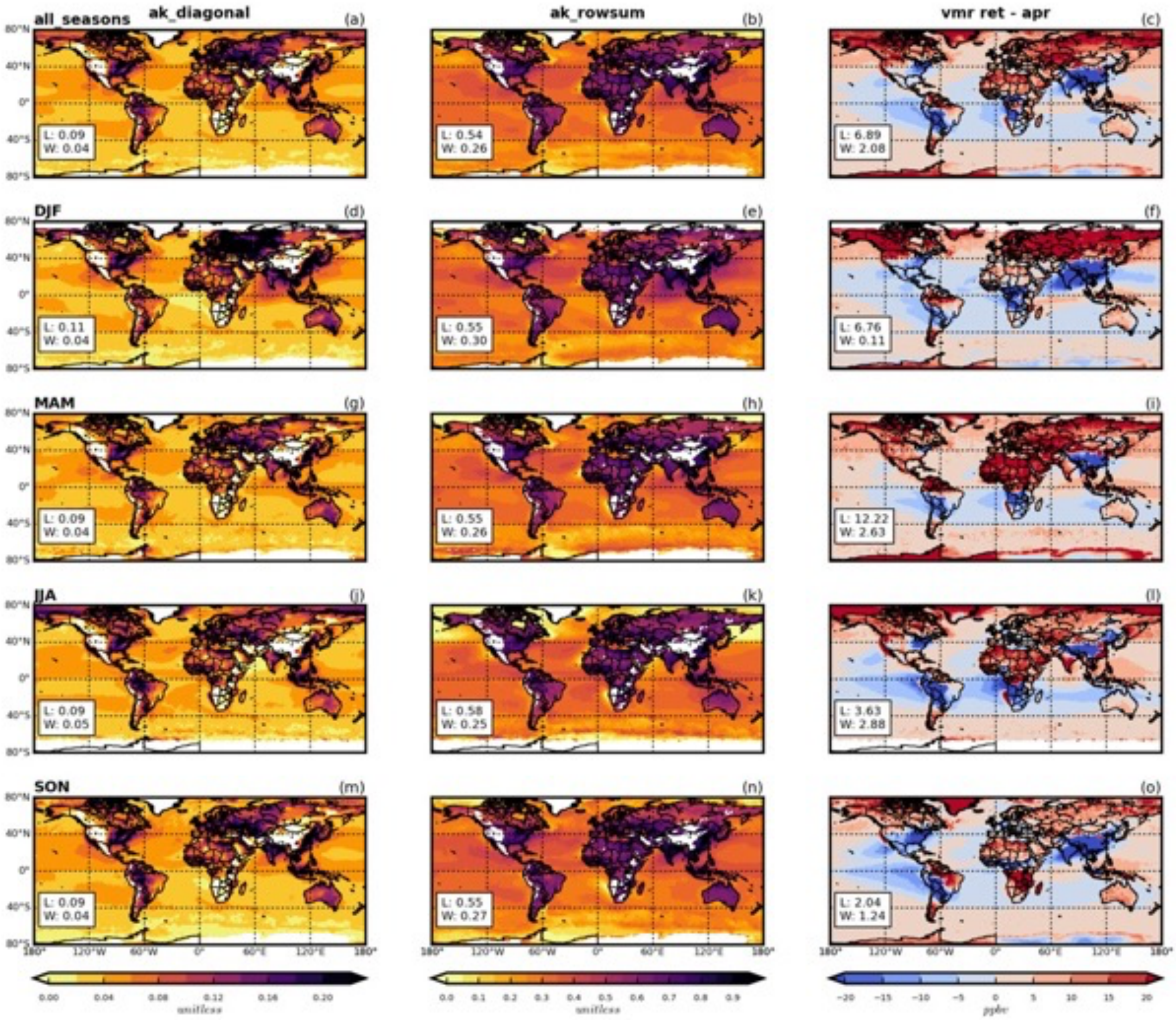
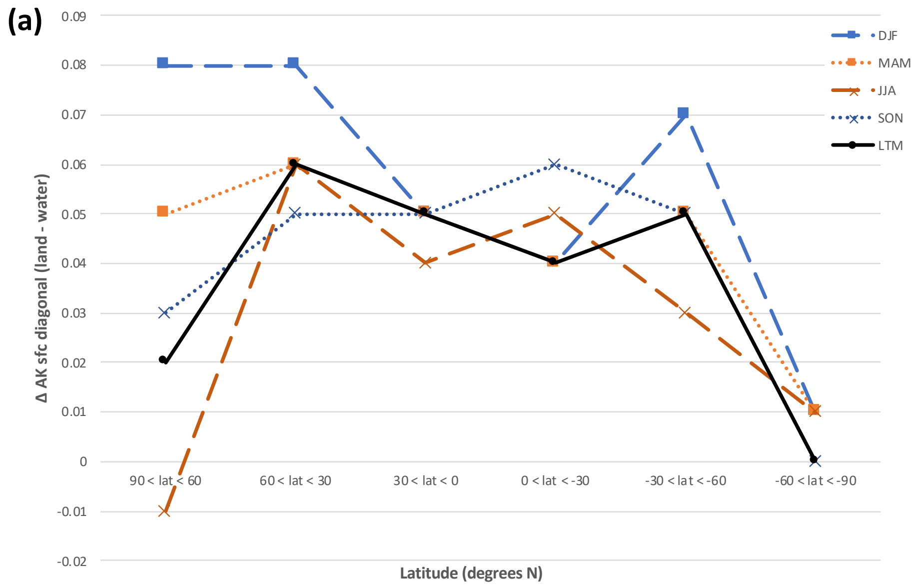
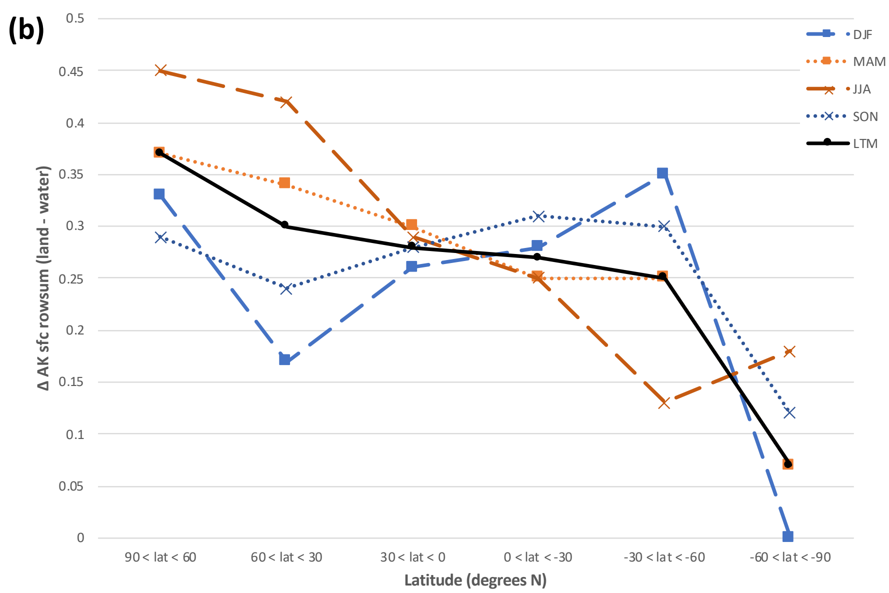


