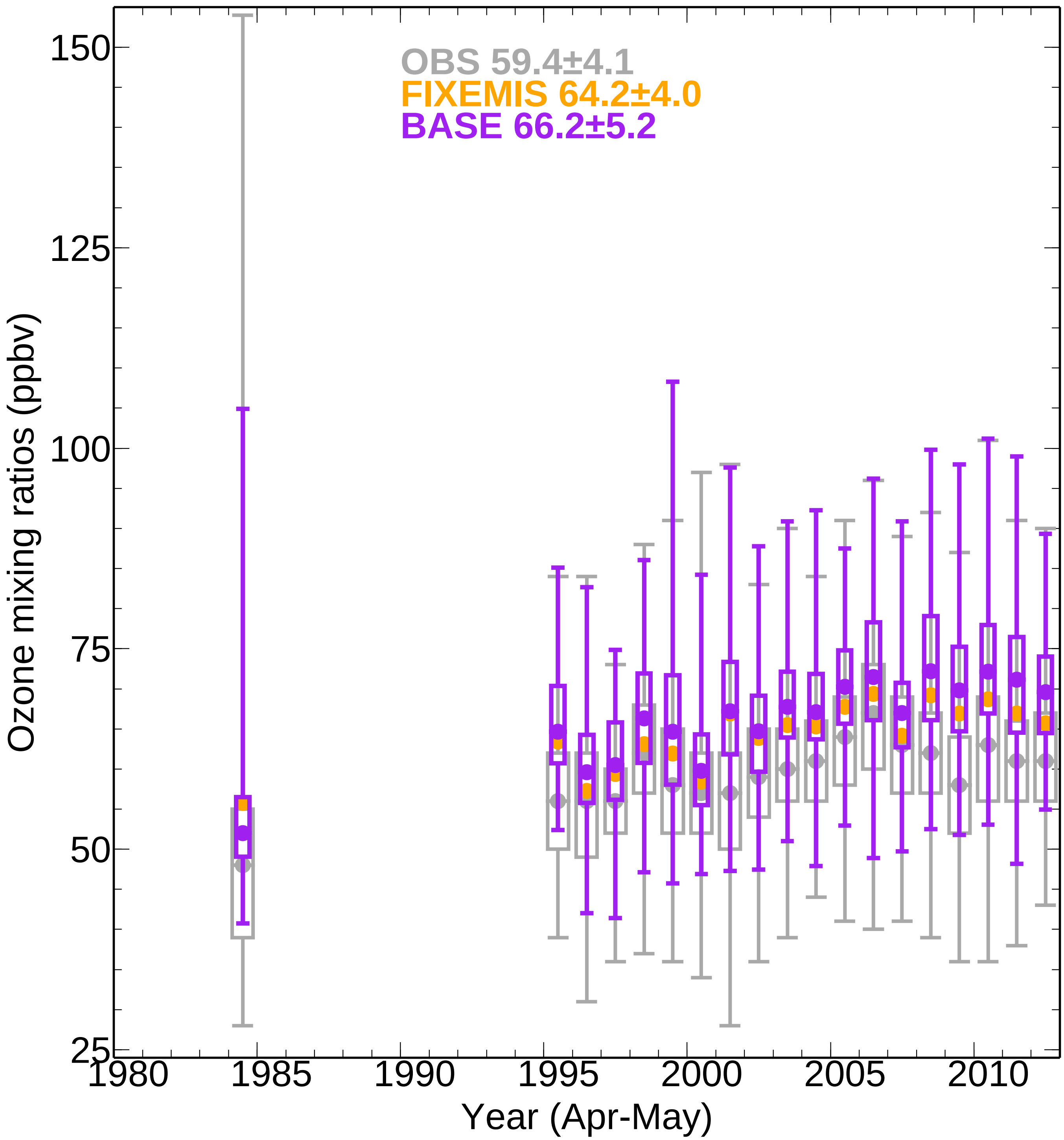
