We thank the referee for the comments.
Here the comments are listed with our reply.

General Comments

The authors should really explain how the selection of stations was made. It is even stated in the title that they use NOAA stations for methane, but only the Mauna Loa Observatory is used. Also the number of stations seems quite limited to accurately represent the global methane.

Indeed only a limited amount of stations were used in the comparison. Following the referee's advice, the simulation was repeated using 16 NOAA stations and CGO(AGAGE) :

Code	Station Name	Country	Lat °	Lon °	elevation / m
ALT	Alert	Canada	82.45	-62.51	190
ASC	Ascension Island	UK	-7.97	-14.40	85
AZR	Terceira Ile., Azores	Portugal	38.77	-27.38	19
BRW	Barrow, Alaska	USA	71.32	-156.61	11
CGO	Cape Grim, Tasmania	Australia	-40.68	144.69	94
CRZ	Crozet Island	France	-46.43	51.85	197
EIC	Easter Island	Chile	-27.16	-109.43	47
GMI	Mariana Islands	Guam	13.39	144.66	0
HBA	Halley Station,	Antarctica, UK	-75.61	-26.21	30
MLO	Mauna Loa, Hawaii	USA	19.54	-155.58	3397
RPB	Ragged Point	Barbados	13.17	-59.43	15
SEY	Mahe Island,	Seychelles	-4.68	55.53	2
SHM	Shemya Island, Alaska	USA	52.71	174.13	23
SMO	Tutuila, Am. Samoa	USA	-14.25	-170.56	42
SPO	South Pole	USA	-89.98	-24.80	2810
ZEP	Ny-Alesund, Svalbard	Norway, Sweden	78.91	11.89	474

The choice of meteorological data seems a bit strange. The operational data of ECMWF have changed vertical resolution at least twice within the study period, definitely affecting the height of each level. This must have an impact on the nudged values and the model results. How did you deal with these issues? Did the meteorological data vertical resolution near tropopause match the model vertical resolution? Also a validation of the computed meteorology is missing from the manuscript.

We thank the referee for pointing this out. The new simulations has been performed using the ERA interim data (Dee et al., 2011), which is consistent for the entire simulation period.

Specific Comments

P1 L25: RMS abbreviation used before defining.
Abstract L27: Root Mean Square deviation (RMS)

P4 L153: Which GFED? GFED4s? Clearly state the version.
GFEDv4: Randerson, J.T., G.R. van der Werf, L. Giglio, G.J. Collatz, and P.S. Kasibhatla. 2018. Global Fire Emissions Database, Version 4, (GFEDv4). ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1293

P5 L178: emission flux, the "e" is missing. – this has been corrected in the manuscript.

P7 L231: A higher resolution of sampling should be used for the CARIBIC data. Daily samples for flight data is far too long.

We apologize for the unclear formulation: the model was sampled daily at 12 UTC at the Stations' location, while for comparison with CARIBIC data the highest possible sampling was used (2 min time-step). Nevertheless, in the new simulation also station values are sampled continuously and then averaged monthly for comparison with observations.

"… calculated CH_4 mixing ratios are recorded and stored at all sampling positions and -times at selected (NOAA (Dlugokencky, 2018) and AGAGE (Prinn et al., 2013) observation sites and along the CARIBIC flight tracks"

P8 L297-305: As mentioned in the general comments, 6 stations are not enough to reach definite conclusions.

This has been changed in the largely revised manuscript.

P10 L360-368: Couldn't this be because of the meteo data?

We exchanged the "operational analyses" with ERA interim with the same effect.
However the high altitude mixing-ratios are averaged over ~500 m grid boxes smoothing down the amplitudes.

P11 L418 and P12 L434 and P13 L495 and Fig13b caption: Be consistent when reporting these numbers.

We corrected the caption.

"…For the years 2007 through 2013 it turns out that a total emission of 25.47 Tg CH_4/y composed of 19.44 Tg TRO and 5.74 Tg SHA optimally reduces the RMS to 0.55 % and approximates the observed ΔNH/SH up to 98%. Fig, 10 …"