Figure S6. Mean sensitivity metrics from MOPITT L3 data, averaged across the entire study period (top row), and only for the seasons DJF (second row), MAM (third row), JJA (fourth row), and SON (bottom row). Shown are AK diagonal values (left column), AK rowsums (center column) and VMR retrieved minus a priori values (right column) for the surface levels of the retrieved profile only. Values in white boxes correspond to mean values across all land ("L") and water ("W") L3 grid boxes.
(DJF = December – January – February; MAM = March – April – May; JJA = June – July – August; SON = September – October – November)

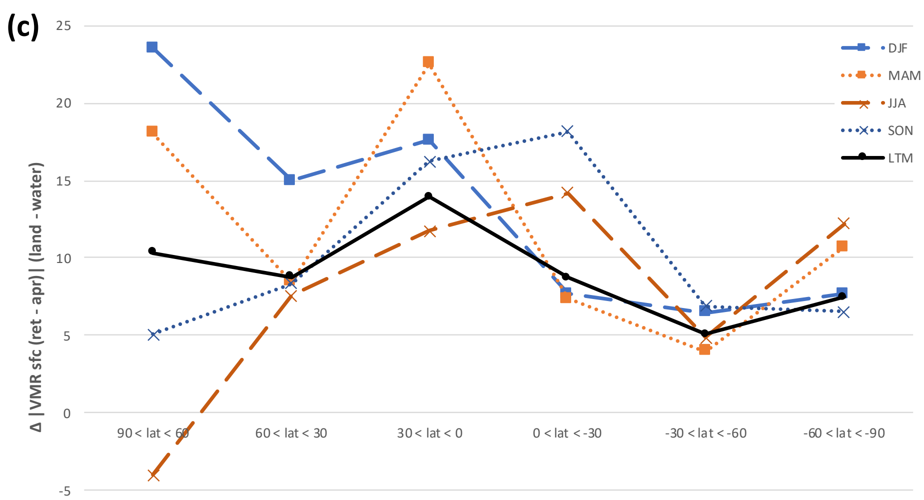
Figure S7 (next page). Land – water differences in mean sensitivity metrics from MOPITT L3 data, averaged across the entire study period (solid black line with filled circles), and only for the seasons DJF (dashed blue line with filled squares), MAM (dotted orange line with filled squares), JJA (dashed orange lines with X symbols), and SON (dotted blue line with X symbols). Averages are taken for the 30^{o} latitudinal bands shown on the x-axes.
Sensitivity metrics shown are (a) AK diagonal values, (b) AK rowsums, and (c) VMR retrieved minus a priori value magnitudes (i.e. ignoring whether the difference is positive or negative), for the surface levels of the retrieved profile only.



(a)
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
Δ AK sfc diagonal (land - water)
90 < lat < 60
60 < lat < 30
30 < lat < 0
0 < lat < -30
-30 < lat < -60
-60 < lat < -90
Latitude (degrees N)
DJF
MAM
JJA
SON
LTM



(b)
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Δ AK sfc rowsum (land - water)
90 < lat < 60
60 < lat < 30
30 < lat < 0
0 < lat < -30
-30 < lat < -60
-60 < lat < -90
Latitude (degrees N)
DJF
MAM
JJA
SON
LTM



(c)
25
20
15
10
5
0
-5
Δ |VMR sfc (ret - apr)| (land - water)
90 < lat < 60
60 < lat < 30
30 < lat < 0
0 < lat < -30
-30 < lat < -60
-60 < lat < -90
DJF
MAM
JJA
SON
LTM

S4. Alternative version of Fig. 4

Figure S8 is an alternative version of Fig. 4, with signed values for retrieved minus a priori (ret-apr) VMRs, as opposed to absolute values in the version presented in the main text. The point made in the main text (Sect 3.1.2. “Land-water contrast in MOPITT sensitivity; analysis of coastal L3 grid boxes”), that ret-apr VMRs deviate more strongly from their a priori values in the lower troposphere (LT) than at MT and UT levels, holds with this version of the figure.

S5. Alternative version of Fig. 5

Figure S9 is an alternative version of Fig. 5, with signed values for retrieved minus a priori (ret-apr) VMRs, as opposed to absolute values in the version presented in the main text. The point made in the main text (Sect 3.2.2 “Differences in retrieved VMRs and temporal trends, and their relation to the land-water sensitivity contrast: L3L vs L33”), that greater land-water sensitivity differences tend to be associated with greater retrieved VMR differences holds with this version of the figure.

