Supplement to: *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)

SM1. Alternative version of Fig. 3

Figure S1 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.

SM2. Alternative version of Fig. 4

Figure S2 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.

SM3. Alternative version of Fig. 5

Figure S3 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.

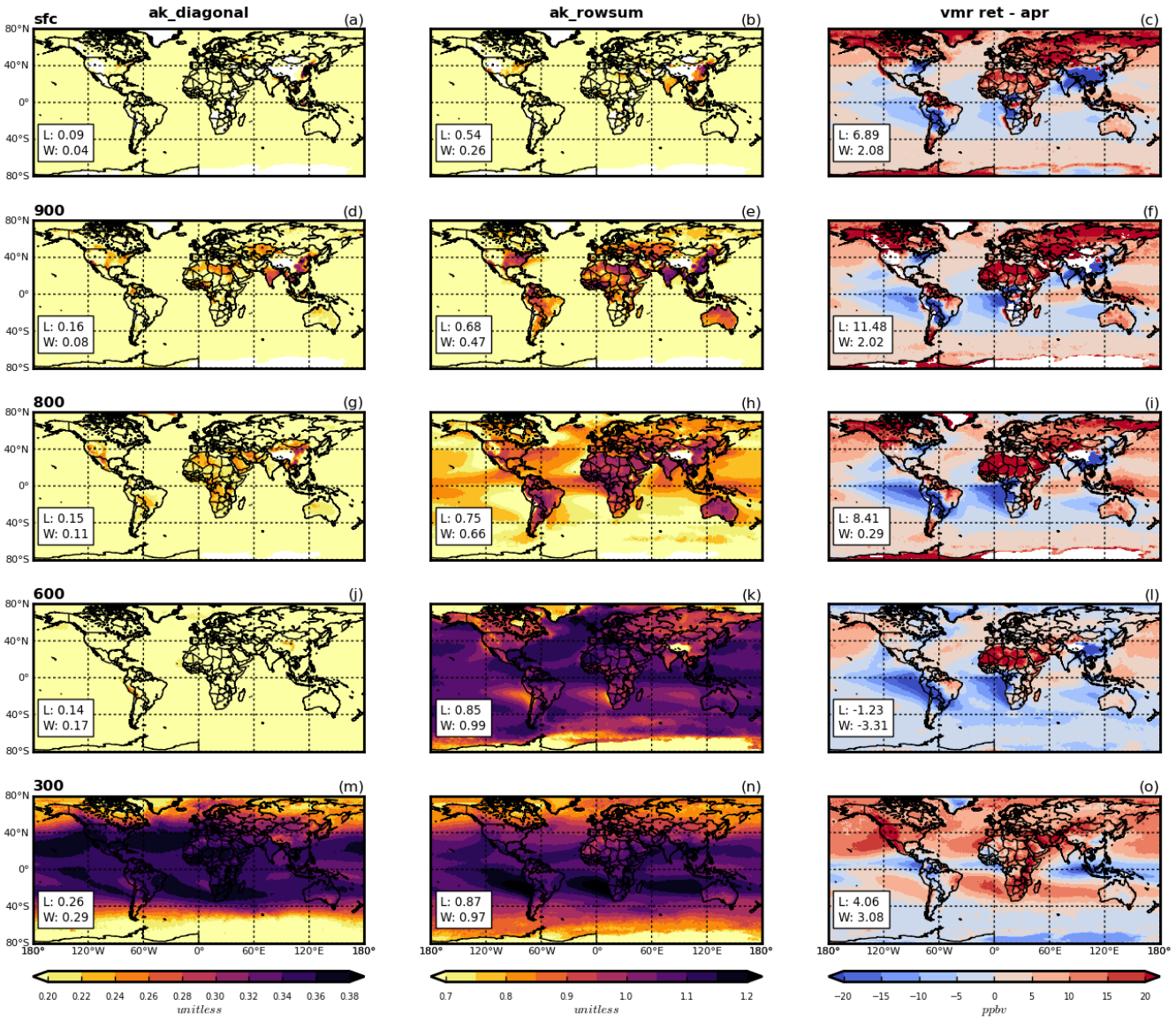

Figure S1. Mean sensitivity metrics from MOPITT L3 data, averaged across the entire study period (September 2001 – February 2019, inclusive). Shown are AK diagonal values (left column), AK rowsums (center column) and VMR retrieved minus a priori values (right column) for the following levels of the retrieved 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.

