Response to comments of referee #2

General comments:

PAHs are substances the can severely impair human health and ecosystems. However, still little is known about their fate in the environment, in particular in the atmospheric environment. By investigating the atmospheric transport of some important PAHs with the state-of-the art model WRF/Chem in China, a region where the pollution by PAHs is supposed to be significant, the authors provide an important contribution to PAH research. The physical-chemical processes they implement into WRF/CHEM are based on sound science. Also, the general setup of the model study appears sound to me.

Response:

Many thanks for the encouraging words. Please kindly find our point-to-point response to your questions/comments below.

1. However, I'm not convinced that the authors treat the PAH emissions - a crucial part of modeling studies - in an appropriate way. They use emissions of 2008 and compare model results to measurements of 2013 and 2003. This only makes sense if PAH emissions did not change within this period. On the other hand, the authors convincingly explain the increasing relevance of PAHs in Asia due to rapidly increasing emissions. If the authors applied inter-annual scaling factors for modeling the years where the measurements were carried out, they should explain in detail how this was done. Otherwise, they should comment on this contradiction. An agreement with measurements alone could also be "right for the wrong reason".

Response:

Thanks for raising this concern. We indeed scaled the anthropogenic PAH emissions from the base year 2008 to the simulated years 2003 and 2013, with scaling factors explained in page 8 line 4. The inter-annual scaling factors are taken from Shen et al. (2013), the same paper that introduces the global PAH emission inventory in 2008 used in our study. In this paper, historical time trends (e.g. our modeling year 2003) are based on the historical fuel consumption data and time-dependent emission factor of PAH. Future time trends (e.g. our modeling year 2013) are predicted using the IPCC SRES A1 scenario supposing a future world of rapid economic growth (Nakićenović et al. 2000). Besides, we have also applied monthly scaling factors following Zhang and Tao (2008). For clarity, we add a new Figure R1 (SI Figure S3) to illustrate the scaling of PAH emissions and rephrase the sentence in page 8 line 4 into *"For specific simulation period, inter-annual scaling factors in the simulated domain are taken from Shen et al. (2013) based on historical fuel consumption data and IPCC SRES A1 scenario supposing a future world of rapid economic growth. Monthly scaling factors are taken from Zhang and Tao (2008)"*.

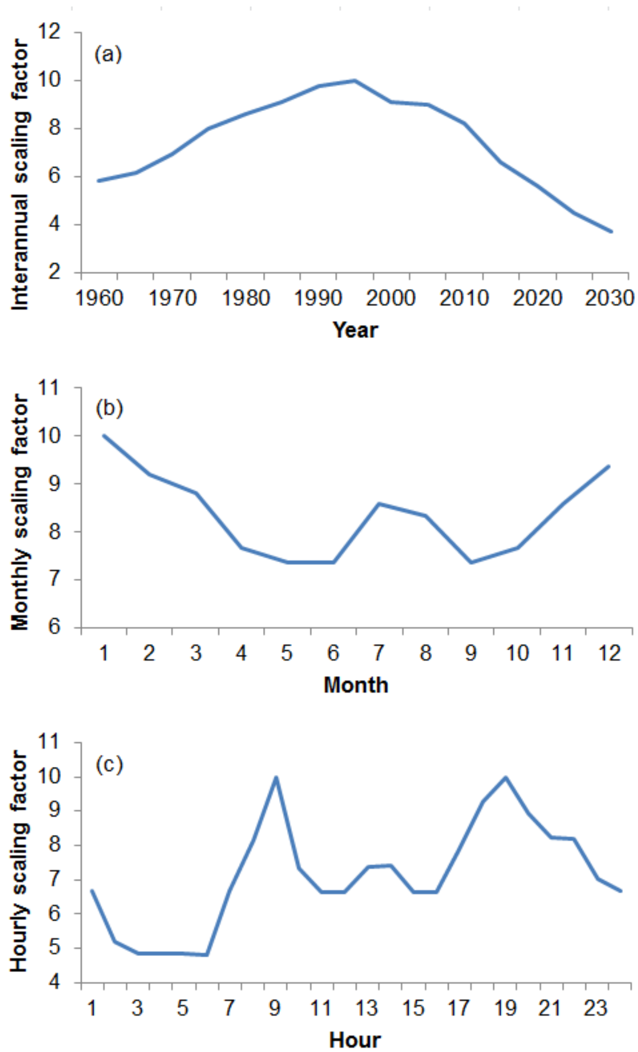

Figure R1 (SI Figure S3). (a) Inter-annual, (b) monthly and (c) hourly scaling factors for PAH emissions.