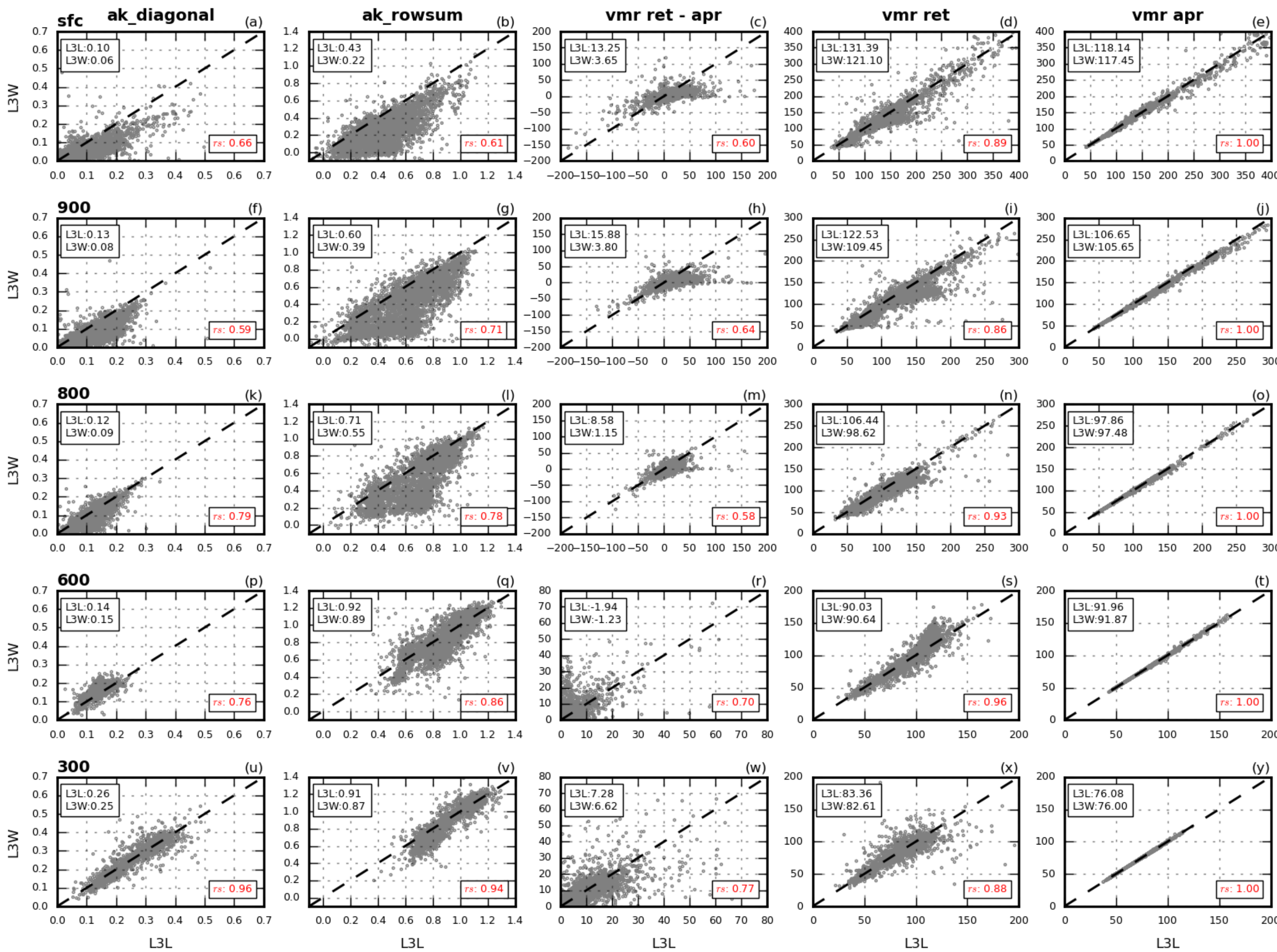


Figure S8. Mean sensitivity metrics and VMRs (retrieved and a priori) from coastal L3 grid boxes. Values compared in the scatterplots are mean values from matched L3L and L3W retrievals within these grid boxes. "Matched" means that only days when both L3L and L3W are present, and the L3O surface index is mixed, are used to create the mean values analysed. Shown are AK diagonal values (left column), AK rowsums (second column), VMR retrieved minus a priori values (third column), retrieved (fourth column) and a priori (fifth column) VMRs, for the following levels of the retrieved profile: surface (top row), 900 hPa (second row), 800 hPa (third row), 600 hPa (fourth row), and 300 hPa (bottom row). Values in boxes in the top-left corner of each panel correspond to mean values across all L3L and L3W grid boxes. These means are significantly different using a 2-tailed t-test (unequal variance) with $p < 0.005$ in all cases except ak_diagonal at 300 hPa where $p = 0.13$, vmr_ret at 600hPa where $p = 0.30$, vmr_ret at 300hPa where $p = 0.11$. No vmr_apr mean differences are significant. Values in the bottom-right corner of each panel correspond to the Spearman's rank correlation coefficient ($p < 0.005$ in all cases).

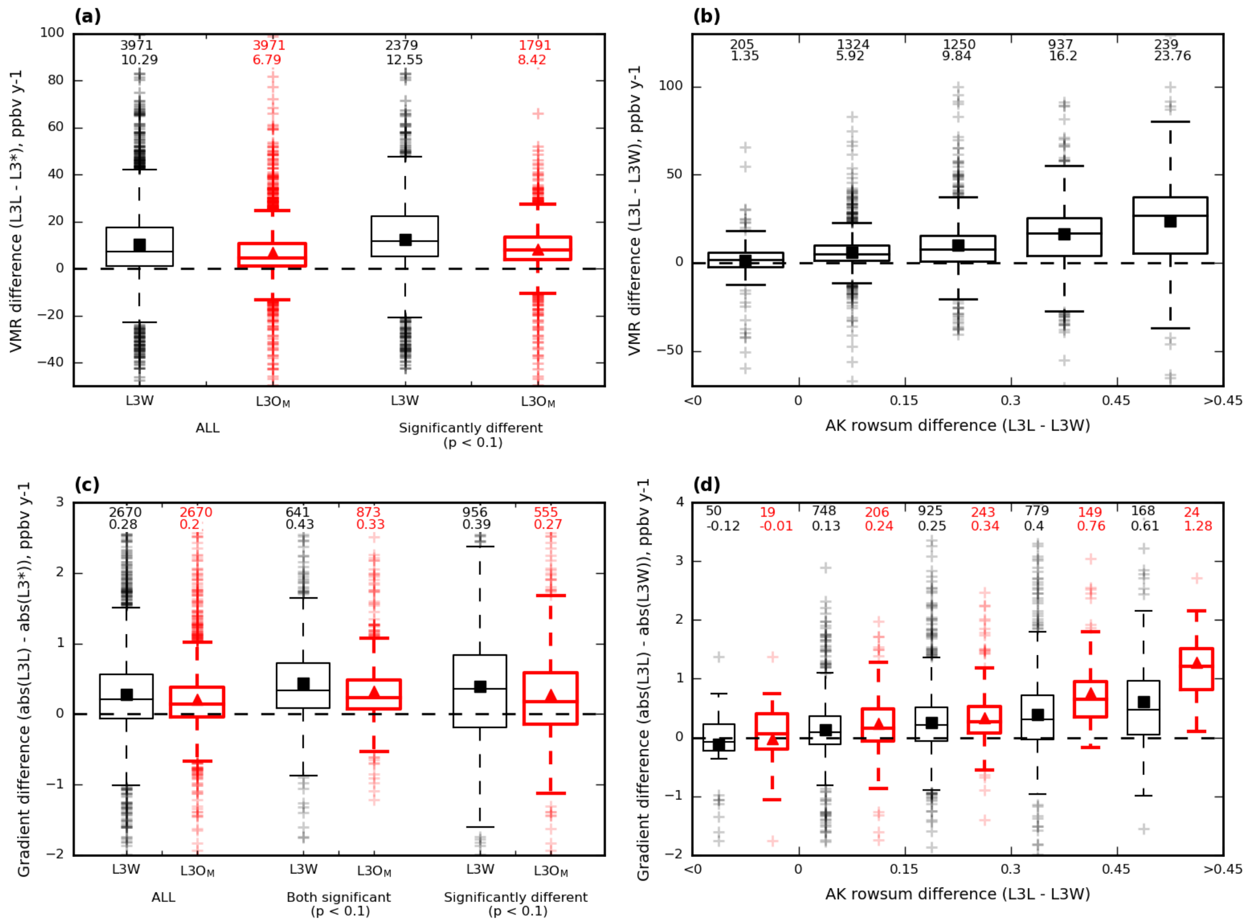


Figure S9. Boxplots showing how mean VMRs and trends from WLS analysis compare for coastal L3 grid boxes, calculated from matched retrievals within these grid boxes. "Matched" means that only days when both L3L and L3W are present and the L3O surface index are mixed are used to create the mean values analysed. Mean values are represented by filled squares/triangles, and values above the boxplots correspond to number of grid boxes with data for that boxplot, and the mean value, respectively. **(a)** Mean VMR differences for L3W (black, mean values represented by filled squares) and $L3O_M$ (red, thicker lines, mean values represented by filled triangles) compared to L3L (L3L – L3* in both cases). Shown are the differences for all coastal grid boxes, and only for those grid boxes where the difference is significant ($p < 0.1$), determined using a 2-tailed t-test. **(b)** Mean VMR differences between L3L and L3W, stratified according to corresponding AK rowsum difference (L3L – L3W in both cases). **(c)** Absolute differences in gradients detected using WLS regression analysis for L3W (black, mean values represented by filled squares) and $L3O_M$ (red, thicker lines, mean values represented by filled triangles), compared to L3L (L3L – L3* in both cases). Shown are differences for all coastal grid boxes where WLS analysis could be performed, for grid boxes where both trends compared are significantly different to zero ($p < 0.1$), and for grid boxes where the trend difference is significant ($p < 0.1$). **(d)** Absolute differences in gradients detected using WLS regression analysis between L3L and L3W, stratified according to corresponding AK rowsum difference (L3L – L3W in both cases). Shown are the differences for all coastal grid boxes where WLS could be performed (black, mean values represented by filled squares), and only for those grid boxes where the detected trend is significant ($p < 0.1$) in both L3L and L3W (red, thicker lines, mean values represented by filled triangles).

