Response to reviewers' comments (Paper ID: 1711)

We would like to thank the reviewers for their feedback and constructive comments, which we respond to below. The blue parts are the value-added contributions of our paper. We would be grateful if you could check it carefully.

To Reviewer 1.

1. **Related work:** The concept related to our aggregation processes has been used in several methods [5, 35]. The linear model of coregionalization (LMC) is the general formulation for multivariate modeling in geostatistics; our baseline (i.e., SLFM) is an instance of LMC, in which latent functions are assumed to be Gaussian processes (as described in [36]). As the reviewer commented, our proposal is based on these two key concepts. We still believe that our paper makes significant technical contributions: 1) integrating aggregation processes with multivariate Gaussian processes that is defined by LMC-based approach; 2) parameter estimation based on the marginal likelihood in which latent GPs $\boldsymbol{g}(\boldsymbol{x})$ and $\boldsymbol{f}(\boldsymbol{x})$ are analytically integrated out; 3) explicit derivation of the posterior GP given multivariate areal observations (see also Appendix B). We will cite the references and discuss them.
2. **Integral over regions:** As the reviewer suggested, it could be an additional contribution to consider an alternative to performing integral over regions. However, we strongly believe that our paper contains enough valuable contributions (as described above); moreover our experiments on real-world data sets from multiple cities show that our model significantly improves the performance in predicting the fine-grained mappings from coarse-grained areal data. This is very helpful for many disciplines including socio-economics [24], epidemiology [25], public security [2], public health [13], and urban planning [34].
3. **Handling observations:** We agree with the reviewer in that introduction of nonlinear link functions (as in the warped GP) and/or alternative likelihoods might help handle some kinds of observations (e.g., rates). One can find, however, in the literature many successful models (e.g., [26]) based on similar assumptions to ours.
4. **Compared methods:** As shown in Table 1 of the manuscript, we confirmed that our model (i.e., SAGP) yields better prediction performance than all the baselines, that is, GPR, 2-stage GP, and SLFM. The detailed comparison with SLFM is important and reasonable, because SAGP is regarded as the extension of SLFM; thus this can clarify the contribution of SAGP. Meanwhile, we also agree with the reviewer in that a more thorough examination from another point of view (e.g., prediction variances) might be helpful, which is one of our future works.
5. **Definition of domain:** In the "Areal data" paragraph (lines 121–130), we describe the case of a single domain (i.e., a city); thus $\mathcal{X}$ denotes a city. For the case of multiple domains, a set of cities is denoted by $\mathcal{X}^{\mathrm{u}}$ on line 176.
6. **Transfer learning across cities:** Given the latent GPs $\{g_l(\boldsymbol{x})\}_{l=1}^{L}$, the observations $\boldsymbol{y}^{(1)}$ and $\boldsymbol{y}^{(2)}$ from respective cities are treated independently (see Figure 2(b)). Although $\boldsymbol{y}^{(1)}$ and $\boldsymbol{y}^{(2)}$ are not directly correlated across cities, the *shared* covariance functions $\{\gamma_l(\boldsymbol{x}, \boldsymbol{x}')\}_{l=1}^{L}$ for the latent GPs can be learnt by transfer learning based on the data sets from multiple cities; thus the spatial correlation for each data set could be more appropriately output via the covariance functions, even if we have only a few data sets available for a single city. We will clarify it.

To Reviewer 2.

1. **Difference from LMC [36]:** Please see the 1st response to **Reviewer 1**.
2. **Extension to multiple domains:** In the experiments, the data were normalized so that each variable in each city has zero mean and unit variance. We will clarify it in our proceeding. Please see also the 6th response to **Reviewer 1** for additional information about the transfer learning across multiple domains.
3. **Significance:** As described in the 2nd response to **Reviewer 1**, we believe that our paper provides significant contributions not only to the field of machine learning but also to various other fields. Moreover, our formulation provides a general framework for handling aggregated data and offers a potential research direction that can be explored in future work; for instance, it has the capability to consider the data aggregated on a higher dimensional input domain $\mathcal{X}$, e.g., spatio-temporal aggregated data.
4. **Using inducing points:** Extension of inference algorithm with inducing points is possible, which is one of the options for the efficient computation. However, the computation complexity $O(|\mathcal{P}_s||\mathcal{P}_{s'}||\mathcal{D}|)$ of our model (see line 230) is not too high; actually the average computation times for inference were 1728.2 and 115.1 seconds for the data sets from NYC and CHI, respectively, where the experiments were conducted on a 3.1 GHz Intel Core i7. The results show that our inference algorithm is efficient enough to run in realistic-time scales.

To Reviewer 3.

1. **Related work:** We will cite the reference [36]. Please see the 1st response to **Reviewer 1** for more details.
2. **Experiments:** In this paper, we focused on the refinement task of poverty, crime, and PM2.5 data sets. The main reasons are as follows: 1) refining these data sets is strongly desired in the practice of socio-economics [2, 24] and public health [13]; 2) our experiments basically replicate the experimental setup in the existing work (2-stage GP) that has proposed in [26]. We agree with the reviewer in that the extensive experiments on more data sets from more cities are left for future work.

References

[35] R. Murray-Smith and B. A. Pearlmutter. Transformations of Gaussian process priors. In *DSMML*, pages 110–123, 2004.

[36] M. A. Álvarez, L. Rosasco, and N. D. Lawrence. Kernels for vector-valued functions: A review. *Foundations and Trends ® in Machine Learning*, 4(3):195–266, 2012.