

Real-Time Information Processing of Environmental Sensor Network Data using Bayesian Gaussian Processes

M. A. Osborne and S. J. Roberts, A. Rogers and N. R. Jennings

4th April 2011

Thank you for comments on our submitted manuscript. We would like to thank all the reviewers for their insightful and useful comments. We have added additional explanation and performed additional empirical evaluations, as requested, and we believe that the changes we have made in response have significantly improved the quality of the paper.

In summary, we have made the following broad changes:

1. To better integrate the theoretical and practical contributions of the paper we have added an additional expository paragraph at the end of Section 2 which makes the application of the Gaussian process approach clearer, before the details are presented.
2. We have added an additional section (Section 7) after the discussion of the computational complexity of our algorithms to discuss how they may actually be used within a real deployment.
3. We have revised the Kalman filter implementation, improving its performance in comparison to the naive approach, and showing that it is still clearly outperformed by our GP approach. The revised results are derived by carefully optimising the use of the Kalman filter to each dataset. Specifically, we now optimally choose the length of history used to make prediction in each dataset (rather than using a default value) in order to present the best prediction performance that was achieved.
4. We have performed the additional experiments requested and now also present the results for the independent GP in Tables II and IV. We note that for the air pressure results presented in Table III, we only have one sensor, and thus, we cannot directly compare our multi-output GP and the independent GPs in this case. Rather, the aim here is to demonstrate the effectiveness of the censored data noise model in contrast to the Gaussian noise model used by the Kalman filter.
5. Furthermore, for the MIDAS dataset presented in Section 5.4 (for which the Kalman filter is inappropriate since it cannot express spatio-temporal correlations) we have now also compared our approach to a state-of-the-art Gaussian basis function predictor.
6. We have added a definition of NMSE in Section 5.

In the following pages, we describe in detail the changes that we have made to address the specific points of each of the reviewers.

Referee: 1

- The paper reads almost like two papers that were combined. The first one deals with the GP optimization and may not be that easily accessible to those who focus on the most practical aspects of sensor networks. The second one provides a thorough evaluation with some real datasets, providing implementation details that may not be of significant interest to those who are more interested by the GP estimation. With respect to this, Algorithm 1 helps to some extent but somehow does not provide the roadmap that the other reviewer asked for. In my opinion having a more informal description earlier on, would help those non-specialists follow GP track approach.

We would contend that the very strength of our paper is that it presents a number of novel theoretical development that are necessary for the practical use of Gaussian processes within this domain (specifically, presenting a model that describes correlations and delays between readings from multiple sensors, providing computationally efficient approaches to update and downdate computation to address streaming data, and to handle realistic noise models whereby data may be rounded to integer values, or censored), and also that these developments are thoroughly evaluated using data from three different real-world sensor networks. The motivation of this work has been to develop theoretically principled approaches that address the challenges of dealing with real-world data, and thus, presenting both aspects of the work together is central to this aim. Previous work tends to keep these aspects separate and this, we believe, is a mistake.

That said, we have tried to balance the demands of readers whose principle expertise in one of these areas. To this end, we have added an additional expository paragraph at the end of Section 2 which makes the application of the Gaussian process approach clearer. We do this before the details of the approach are presented.

- An effort has been made to analyze the advantages of the proposed method in terms of complexity. What is missing, in my view, is clearly articulating what would be the constraints in practical deployment and how the proposed approach meets those constraints. The authors mention in their response that their algorithm would run in a base station. A base station could be almost anything: i) one very simple node that coordinates with other sensors, or ii) a powerful centralized processor located elsewhere in the network, etc. In the results presented comparisons are provided for an off the shelf PC, which may have significantly more computation power than a "base station" as in case i). On the other hand, in case ii), even if data is collected online, computational resources would be plentiful. In summary, reducing complexity (again the main contribution here) seems to be desirable, but the authors have not clearly spelled out for what kinds of scenarios it becomes necessary (or shown that they meet the complexity/delay constraints required for those scenarios).

You are correct with this point. Thus we have now added an additional section (Section 7) after the discussion of the computational complexity of the algorithm to address exactly these issues. In doing so, we have moved the discussion of our live implementation from Section 4 to Section 7, where we discuss how the GP formalism is actually used in this case.

Referee: 2

In looking at the comparison results that follow:

(a) the performance of Kalman filtering shown in Tables I and II is strikingly bad, even compared to the naive approach. This needs to be explained with reference to the above statement, where the naive approach is supposed to provide a lower bound. Even otherwise, such an 'off the charts' performance by a method introduced as the state of the art alternative is cause for serious concern. The technical reasons for this need to be explained at some detail.

(b) in all the results presented, in Tables I, II and III, the performance of Kalman filtering is either significantly worse or very close to the naive approach. In such a case, is it not better to drop Kalman filter as a benchmark?

In line with your observations, we have revised the Kalman filter implementation, improving its performance in comparison to the naive approach, and showing that it is still clearly outperformed by our GP approach. The revised results are derived by carefully optimising the use of the Kalman filter to each dataset. Specifically, we now optimally choose the length of history used to make prediction in each dataset (rather than using a default value) in order to present the best prediction performance that was achieved. This process is described in detail on page XXXX.

(c) in view of (b), it might be better to use the only Independent GP and naive approach as the benchmarks and present a uniform set of tables with the performance of three schemes i.e. naive, Indep. GP and Multi. GP. Currently tables II, III, and IV are missing the independent GP performance.

We have performed the additional experiments requested and now also present the results for the independent GP in Tables II and IV. We note that for the air pressure results presented in Table III, we only have one sensor, and thus, we cannot directly compare our multi-output GP and the independent GPs in this case. Rather, the aim here is to demonstrate the effectiveness of the censored data noise model in contrast to the Gaussian noise model used by the Kalman filter.

Furthermore, for the MIDAS dataset presented in Section 5.4 (for which the Kalman filter is inappropriate since it cannot express spatio-temporal correlations) we now also compare our approach to a state-of-the-art Gaussian basis function predictor.

(d) (minor) include a formal definition of NMSE in Section 5.

This has been added.