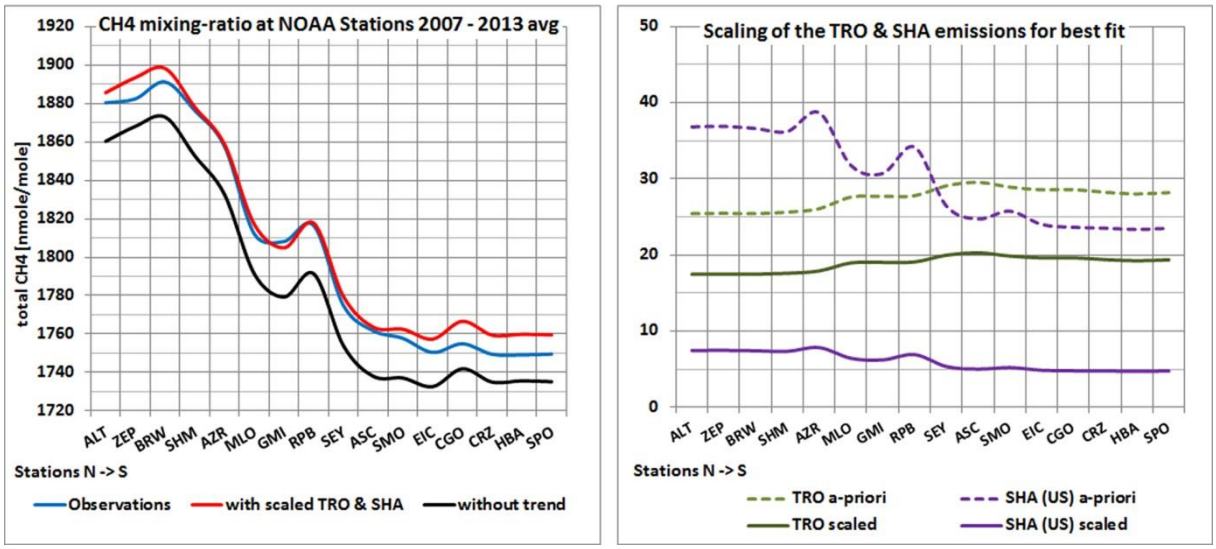


Figure 10: Scaling TRO and SHA emission fractions to fit the all-station observations within smallest RMS: Left: Observations (blue) and total calculated CH_4 without- (black), and with (red) trend period emissions (solid lines right panel).

Right: A-priori estimates (dashed) and solver-scaled (solid) TRO (19.44)- and SHA (5.74 Tg/y) emissions for trend years.

Fig1: A different color code for the different periods would be helpful.

Following referee's suggestion, we change the figure.