S6. Case study analysis of the potential effect of wind direction on L3L – L3W retrieved VMR differences

To evaluate any potential impact of wind direction on the results presented in the paper, wind direction (taken from ERA Interim data, u and v wind vectors at 10-m level) is compared for 6 of the case study L3 grid boxes containing large cities considered in Sect 3.4 for days when retrieved surface-level VMR in L3L > L3W and L3L < L3W. Figures S10 – S15 show wind roses representing the distribution of wind direction for these subsets of days, as well as the mean wind direction distribution across all days considered, for the 6 case studies. There are no marked differences in wind direction on these days for any of the grid boxes considered, giving confidence that wind direction does not have a large impact on the results presented, as discussed in Sect. 3.2.1.

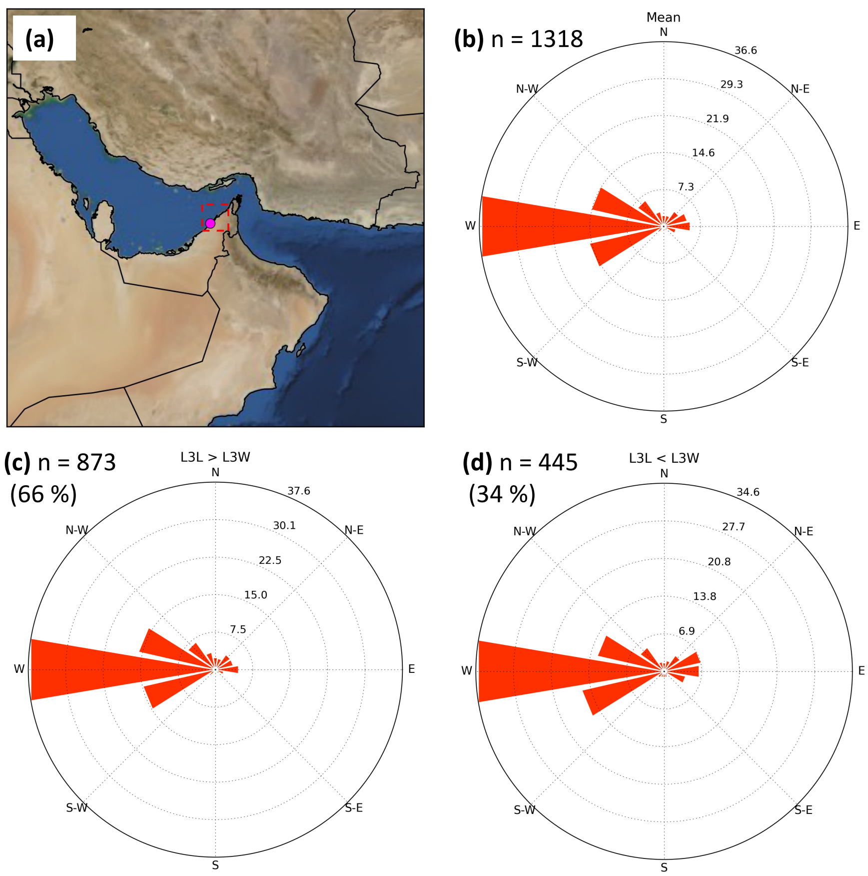


Figure S10. Surface wind direction for days when there is a surface level VMR retrieval over both land and water in the L3 grid box containing Dubai ("L3L" and "L3W", resp.) Wind direction data are calculated from u and v vector components at 10-m above the surface, taken from daily mean ERA Interim data. Wind rose barbs in panels b-d depict the direction in which wind is blowing *from* – e.g. predominantly *from the west* in all cases shown. Note that only the sub-period 01-01-2002 to 31-08-2017 is considered in this analysis (as opposed to the full study period of 2001-08-25 to 2019-02-28 in the paper) owing to local availability of ERA Interim data. **(a)** NASA Blue Marble image of the region surrounding Dubai. The boundaries of the L3 grid box containing the city are shown by red-dashed lines, with the city location indicated by the pink marker. **(b)** Wind rose showing wind direction taken from the grid box containing Dubai in ERA Interim data. Wind direction data are taken for all days with L3L and L3W data for this grid box for the study period, and the wind rose displays the mean wind direction for all these days. Number of days represented is given by n value in panel label. **(c)** As b, but only for days when retrieved surface-level VMR is greater in L3L than L3W ("L3L > L3W"). **(d)** As b, but for days when L3L < L3W.

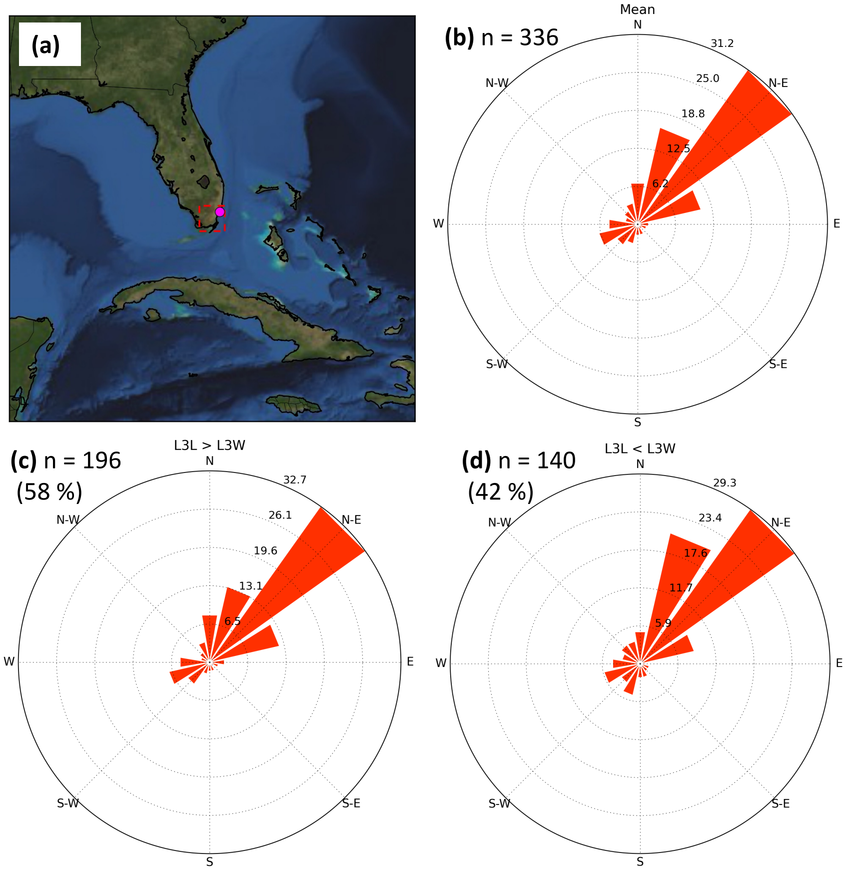


Figure S11. As Fig. S10, but for the L3 grid box containing Miami.