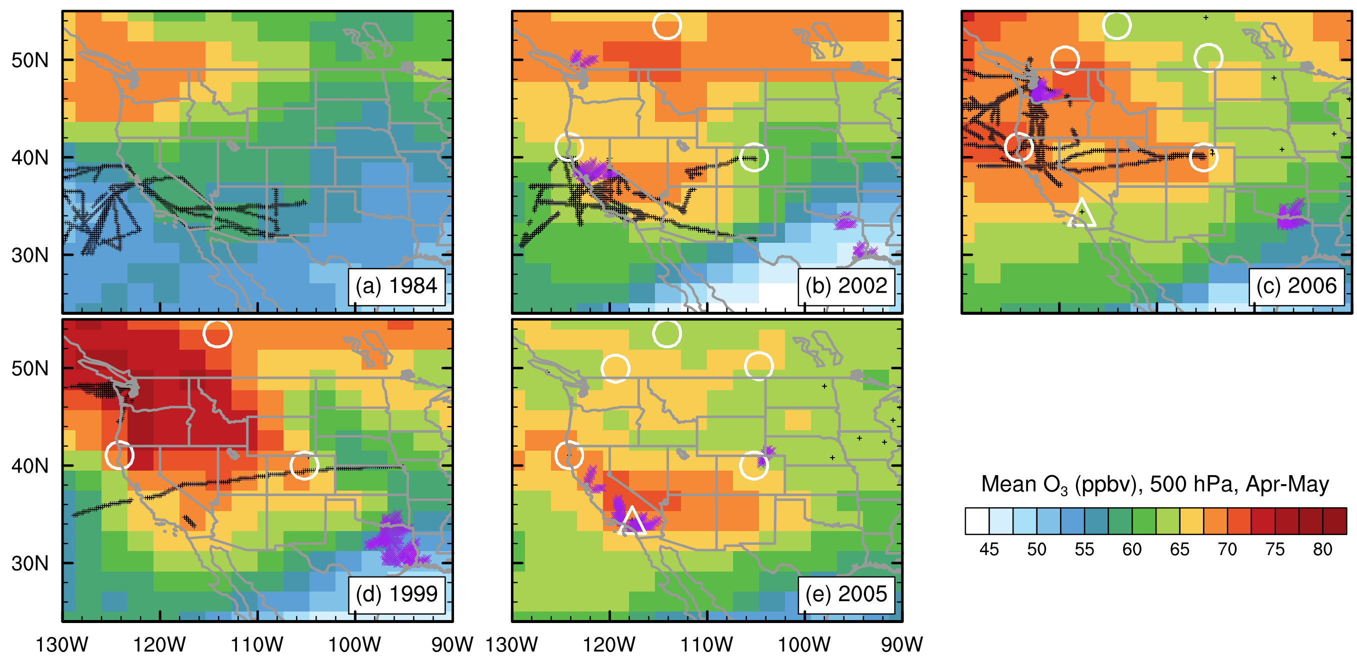
