Response to the short comment on "Global methane budget and trend, 2010-2017: complementarity of inverse analyses using in situ (GLOBALVIEWplus CH4 ObsPack) and satellite (GOSAT) observations" by Dr. Luke Western

Xiao Lu, Daniel Jacob, Yuzhong Zhang, on behalf of co-authors
Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

We appreciate Dr. Luke Western's attention and comments (hereafter Western (2020)) on our work (hereafter Lu et al. (2020)). We organize our response as 1) clarification of some misinterpretation of Lu et al. (2020) in Western (2020), 2) response to the need of parameterization parameter γ (rather than using other methods) and its role in the Bayesian inversion framework (respond to Section 3 in Western (2020)), and 3) response to the concerns of using $J_A(\hat{x}) \approx n$ and $J_O(\hat{x}) \approx m$ to inform the optimal parameterization parameter γ (respond to Section 2 in Western (2020)).

1). We believe that Western (2020) at least partly misinterprets the statements in Lu et al. (2020). Western (2020) states in Page 3, the end of Section 2 that "If $J_A(\hat{\boldsymbol{x}})$ < n, and $J_O(\hat{\boldsymbol{x}})$ < m, why would this suggest an overfit?" and in Page 3, the beginning of Section 3 "My interpretation of, for example, equation 2 (Lu et al., 2020, equation 7), is that if $J_O(\hat{\boldsymbol{x}}) \leq$ m in equation 2, one would assume that the inversion is over confident in its estimated value, and hence the uncertainty is smaller than it should be.". Our statement was "Nevertheless, $J_A(\hat{x}) \gg n$ implies overfit to the observations because the posterior state vector estimates are far outside the estimated errors on the prior estimates." (Line 266). This is not contradict to Western (2020).

2). The need for regularization parameter γ is to avoid overfitting to the observations because the number of observations (1.6 million for GOSAT) is much larger than the number of state vector elements (3378), and the error covariance of the observations cannot be properly quantified. We agree with Western (2020) that there are alternatives to adjust probability distributions rather than using regularization parameter, but they are not really practical in our case. The first alternative mentioned in Western (2020) is to explicitly include uncertainty in parameters within the inversion itself following either an empirical Bayes or hierarchical approach. However, this method does not provide explicitly the analytical solution (which is critical to our objective to quantify the information from inversion) and requires additional analyses, e.g. applying Markov chain Monte Carlo (MCMC) (Ganesan et al., 2014). The second is to create a better prior probability distribution that is representative of actual prior. This is mainly limited by the fact that we do not actually have objective information of the covariance for both state vectors and observations. Therefore, we argue that applying the regularization parameter is the simplest and most applicable method in our inversion framework. The regularization factor itself does not alter the Bayesian nature of the inversion--it is intended to address the lack of covariance structure in the error covariance matrices by modifying the weighting of the prior and observational terms.

3). We agree with Western (2020) that the $J_A(\boldsymbol{x})$ (similarly $J_O(\boldsymbol{x})$) should be expressed in the context of Chi-square distribution with *n* degrees of freedom (*n*=3378 in this case), an expected value of *n* (E($J_A(\boldsymbol{x}) = n$), and a standard deviation of $\sqrt{2n}$. Note that the Chi-square distribution

converges toward a normal distribution for large degrees of freedom (which is the case here). Interpretation of Figure 4 in Lu et al. (2020) shows that the solution with $\gamma = 1$ for the GOSAT-only inversion yields the $J_A(\hat{x}) = 6n \gg n \pm \sqrt{2n}$, suggesting that the posterior state vector estimates are far outside the estimated errors on the prior estimate which indicates overfitting. Our application of comparing $J_A(\hat{x})$ to n for overfitting checking yields consistent result with Zhang et al. (2018), which used the L-curve method (Hansen, 2000) to determine the optimal regularization parameter γ to be 0.05-0.1 for GOSAT observation in a global inversion at 4°× 5° resolution. This method also provides a way to properly weigh the in situ and GOSAT observations in the inversion, by comparing $J_A(\hat{x})$ from the two inversions as shown in the Figure 4. We have revised the text accordingly.

In summary, we appreciate that Dr. Luke Western raises the discussion on this issue and helps to improve the clarification and presentation of the methods. Our method by applying a regularization parameter provides a means to account for unknown error covariance and to weigh different observations in the inversion, while we agree there are alternatives that can be more rigorous and advantageous in some cases. This will be an open question and we welcome further discussions and studies on this issue.

Reference:

Ganesan, A. L., Rigby, M., Zammit-Mangion, A., Manning, A. J., Prinn, R. G., Fraser, P. J., Harth, C. M., Kim, K.-R., Krummel, P. B., Li, S., Mühle, J., O'Doherty, S. J., Park, S., Salameh, P. K., Steele, L. P., and Weiss, R. F.: Characterization of uncertainties in atmospheric trace gas inversions using hierarchical Bayesian methods, Atmos. Chem. Phys., 14, 3855–3864, https://doi.org/10.5194/acp-14-3855-2014, 2014.

Hansen, P. C.: The L-Curve and its Use in the Numerical Treatment of Inverse Problems, in: Computational Inverse Problems in Electrocardiology, edited by: Johnston, P., Advances in Computational Bioengineering, WIT Press, 119–142, 2000.

Lu, X., Jacob, D. J., Zhang, Y., Maasakkers, J. D., Sulprizio, M. P., Shen, L., Qu, Z., Scarpelli, T. R., Nesser, H., Yantosca, R. M., Sheng, J., Andrews, A., Parker, R. J., Boech, H., Bloom, A. A., and Ma, S.: Global methane budget and trend, 2010–2017: complementarity of inverse analyses using in situ (GLOBALVIEWplus CH_4 ObsPack) and satellite (GOSAT) observations, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-775, in review, 2020.

Western, Luke: Short comment to Lu et al.: Global methane budget and trend, 2010–2017: complementarity of inverse analyses using in situ (GLOBALVIEWplus CH4ObsPack) and satellite (GOSAT) observation'

Zhang, Y., Jacob, D. J., Maasakkers, J. D., Sulprizio, M. P., Sheng, J.-X., Gautam, R., and Worden, J.: Monitoring global tropospheric OH concentrations using satellite observations of atmospheric methane, Atmos. Chem. Phys., 18, 15959-15973, http://doi.org/10.5194/acp-18-15959-2018, 2018.