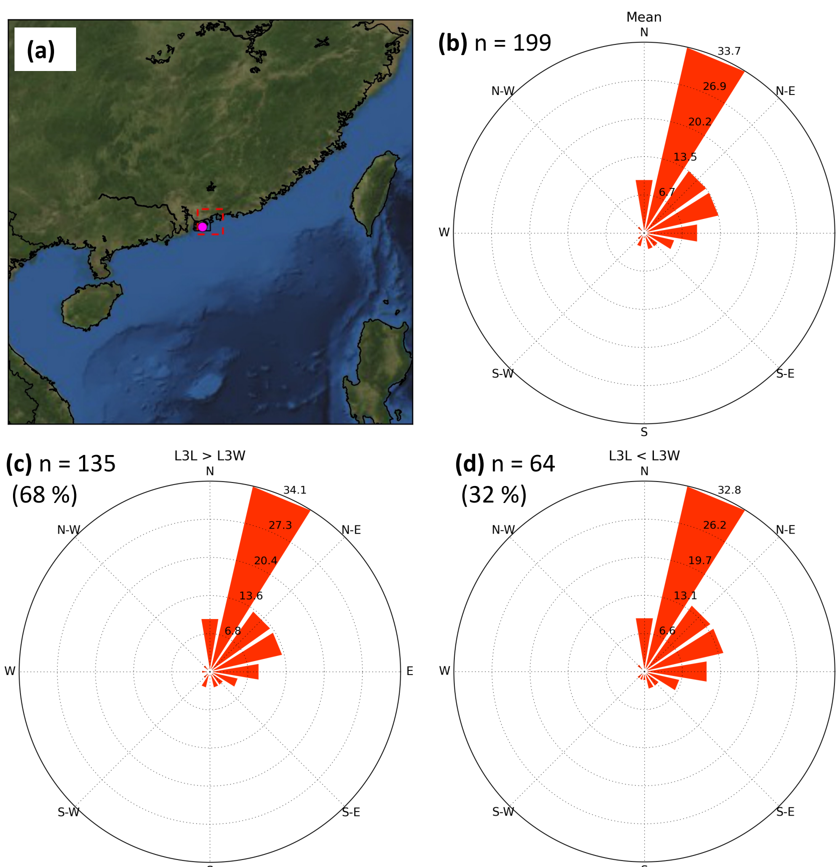


Figure S12. As Fig. S10, but for the L3 grid box containing Hong Kong.

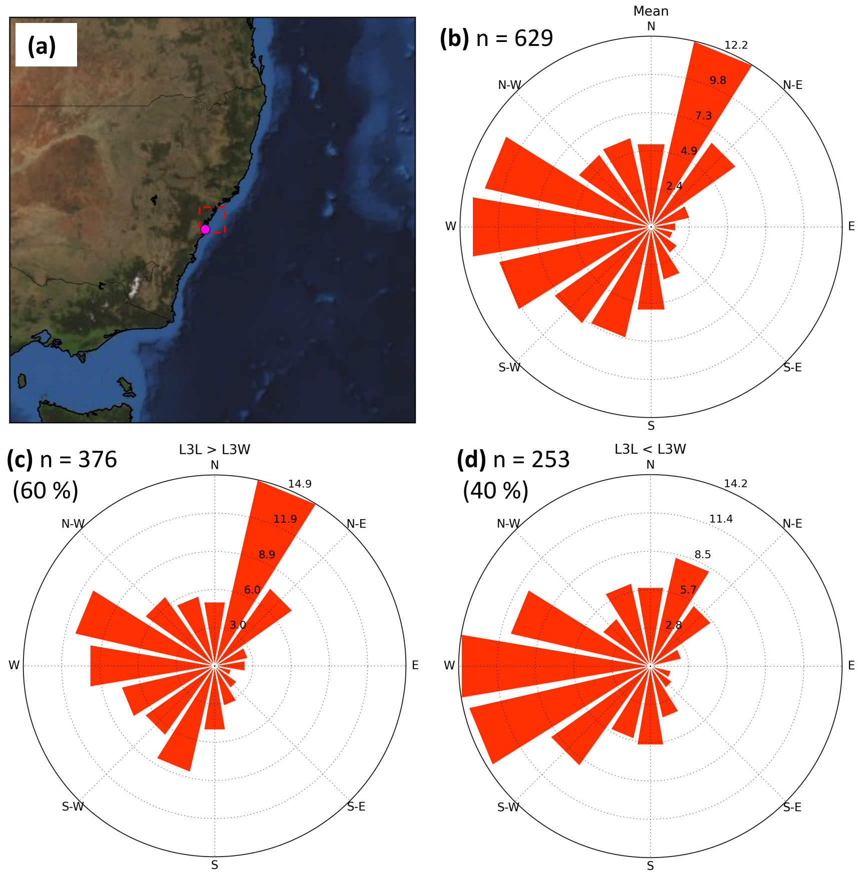


Figure S13. As Fig. S10, but for the L3 grid box containing Sydney.

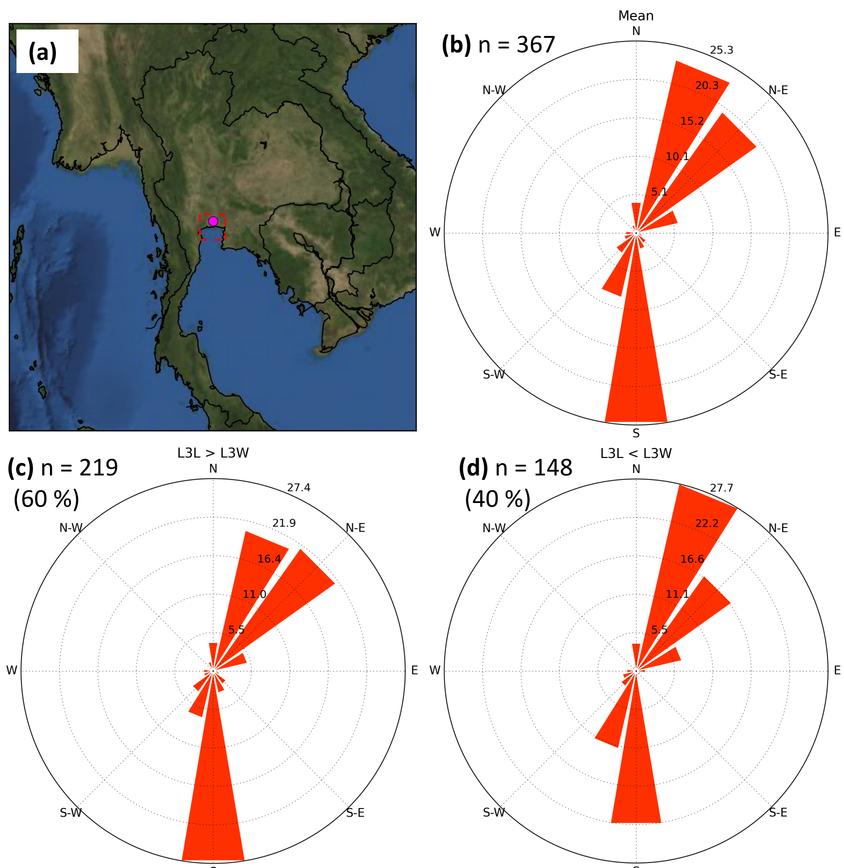


Figure S14. As Fig. S10, but for the L3 grid box containing Bangkok.

