Report on "Bayesian Functional Principal Components Analysis via Variational Message Passing"

This paper develops a variational Bayesian functional principal components analysis using Karhunen-Loève decomposition without smoothing and estimating a covariance surface. More specifically, they incorporate the notion of variational message passing(VMP) over a factor graph because it removes the need for rederiving approximate posterior density functions if there is a change in the model. The performance of the proposed model is investigated by means of simulation studies and an application to temperature data collected from various United States weather stations. Overall, the paper is well written, but there are some concerns further to be considered to better understand the proposed approach with VMP. My concern is that the proposed approach and numerical results here, although somewhat useful, are not compelling from the perspective of novelty or application. Some novel features of the proposed approach can be emphasized and clearly explained in the paper. In the following, I have some specific suggestions/concerns to improve the paper for revision.

- In Section 2, each function is represented by the sum of a mean function and the sum of the infinite linear combination of the orthonormal eigenfunctions. However, in practice, this infinite sum is truncated by the first L orthonormal eigenfunctions, as the authors say. In the real dataset, naturally we don't know the value of L, thus to select this L is also a important problem. Therefore, it would be helpful to explain some criteria for selecting the number of first eigenfunctions for truncation. Additionally, in the first subsection of the Section 2, they simply say that they approximate the unknown functions to splines using O'Sullivan basis without justification. Since there are many other Basis expansions such as wavelets basis or orthogonal basis functions and each expansion has its own desirable theoretical properties, at least some motivations/explanations would be helpful for readers to see why they consider O'Sullivan penalized splines

- They briefly comment that the method using Karhunen-Loève decomposition has several advantages over the other class of methods using covariance decompositions. Although they provide why these advantages exist shortly, I think more justifications will be better. For instant, it will be better to provide why the method using Karhunen-Loève decomposition is more appropriate handling sparse or irregular functional data over the conventional methods. Also, in the first subsection where the model construction proceeds, the prior specifications are not complete since the covariance matrices and the scale parameter of half-Cauchy distribution are not specified. Thus, it is desired to make a complete description of the hyperparameter specification.

- In Section 3, they mention that if there is a change to the model, the parameter vector updates must be rederived in the MFVB method, while this kind of problem does not arise in the VMP method. However, since VMP method is somewhat foreign to the statisticians, it is better to illustrate why this difference happens between the MFVB and VMP method. In Section 5, there are some proposals for this work to be more persuasive. Firstly, it will be better to compare the results in terms of accuracy and computation speed with that of conventional methods in Section 1. Although this work's own contribution is to address a functional data model via variational message passing and it works well considering the results in Section 5, as a newly proposed methodology corresponding for

FPCA model, it will be also indicative to show what aspects get better compared to the conventional FPCA methods. Along with this context, I think results under more complex settings should be given.

- In the simulation setup, all functions are just the linear combinations of a mean function and two eigenfunctions, while real functional data would have more complex structures. That is, the simulation settings used in this section seems to be too simple, considering the structure of real functional dataset. Thus, additional experiments and simulation studies under more diverse settings would be expected, such as with more/less observational points or more eigenfunctions. Plus, the estimation and computational costs seem to heavily depend on the number of basis of O'Sullivan spline. If the number of basis is increasing, the precision of estimation measured in terms of ISE will be stronger while the computational cost gets more burdensome and vice versa. This is because the volume of design matrix depends on the number of basis. Therefore, it will be also meaningful to find the optimal number of basis in terms of ISE and computational speed.

- Additionally, convergence of the algorithm would be an issue since VMP algorithm is one of the deterministic algorithms. Therefore, showing concrete evidence that the algorithm successfully converges would be better for the validity of this algorithm. In Section 6, I think it would be also available to expand the proposed model to a multilevel one. With this expansion, analysis on the yearly data on each state, not the averaged data over 25 years seems to be possible. Naturally, it would be more meaningful to analyze the yearly data rather the averaged one since this averaged one has little variation. Also, analyzing the yearly data, we could catch some tendency as time goes on. Thus, if it is available, expansion to the multilevel model should be considered and new analysis with this expanded one will be desirable.

- Finally, as the computational technicalities associated with the proposed VMP algorithm are quite challenging and important for fast computation and scalability, it would be useful to provide code or make it accessible in the github for public use, so that implementation and reproduction of the authors work can be useful to practitioners and other statisticians.