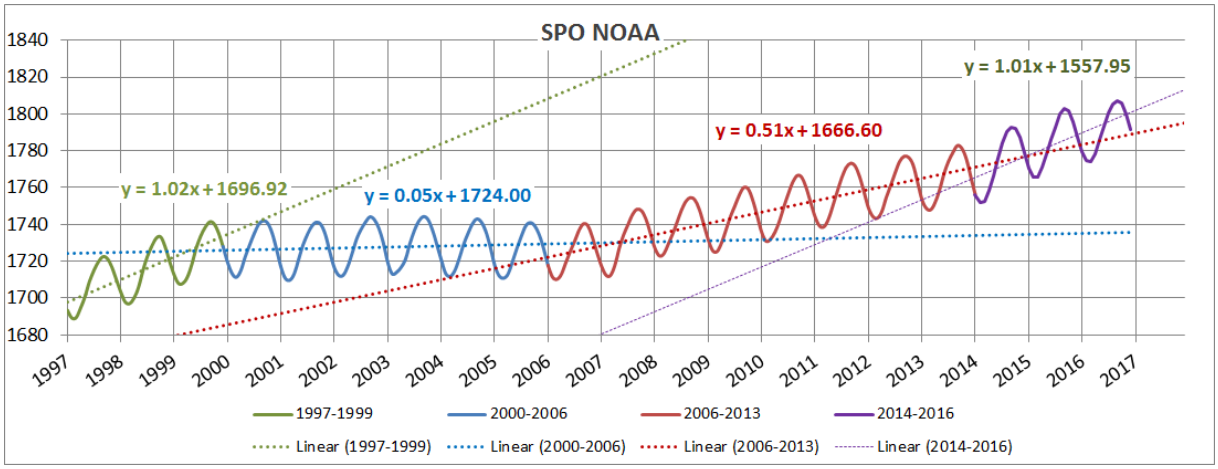
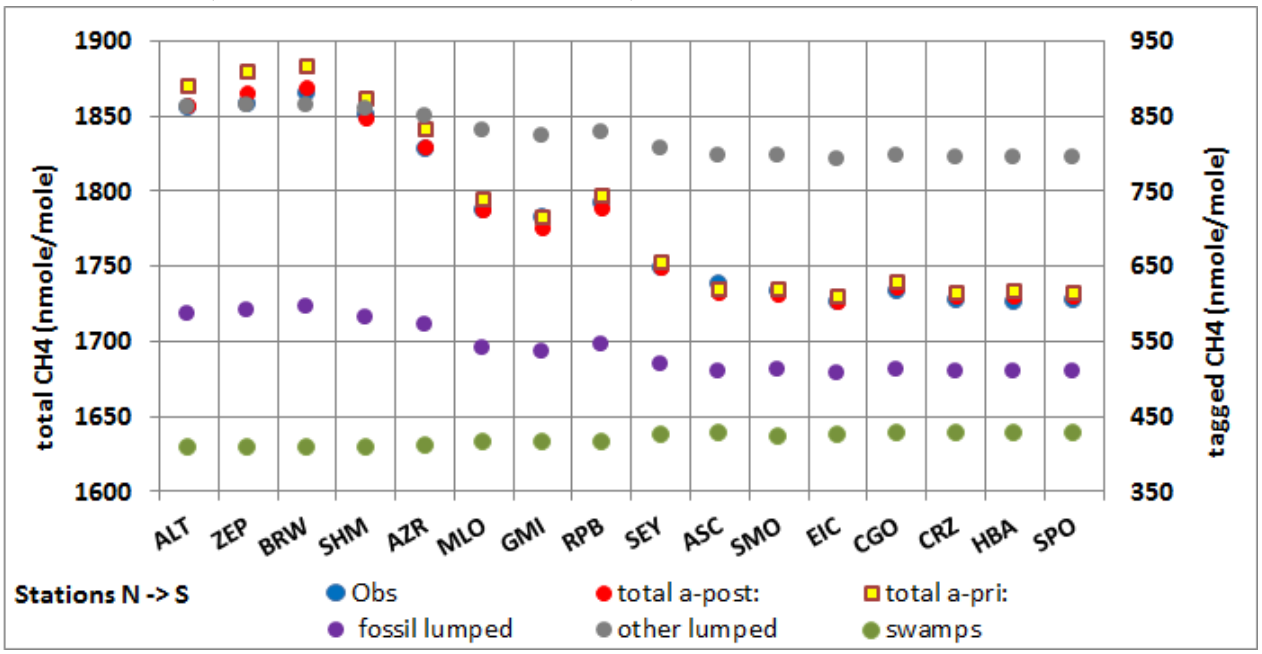


Fig4a and b: I believe the lines connecting the circles are misleading.

New simulation: (I am not sure if looks is better . . . ?)



Figs 5, 6, 9, 11, 14, 15: Really hard to read because of size.

Vector graphics of all figures will be provided for the final publication.

Fig13b: again I fail to see the need for the line connecting stations.
New simulation: (I am not sure if looks is better . . . ?)