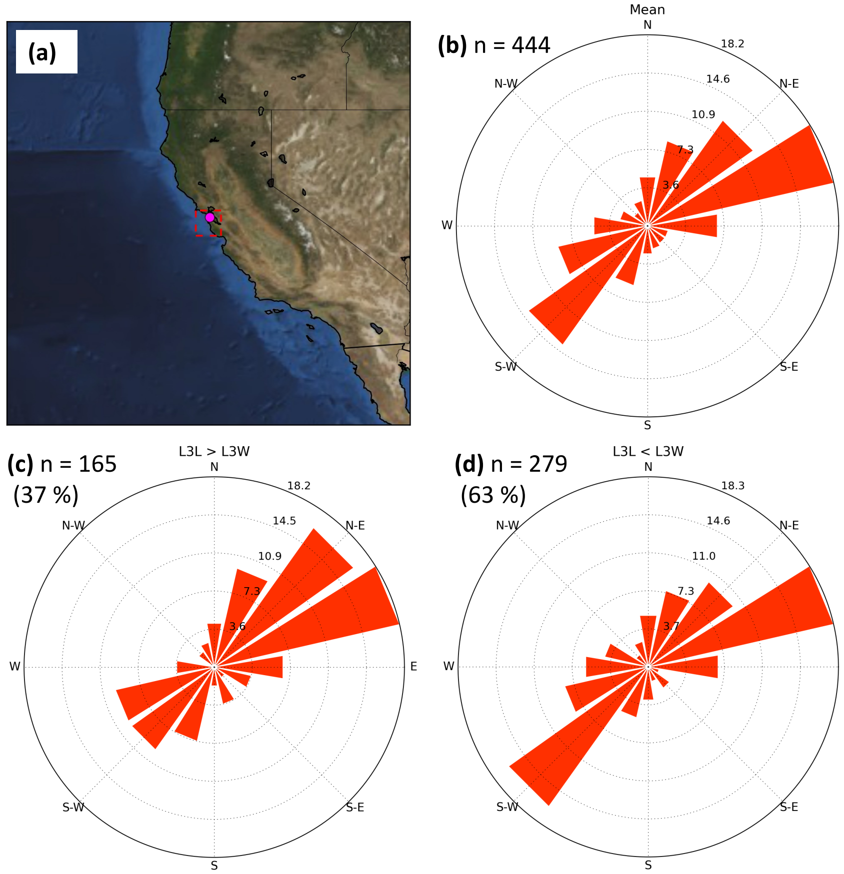


Figure S15. As Fig. S10, but for the L3 grid box containing San Francisco.

S7. L3O misclassification examples

Here is provided case study evidence of L3O retrievals incorrectly (as far as the author's understanding goes) being given the surface index of mixed for specified coastal grid boxes. To recap (text from Sect. 2.2): For a given 1° x 1° L3 grid box, how the L2 retrievals that fall within its boundaries are processed to produce the L3 product depends on how their surface indexes vary: If more than 75 % of the bounded L2 retrievals have the same surface index, only those retrievals are averaged to produce the L3 gridded value, and the L3 surface index is set to that surface type (the other L2 retrievals are discarded). Otherwise, all L2 retrievals available in the L3 gridbox are averaged together and the L3 surface index is set to "mixed".

Table S2 presents data extracted from the original, as-downloaded MOPITT V8 L3 TIR-NIR combined file ('MOP03J') for the L3 grid box containing the cities of San Francisco (a; longitude = -122.447° E, latitude = 37.734° N) and Istanbul (b; longitude = 28.980° E, latitude =41.015° N) for selected days, as indicated by the date column. It shows the surface index ascribed to the retrieval for that grid box and day, and the number of L2 retrievals that are averaged together to create the L3 retrieval. Also presented in Table S2 is a breakdown of the surface indexes of all L2 retrievals which fall within the respective L3 grid box on the specified day that are used to create the L3 retrieval. Note that these retrievals are first screened for data quality, following the criteria outlined in Sect. 2.4 and specified in the data user's guide (MOPITT Algorithm Development Team, 2018). For all cases presented, the L3O surface index is "mixed", yet the *only* L2 retrievals that contribute to the L3 retrieval are retrievals with a surface index of "water" (as verified by n_ret(L3O) equalling n_ret($L2_W$) in all cases). It has also been confirmed for all cases shown that the retrieved surface level VMR reported in L3O is created only from averaging the bounded L2 retrievals over water.

A note on case study data selection:

These case studies were chosen for further analysis because the grid boxes were already being analysed in Sect. 3.4 ("Illustrative examples comparing L3O and L3L: analysis of the most populous coastal cities"). The total number of cases of apparent incorrect surface classification in the L3O, as-downloaded data, has not been quantified, as it is beyond the scope of this paper. However, the relative ease with which case studies demonstrating this issue was found suggests that this effect could be large, going some way to explaining how the ratio n_days($L3L/L3O_{LM}$) can be less than 1 for certain grid boxes (i.e. n_obs($L3O_{LM}$) > n_obs(L3L)), contrary to expectations based on understanding of how the L3O data are created, as discussed in Sect. 3.3.1 ("Loss of available data").

Tentative explanation for apparent misclassification

It is plausible that the L3O surface index could be determined based on the surface indexes of bounded L2 retrievals *before* they are screened for data quality following the criteria outlined in Sect. 2.4 and specified in the data user's guide (MOPITT Algorithm Development Team, 2018). For each case study presented in Table S2, no one surface type accounts for more than 75 % of bounded L2 retrievals, if all L2 retrievals are counted prior to screening. This is visualized in Fig. S16, the San Francisco case study of 20051121. In this example, a total of 13 L2 retrievals (8 (5) with a surface index of water (land)) are bounded by the L3 grid box containing San Francisco, but only 5, all with a surface index of water, are used to create the L3O product – yet its surface index is "mixed". However, as noted above, it has been confirmed for all cases shown that the retrieved surface level VMR in L3O is created only from averaging the bounded L2 retrievals over water.

Table S2. L3 and L2 surface classification details for selected L3 surface misclassification case studies for the coastal L3 grid box containing the cities of **a)** San Francisco (-122.447° E, 37.734° N) and **b)** Istanbul (28.980° E, 41.015° N). "L3O sfci" = surface index of the L3 grid box for day given by "date" (note that the number "2" is the code for surface index "mixed" in the MOP03J data files); n_ret(L3O) = number of L2 retrievals averaged together to create the L3 product; n_ret(L2$_\mathrm{W}$) = number of L2 retrievals with a surface index of "water"; n_ret(L2$_\mathrm{L}$) = number of L2 retrievals with a surface index of "land"; n_ret(L2$_\mathrm{M}$) = number of L2 retrievals with a surface index of "mixed". Subscripted, italicised numbers in square brackets correspond to n_ret(L2$_\mathrm{W/L/M}$) *before* the bounded L2 retrievals are screened for data quality during L3 product creation.