2. For evaluating the model against measurements the authors compare arithmetic means and Pearson's correlation coefficients of time series, I assume (they don't mention it explicitly). In figure 3 and 5 one can see that the error bars reach negative values. This indicates non-normal distributions and statistical measures like arithmetic mean and standard deviations cannot be applied. The authors must check the distribution and decide based on this which measures to use.

Response:

The correlation coefficients shown in the figures use combined daytime and nighttime data sets, i.e., between 24 observation and 24 simulation samples (12 daytime and 12 nighttime). These correlation coefficients have all passed the Student's t-test with a significance level of 0.05. Correlation coefficients for only daytime or nighttime samples are not included, because of small data sets. In Fig. 5, correlation coefficients are also not included for the same reason. We have added the clarifications of correlation coefficients into the caption of Fig. 3: *The correlation coefficients R use combined daytime and nighttime data sets, passing the Student's t-test with a significance level of 0.05.*

In Fig. 3 and Fig. 5, some of the error bars reach negative values because the error bars indicate the standard deviations which are larger than the arithmetic means. This implies that the data is widely spread out to the mean. The same applies for the observed and simulated PAH data in another similar modeling study applied to this region (Inomata et al., 2012). Although the concept of standard deviation and arithmetic mean are not limited to normally distributed data, we do admit that two statistic metrics are not enough, so that more metrics are now provided in the Table R1 (SI Table S3) and Table R2 (SI Table S4) to characterize model performance in the Xianghe summer case (Fig. 3) and the Gosan winter case (Fig. 5), respectively.

Table S1 (SI Table S3). Observed (obs) and simulated (sim) mean concentration, median, standard deviation (σ), mean bias (MB), root mean square error (RMSE), mean absolute deviation (MAD) in unit ng m^{-3} and correlation coefficient (R) at the Xianghe site averaged over 11–22 July, 2013. The correlation coefficients use combined daytime and nighttime data sets, passing the Student's t-test with a significance level of 0.05.

	gaseous PHE											
	daily				day				night			
	obs	sim	sim-obs	(sim-obs)/obs	obs	sim	sim-obs	(sim-obs)/obs	obs	sim	sim-obs	(sim-obs)/obs
Mean	25.6	22.2	-3.4	-13.4%	13.6	11.7	-1.9	-14.1%	36.5	32.6	-3.9	-10.7%
Median	23.9	15.5	-8.5	-35.5%	13.0	12.5	-0.5	-4.2%	33.3	35.3	2.0	6.0%
σ	13.8	14.8	0.9	6.7%	5.1	5.2	0.1	2.5%	9.6	13.8	4.2	43.8%
MB	-3.0				-2.0				-3.9			
RMSE	10.5				5.4				13.6			
MAD	7.9				4.5				11.1			
R	0.72											

	gaseous CHR											
	daily				day				night			
	obs	sim	sim-obs	(sim-obs)/obs	obs	sim	sim-obs	(sim-obs)/obs	obs	sim	sim-obs	(sim-obs)/obs
Mean	0.63	1.27	0.65	102.8%	0.48	0.65	0.17	34.5%	0.76	1.90	1.14	149.2%
Median	0.56	1.08	0.52	94.0%	0.42	0.61	0.19	45.5%	0.81	1.98	1.17	143.8%
σ	0.28	0.77	0.49	170.9%	0.19	0.27	0.08	42.6%	0.29	0.57	0.28	97.5%
MB	0.67				0.16				1.14			
RMSE	0.97				0.35				1.31			
MAD	0.74				0.30				1.14			
R	0.42											

	particulate CHR											
	daily				day				night			
	obs	sim	sim-obs	(sim-obs)/obs	obs	sim	sim-obs	(sim-obs)/obs	obs	sim	sim-obs	(sim-obs)/obs
Mean	0.85	2.46	1.61	189.6%	0.78	0.52	-0.26	-33.4%	0.92	4.40	3.48	378.0%
Median	0.58	0.98	0.40	69.1%	0.45	0.45	0.00	0.7%	0.97	4.89	3.92	404.4%
σ	0.75	2.65	1.90	252.7%	0.99	0.35	-0.64	-64.7%	0.37	2.52	2.15	579.3%

