Review of "Informative autonomous sampling of oceanographic variables using joint excursion sets

January 1, 2020

1 General Comments

1. Overall I found this paper interesting. Moreover it addresses the topic of spatial design a topic that does not receive the attention it deserves. It is driven by the practical need to direct autonomous underwater vehicles where to take samples. The sequential sampling strategy may be a valuable addition to the designer's toolkit.

2. The things that stand out as novel in this paper are: the need for a sequential design for driving an autonomous under water vehicle; a way of characterizing water plumes that are washed down from the land into the sea. Papers co-authored by one of the authors (Ginsbourger) of this paper inevitably means some overlap of ideas with the other papers, although the two step ahead greedy design algorithm seems novel. I seem to recall that Jon Lee published that algorithm albeit for determining the maximum entropy optimum by finding optimal points sequentially over space not time as in this paper. The literature on spatial design is huge and often involves spatial sequential designs. The handbook chapters by Zidek and Zimmerman (or is the other around) might be worth citing for the benefit of readers interested in some of the background.

3. Overall the paper is remarkably free of a need for editorial improvements. I will attach a marked up copy indicating a few errors as well as suggestions for minor changes.

4. I am not clear about the intended audience for this paper. Or to put it another way, for whom is this paper important? The case study or the

design methodology? Case studies offer concern specific events deemed of importance e.g. the much–studied O-ring failure long ago on a US rocket launch. Such a case study may then drive the innovation of statistical methods for an analysis of data that were collected because of that case's importance. More likely the case study was intended as a tutorial on the use of a novel design methodology. In fact the Closing Remarks section points "the opportunities available real-time multivariable spatial data-gathering" that "statistician can exploit".

However, the user community is pretty tech-savvy nowadays, and that community might well be the primary beneficiary and more motivated than the statisticians to use the tools for sequentially sampling rivers for mapping plumes. That suggests including the relevant code for running the analyses, especially the simulation studies. An alternative journal to a statistical one might be more appropriate from that perspective.

5. Whatever the intended audience, the paper would benefit a lot from being shortened. I slogged through the first six papers before I discovered what the paper was to be about on page 6:

> "In this paper, with a focus on mapping the river plume, we reward the designs taht improve the classification of water masses as a means to characterize teh frontal zone and thereby the extent of the river plume."

Later we find on page 7:

> "Our goal is to construct adaptive sampling strategies based on the continuous evolution of the EPs."

These sentences should have been put right at the beginning after suitable rewording along with the "In this work..." paragraph on page 3. If statisticians are the intended readers, deleting much of the first five pages would improve the paper. There are other places where technical material can be condensed without much if any loss. I will mention some later.

6. It is not clear why the binary classification and excursion sets are being used. Is there really a sharp boundary at the edge of a plume? Why not mapping the salinity and temperature fields instead? After all, on page 19, "achieving good predictive performance" is mentioned as a desirable feature of a good design. A lot of work has been done

on that topic.

And what about indicator kriging (along with logit regression) that generalizes indicator kriging that has been used to identify plumes? Modelling the $p(x)$ is more like probit analysis-what is the advantage.

7. The sequential sampling approaches are not novel. they are mentioned as examples in the cited paper by Bect et al. What is novel is the study of how all these methods work first in simulation studies as well as well as in the case study. That work will be a valuable contribution.

8. While I can see some advantages of simplicity in the IBV of Bect et al for binary data, one could have used entropy as it has information-theoretical credentials. Furthermore it could cover design both for spatial mapping as well as indicator mapping. Finally model uncertainty could be included in the overall assessment of uncertainty. The more general issue here is what objective function should be optimized in the design? Are we in a multi-attribute criterion situation?

9. On the issue of covariance separability, objections have always been raised when I have used it as its unrealistic. My rejoinder has always worked, that of citing the "all models are wrong", George Box's famous phrase. But what I was able to show in each case is a cross-validatory assessment that showed the model was "useful". I wonder if such a model assessment could be made here to support the assumption of separability and other things as well. By the way, in spatial mapping at least, point predictors are pretty robust against misspecification of the covariance. Where things go sideways is in the confidence bands or ellipsoids they produce. What are the error bands when modelling binary responses? Surely the predicted plumes have fuzzy boundaries.

10. The strategy at the bottom of page 17 seems ad hoc. Could you not use the data collected with the two ahead design to get better estimates of parameters and then with these parameters and thereafter a better myopic design, e -

11. Are these results generalizable? In other words, how much plume-to-plume variability is there? One of the problems with modelling using a lot of parameters, is there instability over replicate experiments. To put it another way which model parameters represent plume-population parameters and which represent random plume effects?

12. The separability assumption is to be relaxed in future work. Should publication of this case study await the proposed improvement? It would make for a much better paper.

2 Specific comments

1. "joint excursion sets" in the title would be a turnoff for most statisticians. Having read the paper pretty carefully I still am not sure why they are so labelled.

2. Why on page 13 were those particular values chosen for the means μ? Would a different choice have made a difference?

3. I don't think Figure 3 conveys too much additional benefit and eliminating it for space might be considered.

4. Why are the ES's called "excursion sets? Why excursion probabilities?

5. I did not spot a definition for the "waypoint graph" a phrase used several times in the paper. Again points to the need to identify the audience-I doubt that most statisticians would know that phrase.

6. In statistics journals, random variables are commonly labelled by Roman characters e.g. Z, parameters in Greek e.g. ξ. I realize that in other scientific domains that distinction is not always made but I do recommend the change if the paper is to appear in a stats journal.

7. There is not much of novelty in the Gaussian Process theory used in the paper- its pretty much standard. I suggest that the derivations stated in detail be summarized by just the results. Statisticians will know these things and users will not care.