	Date	**L3O sfci**	**n_ret(L3O)**	**n_ret(L2$_\mathrm{W}$)**	**n_ret(L2$_\mathrm{L}$)**	**n_ret(L2$_\mathrm{M}$)**
a) *L3 grid box containing San Francisco*	20050925	2 ("mixed")	4	4 *[11]*	0 *[6]*	0 *[1]*
	20051025	2 ("mixed")	2	2 *[4]*	0 *[3]*	0 *[0]*
	20051114	2 ("mixed")	3	3 *[4]*	0 *[5]*	0 *[0]*
	20051121	2 ("mixed")	5	5 *[8]*	0 *[5]*	0 *[0]*
b) *L3 grid box containing Istanbul*	20020613	2 ("mixed")	6	6 *[17]*	0 *[12]*	0 *[1]*
	20020622	2 ("mixed")	5	5 *[18]*	0 *[14]*	0 *[3]*
	20020715	2 ("mixed")	7	7 *[18]*	0 *[11]*	0 *[2]*
	20020717	2 ("mixed")	2	2 *[17]*	0 *[11]*	0 *[0]*
	20020811	2 ("mixed")	4	4 *[21]*	0 *[12]*	0 *[0]*

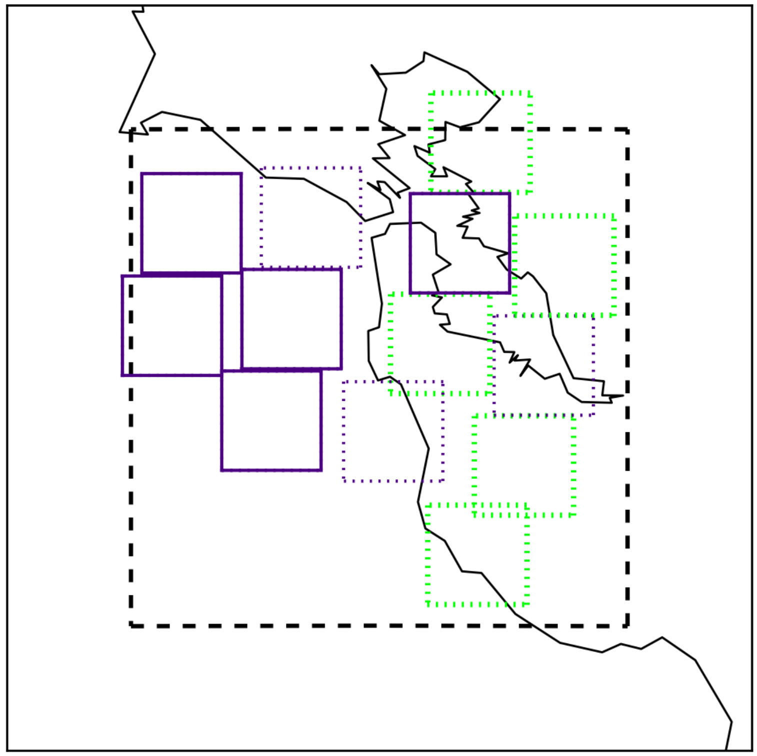


Figure S16. Coastal 1º x 1º L3 grid box containing the city of San Francisco (black dashed box) and bounded L2 retrievals for the case study day 20051121 where the L3 surface index is “mixed”. Purple (green) boxes correspond to L2 retrievals with a surface index of “water” (“land”). Where the purple/green box is solid, the L2 retrieval is used to create the L3O product; dashed purple/green boxes indicate an L2 retrieval that was discarded before L3O product creation following the data filtering criteria outlined in Sect. 2.4. More information on surface indexing and L3 product creation is given in Sect. 2.2. “Coastal” L3 grid box classification is outlined in Sect. 2.3.

S8. Comparison of trends detected for most populous coastal cities using alternative regression methods

The analysis presented in Sect. 3.4 is based on trends detected using Weighted Least Squares (WLS) linear regression (method outlined in detail in Sect. 2.5). Here, it is demonstrated that the main results presented in the paper hold if the trend analysis is performed using an alternative regression method, namely the Theil Sen slope estimator. Theil Sen analysis is a non-parametric trend estimation method whereby slopes between all pairs of points are calculated and the median of the slopes is taken. This is less sensitive to the impact of outliers compared to least-squares regression methods. Trends from both methods, and a comparison of the conclusions drawn using them, are given in Table S3.

Firstly, results for the L3 grid box containing Dubai (discussed in Sect. 3.4.1) are compared. The trends from both methods are the same for L3L, $L3O_{LM}$, and L3W. The trend for $L3O_{L}$ is slightly weaker in Theil Sen than in WLS, however it is still 1.7 ppbv stronger than the trend in L3L and, based on the differences in temporal coverage between the two datasets, the conclusion that this is a clear over-estimate stands.

Considering all 33 cities compared, on average trends are strongest in $L3O_{L}$ and weakest in L3W using both regression methods, but the trends in L3L – $L3O_{LM}$ are a closer match in Theil Sen than WLS. Across all datasets, the trend detected using Theil Sen is significant ($p < 0.1$) for a couple less cities than using WLS.

Comparing L3L and $L3O_{L}$: As with the Dubai case study, detected trends are slightly weaker in Theil Sen than WLS across all cities analysed, nonetheless on average they still represent an over-estimate of the equivalent trend in L3L, of which $L3O_{L}$ is a sub-sample (as discussed in e.g. Sect.3.3.2). The L3L – $L3O_{L}$ trend difference is significant ($p < 0.1$) in 5 less cases using Theil Sen than WLS. The 4 cities that do display a significant L3L – $L3O_{L}$ trend difference using Theil Sen do so as well using WLS (Qingdao, Saigon, Buenos Aires, Bangkok; these cities are discussed in the paper).

Comparing L3L and $L3O_{LM}$: As with WLS, the negligible mean difference in trends from Theil Sen hides large differences for some cities. In fact, two additional cities (Kuala Lumpur and Accra) see significant differences between trends detected in L3L and $L3O_{LM}$ using Theil Sen than WLS. The five cities that do display a significant L3L – $L3O_{LM}$ trend difference using WLS do so as well using Theil Sen (New York, Istanbul, Saigon, Hong Kong, Dubai; these cities are discussed in the paper).

To summarise: the main conclusion that the results of trend analysis for large cities situated in coastal L3 grid boxes differ depending on whether L3O or L3L is analysed holds, irrespective of the trend detection method used. There are small differences in the details, and the analysis results do not differ significantly for some cities, but there remain a number whereby the differences in results obtained using L3O and L3L are large.

Table S3. Results of trend analysis for the L3 grid boxes containing the 33 cities of interest from each of the L3O subsets considered, L3L, and L3W (for comparison). Trends are from Weighted Least Squares ("WLS"; these are as presented in Table 6 of the paper) and Theil Sen methods. Trends units are ppbv y^{-1}. Dash symbols ('-') indicate that the trend cannot be calculated for a given grid box and dataset owing to lack of data. Yellow shading indicates that a trend value is significantly different to the corresponding value in L3L for that city ($p < 0.1$). Grey shading indicates that the trend value is not significantly different to zero ($p < 0.1$). Diagonally striped yellow-grey shading indicates that the trend value is not significantly different to zero AND that it is significantly different to the corresponding trend in L3L for that city. The right hand columns indicate whether the conclusion drawn from the analysis is different depending on whether the trend is estimated using the WLS or Theil Sen methods.