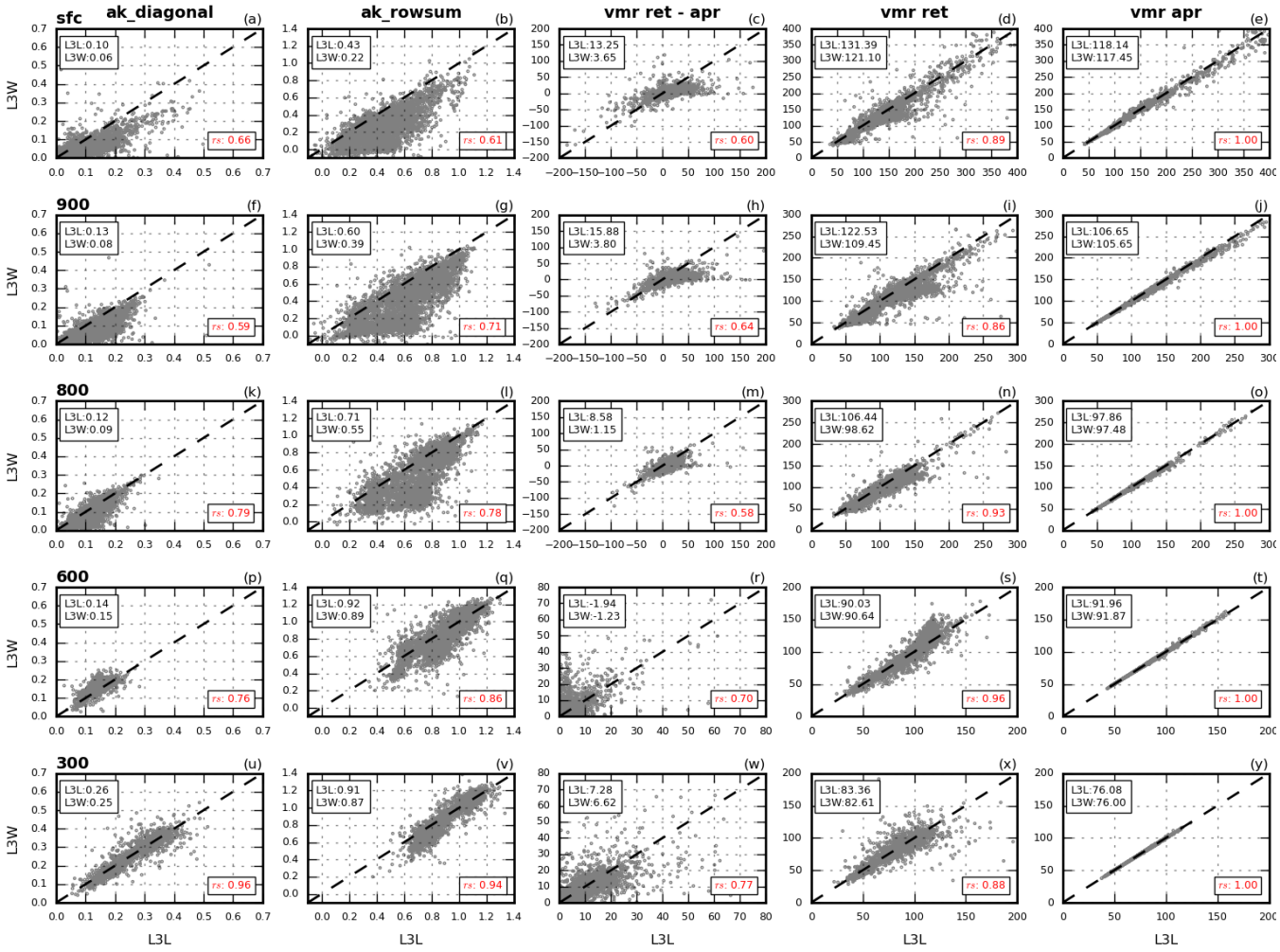

Figure S2. Mean sensitivity metrics and VMRs (retrieved and a priori) from coastal L3 grid boxes. Values compared in the scatterplots are mean values from matched L3L and L3W retrievals within these grid boxes. "Matched" means that only days when both L3L and L3W are present, and the L3O surface index 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).