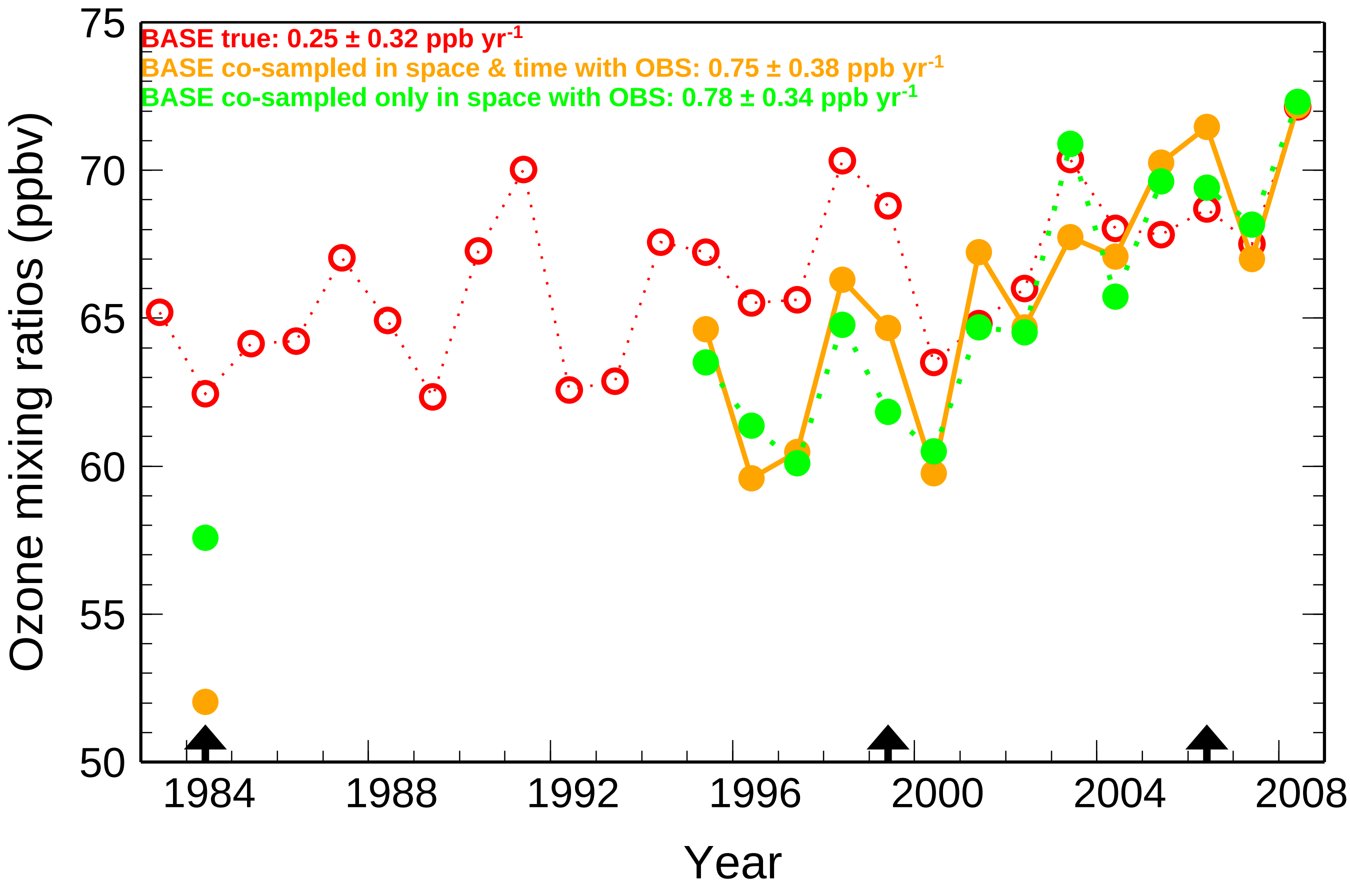