[1] The modified mean corresponds to the mean value that is calculated only for cities where is a corresponding trend in the L3O$_L$ dataset (n = 18). By contrast, the mean value simply represents the mean of all values in that column.

city	WLS trend (± standard error) [ppbv y-1]				Theil Sen trend (± standard error) [ppbv y-1]				WLS & Theil Sen result different?		
	L3L	L3O$_L$	L3O$_{LM}$	L3W	L3L	L3O$_L$	L3O$_{LM}$	L3W	L3L v L3O$_L$	L3L v L3O$_{LM}$	L3L v L3W
Tokyo	-1.7 (0.3)	-2.3 (0.5)	-1.7 (0.3)	-1.7 (0.3)	-1.8 (0.5)	-2.6 (0.8)	-1.7 (0.5)	-1.5 (0.3)	x	x	x
Shanghai	-5.9 (1.4)	-7.0 (1.6)	-5.7 (1.4)	-3.4 (1.2)	-6.0 (1.1)	-7.5 (3.3)	-5.5 (1.5)	-3.3 (1.4)	x	x	x
Manila	-1.3 (0.5)	-	-1.2 (0.4)	-1.3 (0.2)	-1.6 (0.7)	-	-1.2 (0.5)	-1.3 (0.2)	-	x	x
Mumbai	-1.2 (0.9)	-	-0.6 (0.6)	-0.1 (0.3)	-1.1 (1.6)	-	-0.8 (0.6)	-0.1 (0.4)	-	x	x
New York	-1.4 (1.1)	-	-2.3 (0.8)	-1.9 (0.5)	-0.9 (1.7)	-	-1.8 (1.1)	-1.8 (0.7)	-	x	x
Lagos	1.2 (2.0)	-	0.5 (1.7)	0.2 (0.4)	1.3 (2.5)	-	1.2 (2.2)	0.1 (0.6)	-	x	x
Bangkok	-3.0 (0.6)	-8.6 (2.1)	-3.1 (0.7)	-2.0 (0.4)	-3.0 (0.8)	0.8 (3.6)	-3.4 (0.9)	-2.0 (0.5)	x	x	x
Osaka	-2.5 (0.5)	-2.3 (0.5)	-2.3 (0.4)	-1.3 (0.4)	-2.6 (0.7)	-2.3 (0.6)	-2.5 (0.4)	-1.7 (0.5)	x	x	Yes
Karachi	-0.8 (0.2)	-0.6 (0.2)	-0.7 (0.2)	-0.5 (0.3)	-0.8 (0.3)	-0.5 (0.3)	-0.7 (0.2)	-0.5 (0.3)	x	x	x
Buenos Aires	-0.1 (0.1)	-0.5 (0.2)	-0.2 (0.1)	-0.1 (0.1)	-0.2 (0.2)	-0.5 (0.2)	-0.3 (0.1)	-0.1 (0.1)	x	x	x
Istanbul	-1.2 (0.4)	-	-0.4 (0.2)	-0.8 (0.2)	-1.3 (0.4)	-	-0.4 (0.3)	-0.8 (0.3)	-	x	x
Chennai	0.0 (0.8)	-	0.5 (0.5)	-0.9 (0.3)	-0.1 (0.9)	-	-0.0 (0.6)	-0.7 (0.2)	-	x	x
Xiamen	-2.6 (0.9)	-	-4.1 (1.7)	-1.9 (0.4)	-2.9 (0.8)	-	-4.8 (2.1)	-2.1 (0.4)	-	x	x
Taipei	-3.7 (1.0)	-	-	-1.5 (0.4)	-2.2 (2.2)	-	-	-1.5 (0.3)	-	-	x
Kuala Lumpur	-2.7 (1.3)	-3.4 (1.2)	-3.9 (1.0)	-5.1 (1.1)	-2.3 (2.5)	-2.0 (3.3)	-3.8 (1.9)	-4.5 (1.7)	x	Yes	Yes
Saigon	-1.4 (0.9)	-3.6 (1.3)	-2.3 (0.8)	-2.3 (0.8)	-1.0 (1.1)	-2.6 (1.3)	-2.2 (0.9)	-2.4 (0.9)	x	x	x
Luanda	-0.5 (2.1)	-2.6 (3.7)	0.5 (2.2)	-0.2 (1.0)	0.1 (2.3)	-0.1 (3.4)	0.1 (2.5)	-0.2 (1.2)	x	x	x
San Francisco	-1.1 (0.7)	-	-0.7 (0.5)	-1.0 (0.6)	-0.7 (1.1)	-	-0.7 (0.7)	-1.0 (0.8)	-	x	Yes
Singapore	-	-	-	-4.3 (2.4)	-	-	-	-4.3 (3.3)	-	-	-
Shantou	-5.4 (0.5)	-5.9 (1.4)	-5.7 (0.4)	-3.8 (0.7)	-5.6 (0.7)	-6.1 (2.2)	-5.8 (0.5)	-4.1 (0.8)	x	x	Yes
Hong Kong	-8.1 (0.9)	-	-5.1 (1.3)	-3.5 (0.5)	-8.2 (0.9)	-	-5.4 (1.4)	-3.4 (0.6)	-	x	x
Toronto	-1.1 (0.8)	-0.3 (1.1)	-1.2 (0.7)	-2.0 (0.6)	-0.8 (0.9)	-0.0 (1.8)	-0.6 (0.9)	-2.2 (0.7)	x	x	x
Miami	-1.5 (0.4)	-1.2 (1.2)	-1.3 (0.3)	-0.8 (0.2)	-1.4 (0.3)	-2.0 (1.1)	-1.2 (0.2)	-0.7 (0.2)	Yes	x	x
Surat	-0.4 (0.3)	-1.6 (0.7)	-0.4 (0.3)	-0.1 (0.3)	-0.3 (0.4)	-1.3 (0.8)	-0.3 (0.3)	-0.0 (0.4)	Yes	x	x
Dar Es Salaam	-0.3 (0.7)	-	-	-0.2 (0.1)	-0.2 (1.3)	-	-	-0.2 (0.1)	-	-	x
Qingdao	-3.8 (1.5)	-2.0 (1.7)	-3.7 (1.4)	-4.2 (0.9)	-4.3 (1.5)	-3.8 (2.1)	-4.6 (1.2)	-4.3 (1.2)	x	x	x
Yangon	-1.5 (0.8)	-	-1.7 (0.6)	-2.1 (0.5)	-1.9 (0.8)	-	-1.7 (0.7)	-2.1 (0.6)	-	x	x
Abidjan	-2.1 (1.4)	-3.6 (1.8)	-0.3 (1.4)	-0.7 (0.3)	-1.4 (2.0)	-3.4 (3.2)	-1.4 (1.7)	-0.9 (0.4)	Yes	x	x
Wenzhou	-4.2 (0.7)	-10.1 (2.4)	-3.5 (0.6)	-2.8 (0.6)	-4.4 (0.9)	-6.8 (4.4)	-3.4 (0.7)	-2.7 (0.6)	Yes	x	x
Sydney	-0.7 (0.2)	-	-0.5 (0.1)	-0.2 (0.1)	-0.8 (0.2)	-	-0.5 (0.1)	-0.3 (0.1)	-	x	x
Accra	2.9 (1.5)	-	2.9 (0.8)	-0.5 (0.3)	2.3 (1.2)	-	2.6 (0.9)	-0.5 (0.3)	-	Yes	x
Dubai	-2.9 (0.3)	-5.0 (1.2)	-1.6 (0.2)	-0.9 (0.1)	-2.9 (0.3)	-4.6 (1.6)	-1.6 (0.3)	-0.9 (0.2)	Yes	x	x
Chittagong	-0.7 (0.5)	-2.1 (2.2)	-0.7 (0.5)	-0.9 (0.7)	-0.7 (0.6)	-1.4 (3.7)	-0.7 (0.5)	-1.0 (0.7)	x	x	x
Mean	-1.9	-3.5	-1.7	-1.6	-1.8	-2.6	-1.8	-1.6	-	-	-
Modified Mean[1]	-2.3	-3.5	-2.1	-1.8	-2.2	-2.6	-2.2	-1.8	-	-	-
n. significant	19	13	19	27	16	10	19	25	-	-	-
n. sigdiff to L3L	-	9	5	15	-	4	7	13	-	-	-