MB	1.83				0.03				3.48			
RMSE	3.07				0.46				4.23			
MAD	2.00				0.35				3.51			
R	0.59											
particulate BaP												
	daily				day				night			
	obs	sim	sim-obs	(sim-obs)/obs	obs	sim	sim-obs	(sim-obs)/obs	obs	sim	sim-obs	(sim-obs)/obs
Mean	0.78	0.78	0.00	0.3%	0.43	0.14	-0.29	-68.1%	1.10	1.37	0.27	24.5%
Median	0.49	0.29	-0.20	-41.1%	0.39	0.13	-0.26	-67.1%	1.50	0.89	-0.61	-40.6%
σ	0.81	0.70	-0.12	-14.5%	0.24	0.08	-0.16	-67.5%	0.81	0.73	-0.08	-10.1%
MB	0.002				-0.29				0.27			
RMSE	0.61				0.38				0.76			
NMB	0.48				0.31				0.63			
R	0.69											

Table R2 (SI Table S4). Same as SI Table S3 but at the Gosan site averaged over 14–25 February, 2003.

	obs	sim	sim-obs	(sim-obs)/obs
gaseous PHE				
Mean	0.81	1.73	0.92	113.6%
Median	0.54	1.14	0.60	109.8%
σ	0.57	1.95	1.38	242.1%
MB	0.92			
RMSE	2.35			
MAD	1.32			
gaseous CHR				
	obs	sim	sim-obs	(sim-obs)/obs
Mean	0.03	0.03	0.000520	1.8%
Median	0.02	0.02	0.00	-4.6%
σ	0.02	0.03	0.01	69.8%
MB	0.00			
RMSE	0.04			
MAD	0.03			
particulate CHR				
	obs	sim	sim-obs	(sim-obs)/obs
Mean	0.45	0.24	-0.21	-47.5%
Median	0.40	0.06	-0.34	-84.1%
σ	0.35	0.44	0.09	26.2%
MB	-0.21			
RMSE	0.51			
MAD	0.36			
particulate BaP				
	obs	sim	sim-obs	(sim-obs)/obs
Mean	0.020	0.022	0.002	8.3%
Median	0.016	0.018	0.002	13.8%
σ	0.015	0.019	0.004	29.9%
MB	0.000			
RMSE	0.021			
MAD	0.016			

3. The authors use expressions like "good prediction", "fair agreement", "significantly improved", ... to judge their model results. They should explain by which criteria they consider a result (an average or a correlation) as good or not good. As they didn't perform any statistical tests it seems to be pure opinion. I found only one statement where they explain their opinion: Compared with previous studies ... (page 10, line 4).

Response:

Thanks. We add more statistic metrics in Table R1 and Table R2 to characterize model performance in the Xianghe summer case (Fig. 3) and the Gosan winter case (Fig. 5), respectively. Also, we try to discuss model performance by comparing with previous global and regional model studies. For example, in "4.2.2 Evaluation of the Asian outflow" we have the following comparisons *"Model validation so far had been limited to seasonal features (Zhang et al., 2011a; Zhang et al., 2011b), while higher temporal features had not been addressed yet. For example, discrepancies of a factor of 16–476 between predicted and observed average PAH (BaP, CHR, BbF, BkF, IcdP, DahA, BghiP) concentrations at the Waliguan site, a continental background site for ambient air monitoring in western China, were found much larger than at urban or suburban sites (Zhang et al., 2009)"*. Another example in this section is *"Compared with previous studies, our simulated average concentrations of BaP agreed well with the observation (deviation < 10%), while Zhang et al. (2011a) underestimated BaP by about 50%. For the Gosan summer case, our simulated average BaP concentration is 0.006 ng m^{-3} (Fig. S6), much closer to the observed 0.012 ng m^{-3} than the simulated BaP concentration of ≈ 0.001 ng m^{-3} by Zhang et al. (2011a)"*. In "4.2.1 Evaluation at the near source areas", we add comparisons of simulated daily results with one previous regional model evaluation in Beijing (Inomata et al., 2012), since Xianghe is a semi-urban town in the Beijing metropolitan area. Page 9 line 8, "*PAH diurnal variabilities are well captured for both gas- and particulate-phase species at the Xianghe site, with correlation coefficients of 0.42–0.72 (Fig. 3, Table S3) compared with 0.30–0.58 in Beijing (Xianghe is a semi-urban town in the Beijing metropolitan area) by Inomata et al. (2012)"*. Page 9 line 14, "*The model well catches the observed daily average concentration of particulate BaP (observed 0.78 ng m^{-3}, simulated 0.78 ng m^{-3}), while Inomata et al. (2012) underestimated daily concentration of BaP in Beijing by a factor of 2"*.