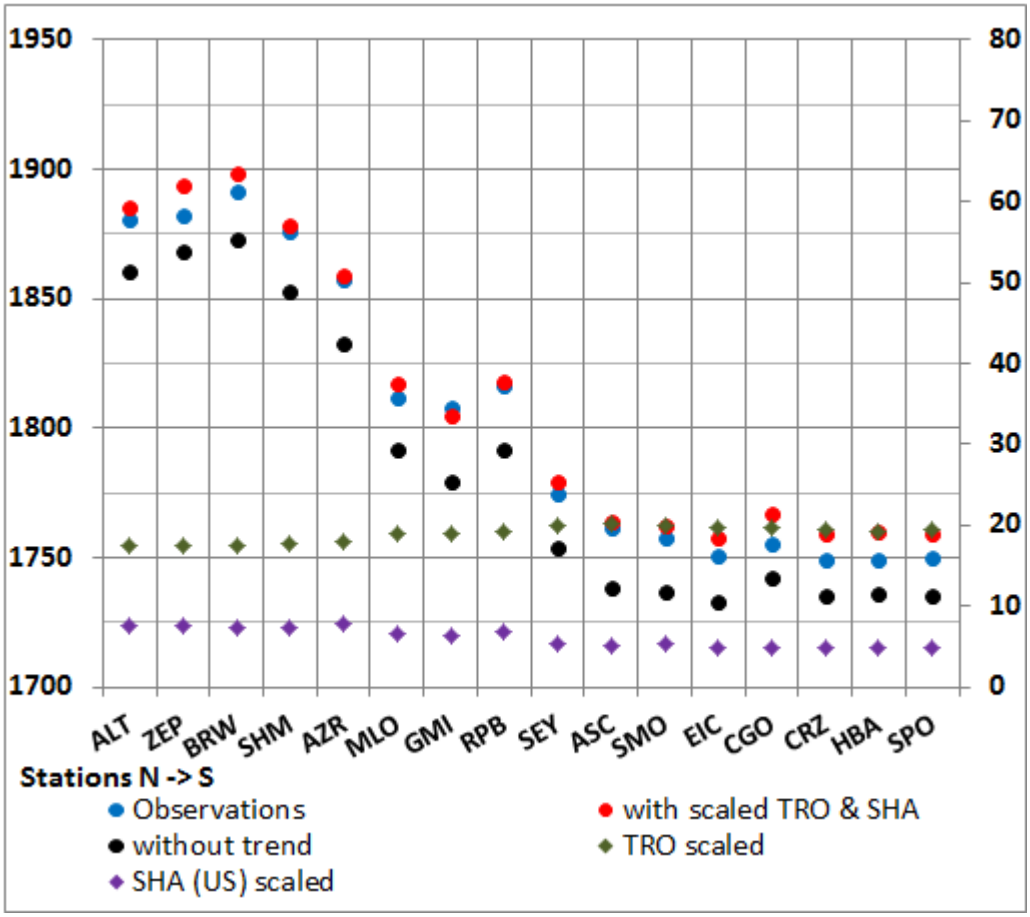


Fig17: State either in the caption or in the legend which set of lines is for every period.
Based on the results of the new simulation, the figure (and caption) has been updated.

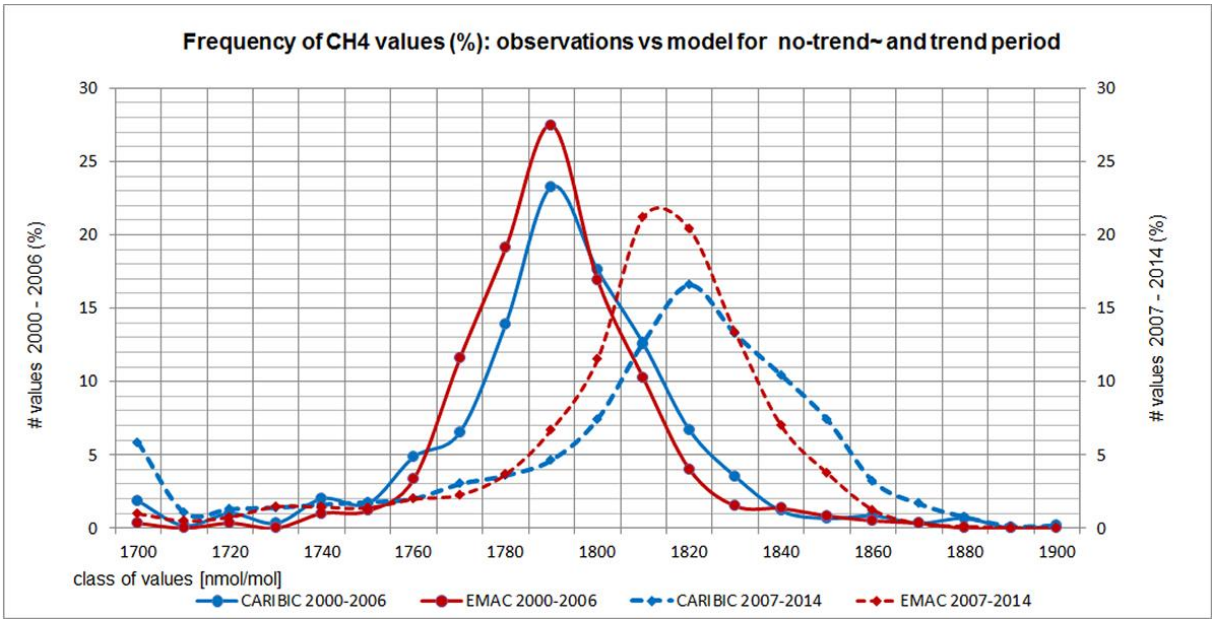


Figure 15: Frequency spectrum of CARIBIC observed and EMAC simulated CH4-mixing-ratios separately plotted for the years 2000-2006 and 2007-2014.

Supplementary material:
There are some inconsistencies between the figures and the captions making it sometimes confusing.
The supplement has been updated.,