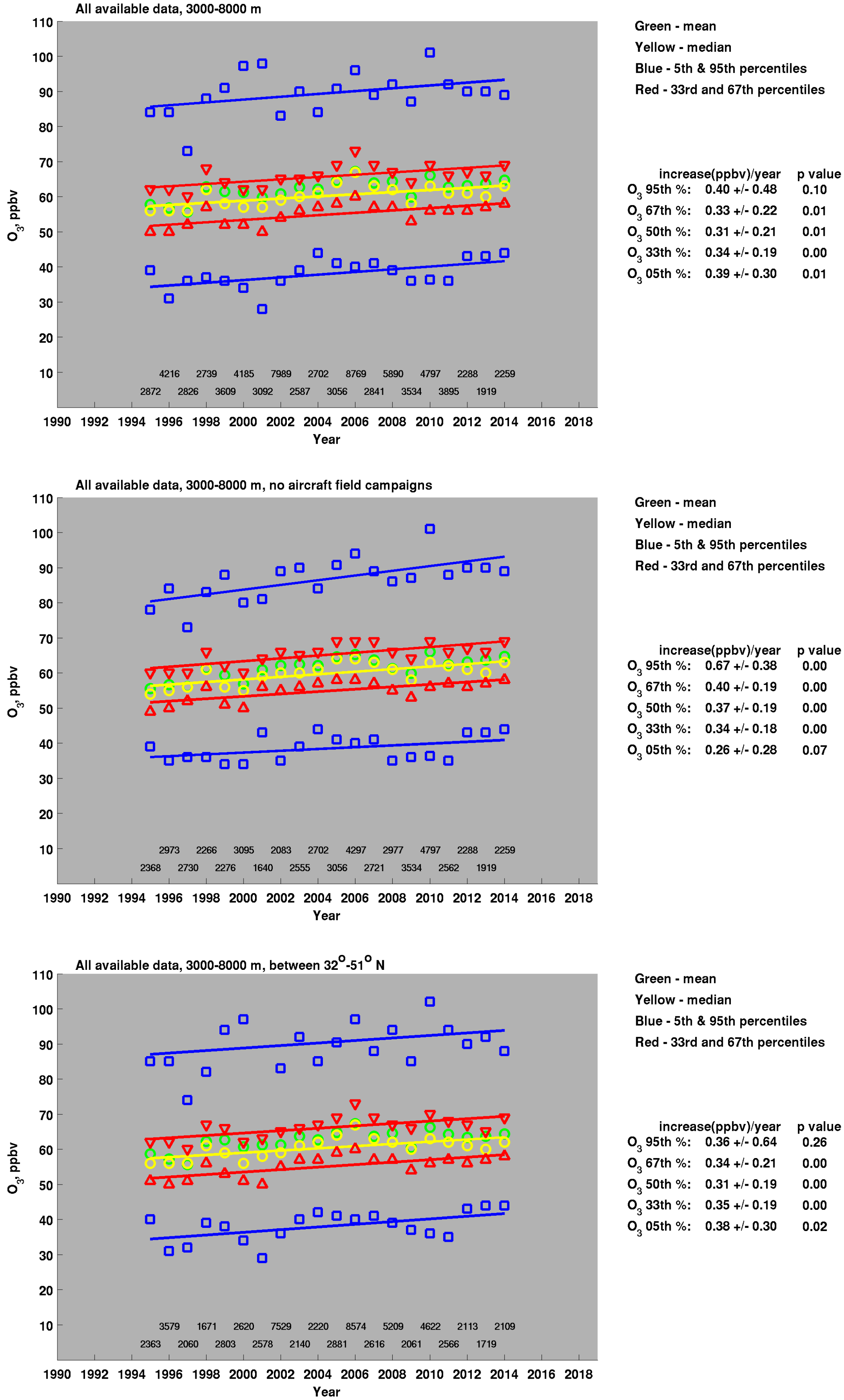
Supplementary Figure S1. Distributions of ozone samples for each April-May from the composite of observations (gray) and AM3 simulations with constant (FIXEMIS; orange) and time-varying (BASE; purple) anthropogenic emissions. The model results co-sampled in space and time with available observations are shown. The box-and-whisker plots represent the 5th, 33th, 50th, 67th, and 95th percentiles. The climatological mean and the year-to-year standard deviation are shown at the top of the graph.