However, it is not easy to find proper criteria for daytime and nighttime concentrations as well as particulate mass fraction of PAH. Previous model evaluation is unavoidably quite limited due to rare and often incomplete (e.g. particulate phase only) monitoring activities (almost none in Eastern Asia) with low temporal resolution. The results shown in this study are beyond the context of previous PAH modeling studies. Considering that to predict the diurnal cycle and the mass fraction of PAH involves higher temporal resolution and more complex processes than to predict daily concentration, we find the model performance good enough since it could meet the similar standard as daily concentration. When more PAH monitoring data becomes available in the near future, more features of simulated results can be evaluated.

Minor comments:

1. page 1 line 29: To my knowledge the word "tracer" is used for inert substances (which PAHs are not). For substances in very low concentrations I would rather suggest to use "trace gases/compounds/substances".

Response:

Thanks. We intend to use the word “tracer” to include both inert and reactive substances in the modification of transport scheme, not only PAHs. In fact, the usage of “tracer” is not totally consistent throughout the literature, e.g., the widely used chemical mechanism Model for Ozone and Related Chemical Tracers version 4 (MOZART-4) uses “tracer” to represent reactive substances. Therefore, we have changed the word “tracer” into “species” to avoid any misunderstanding.

2. page 2 line 27ff: the information of item 4 is included in item 1 and could be left away. Figures 2c and 2d are not necessary because the authors explain in the text (page 6 line 32ff) why the transport behavior of trace substances is inherent to the model.

Response:

Thanks. Item 1 is to implement and introduce all the latest schemes in Section 2, but item 4 shows sensitivity tests to explain why some of the processes are indeed necessary in Section 5. Considering that item 4 proves the indispensability of new processes and contributes an independent section, it might be better to separate item 4 from item 1 as it is.

Figure 2c and 2d demonstrate an example of BC, a species that is already included in the WRF/Chem model, to exclude the possibility that our incorrect method would cause such abnormal transport behavior. The example of BC also proves that this transport problem occurs not only for PAH but also for any low concentrated species. Therefore, we tend to keep Fig. 2c and 2d for the above reasons.

References

Inomata, Y., Kajino, M., Sato, K., Ohara, T., Kurokawa, J. I., Ueda, H., Tang, N., Hayakawa, K., Ohizumi, T., and Akimoto, H.: Emission and atmospheric transport of particulate PAHs in Northeast Asia, Environ. Sci. Technol., 46, 4941-4949, Doi 10.1021/Es300391w, 2012.

Shen, H. Z., Huang, Y., Wang, R., Zhu, D., Li, W., Shen, G. F., Wang, B., Zhang, Y. Y., Chen, Y. C., Lu, Y., Chen, H., Li, T. C., Sun, K., Li, B. G., Liu, W. X., Liu, J. F., and Tao, S.: Global Atmospheric Emissions of Polycyclic Aromatic Hydrocarbons from 1960 to 2008 and Future Predictions, Environ. Sci. Technol., 47, 6415-6424, Doi 10.1021/Es400857z, 2013.

Zhang, Y. X., and Tao, S.: Seasonal variation of polycyclic aromatic hydrocarbons (PAHs) emissions in China, Environ. Pollut., 156, 657-663, DOI 10.1016/j.envpol.2008.06.017, 2008.

Nakićenović, N., et al. Emissions Scenarios: A Special Report of Working Group III of the Inter-Governmental Panel on Climate Change, Cambridge University Press, New York, 2000.