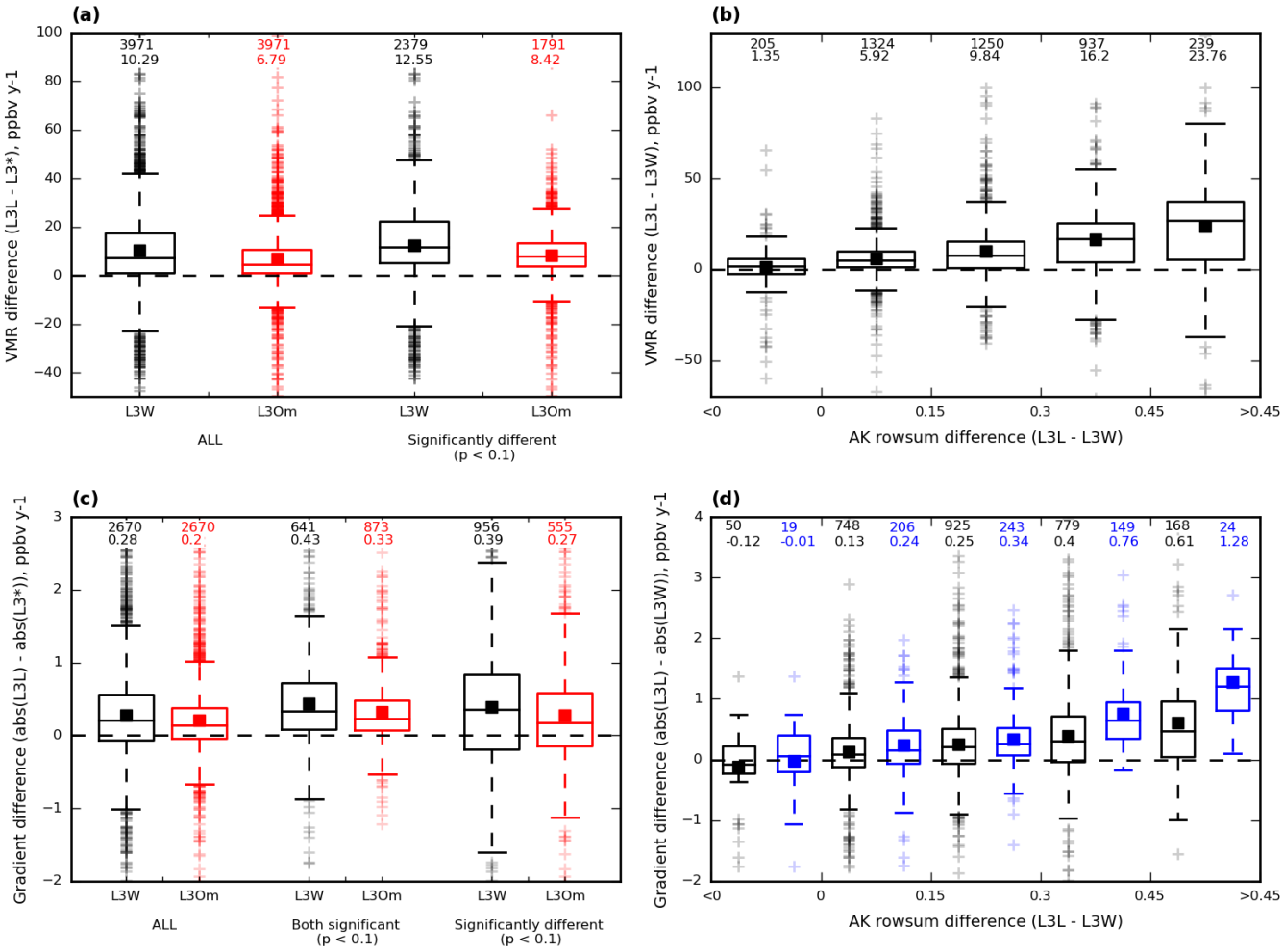

Figure S3. Boxplots showing how mean VMRs and trends from WLS analysis compare for coastal L3 grid boxes, calculated from matched retrievals within these grid boxes. "Matched" means that only days when both L3L and L3W are present and the L3O surface index are mixed are used to create the mean values analysed. Mean values are represented by filled squares, 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) and $L3O_M$ (red) 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) and $L3O_M$ (red), 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), and only for those grid boxes where the detected trend is significant ($p < 0.1$) in both L3L and L3W (blue).

SM4. 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 S1 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 S1 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. 3.1.1 ("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 S1, 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. S4, 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 S1. 