Supplementary Figure S2. Same as Figure 2 in the main article, but showing background ozone simulated with anthropogenic emissions set to zero over North America but varying elsewhere as in BASE.



Supplementary Figure S3. April-May averages of ozone 3-8 km above western North Amerca over 1984-2008: from AM3_BASE true median (red); from AM3_BASE co-sampled both in space and time with observations (orange); and from AM3_BASE co-sampled only in space with observations with monthly ozone fields archived from the model for each April-May (green). The 1995-2008 trends are reported.



Supplementary Figure S4. Observed April-May free tropospheric ozone trends during 1995-2014 over western North America: **(Top)** from all available data, **(Middle)** when research aircraft campaign data are removed, and **(Bottom)** when the latitudinal range of the data set is limited to 32N-51N. Numbers at the bottom of the graph indicate the sample size for each year.

Supplementary Information

Meiyun Lin, L.W. Horowitz, O.R. Cooper, D. Tarasick, S. Conley, L.T. Iraci, B. Johnson, T. Leblanc, I. Petropavlovskikh, and E.L. Yates: *Revisiting the evidence of increasing springtime ozone mixing ratios in the free troposphere over western North America*, Geophysical Research Letter, 2015GL065311.

Supplementary Methods

1. Significance testing for trends with overlapping confidence intervals

If two statistics have overlapping confidence intervals, it is not necessarily true that they are not significantly different. We conduct statistical testing following the instruction in https://www.cscu.cornell.edu/news/statnews/stnews73.pdf. According to the tests, the model co-sampled and 'true' median trends for 1995-2008 (either in BASE or FIXEMIS) are significantly different at the 95% confidence level despite that they have overlapping confidence intervals.

2. Calculation of data representativeness uncertainty

We have suggested that the extensive statistical sensitivity analysis of Cooper *et al*. [2010] was not sufficient to delimit the meteorological variability in their data, since that information was not captured by the data. However, if we can assume that the meteorologically-driven variability in the ozone distribution in the model is approximately correct, at least in magnitude, then the differences between the model "true" median points and the model points co-sampled with observations can be used as a measure of the "data representativeness uncertainty".

The differences between FIXEMIS_true and FIXEMIS_co-sampled over the period 1995-2008 have a standard deviation of 3.0509 ppbv. This is our estimate of the data representativeness uncertainty. This number will be small if the placement of the observations is such as to capture much of the interannual variability in the ozone field, and larger if they are not. Note that this estimate of data representativeness uncertainty depends only on the RMS difference of a large number of model points, not the individual values. It is therefore probably not too sensitive to model errors, as long as the magnitude of variability in the ozone distribution in the model is approximately correct.

The data representativeness uncertainty can be added to the statistical uncertainty on the trend derived from observations; e.g.:

The standard expression for the size of a one-sided confidence interval on a linear trend is

$$\frac{t\sigma_{y/x}}{\sigma_x\sqrt{n-2}}$$

where $\sigma_{y/x}$ is the standard deviation of the residuals to the fitted regression line, σ_x is the standard deviation of the independent variable *n* is the number of data points and t is the appropriate value of the *t* statistic. For the 1995-2008 trend, *n*=14, *t*=2.179, σ_x= 4.0311 and

$\sigma_{y/x}$ = 2.0482, which gives a 95% confidence interval of ±0.32 (as *Cooper et al.* found). Adding the representativeness uncertainty, i.e., $\sigma_{y/x} = \sqrt{(2.0482)^2 + (3.0509)^2}$, gives a 95% confidence interval of ±0.57. The differences between BASE_true and BASE_co-sampled can also be used to estimate the representativeness uncertainty; they yield a standard deviation of 2.8511 which gives a 95% confidence interval of ±0.55.

When representativeness uncertainty is included the new trends for 1995-2014 become 0.31±0.32, and 0.37±0.32 if research aircraft field campaigns are excluded, so we conclude that these trends from observations are still significant when this source of uncertainty is considered.

Interestingly, this suggests that one could use a model to determine objectively how to optimize an observation network to give the best estimate of trends for the least cost. Conversely, one can put a quantitative estimate on the value of an existing station or stations, by calculating the increase in the confidence interval that results when one removes that data from the observation set.

Supplementary Table 1 Summary of mid-tropospheric ozone trends over western North America. The range on the slope represents the 95% confidence intervals, with gray shading indicating insignificant trends. Unit: ppbv yr^{-1}.

References	Time periods	OBS	BASE		FIXEMIS	
			co-sampled (space,time)[a]	true median	co-sampled (space, time)	true median
Cooper et al. [2010]	1995-2008	0.64±0.31	0.75±0.38	*0.25±0.32*	0.68±0.38	*0.14±0.24*
	1984-2008	0.66±0.20	0.77±0.24	*0.21±0.13*	0.56±0.25	0.02±*0.11*
Cooper et al. [2012]; *Parrish et al.* [2014]	1995-2011	0.41±0.26	0.68±0.25	0.35±0.27	0.56±0.26	*0.19±0.20*
	1984-2011	0.52±0.20	0.73±0.18	0.25±0.12	0.52±0.19	*0.06±0.11*
This Study	1995-2014	0.31±0.21	0.53±0.21	0.36±0.18	0.34±0.23	0.14±0.14
	1984-2014	0.41±0.17	0.61±0.17	0.27±0.10	0.38±0.18	0.05±*0.08*