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($L2_W$) = number of L2 retrievals with a surface index of "water"; n_ret($L2_L$) = number of L2 retrievals with a surface index of "land"; n_ret($L2_M$) = number of L2 retrievals with a surface index of "mixed". Subscripted, italicised numbers in square brackets correspond to n_ret($L2_{W/L/M}$) *before* the bounded L2 retrievals are screened for data quality during L3 product creation.

	Date	L3O sfci	n_ret(L3O)	n_ret($L2_W$)	n_ret($L2_L$)	n_ret($L2_M$)
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]*

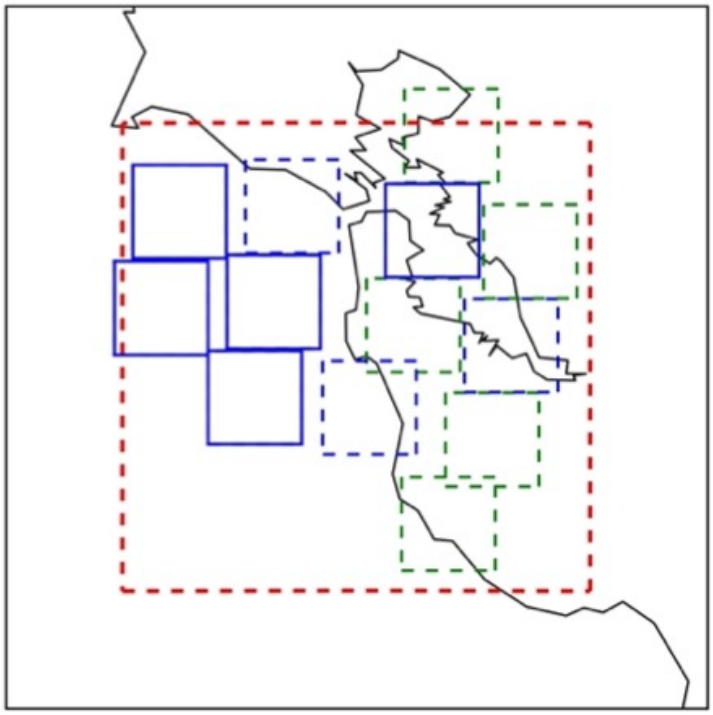

Figure S4. Coastal 1° x 1° L3 grid box containing the city of San Francisco (red dashed box) and bounded L2 retrievals for the case study day 20051121 where the L3 surface index is "mixed". Blue (green) boxes correspond to L2 retrievals with a surface index of "water" ("land"). Where the blue/green box is solid, the L2 retrieval is used to create the L3O product; dashed blue/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.