Untitled EnKF paper

Keiran Suchak

December 8, 2022

Abstract

Write abstract here.

1 Introduction

Contents:

- Provide context and motivation for investigation.
- Outline aims and objectives.

Main aim: show that an Ensemble Kalman Filter (EnKF) can improve the accuracy with which an agent-based model simulates a system of pedestrians.

Points of distinction to highlight:

- Defining an approach for defining whether an agent is active or inactive in an ensemble of models.
- Comparing error in ensemble mean with mean of errors of ensemble-member models.
- Explaining the importance of an appropriate summary statistic (median instead of mean) when calculating the average error over time.
- Explaining the importance of considering time-steps when a sufficient number of filters are still running when collecting summary statistics of multiple filter runs.
- Using EnKF to improve the accuracy with which an ABM simulates a pedestrian system.

Other things to mention in the introduction:

- Pseudo-truth data. "The purpose of the base model for these experiments is simply to provide a state against which to compare the performance of filters."
- Broad overview of the experimental approach (how do the experiments show that the EnKF is/isn't working?)

2 Background

- Discuss previous relevant work:
  - **?**
  - **?**
  - **?**

3 Methods

3.1 Model

Explain about `StationSim_GCS`.

Things to note:

- What an 'active' agent is

3.2 Ensemble Kalman Filter

- Explain about the Ensemble Kalman Filter (**?**), which is based on the Kalman Filter (**?**).
- Point to previous experiments in the thesis. E.g. "Results: The data assimilation scheme was tested for a range of different filter parameter values, and it was found that improvements in filter performance resulted from increases in the ensemble size, reductions in the standard deviation of the observation error and reductions in the number of time-steps between successive attempts to assimilate observational data into the system." (p113)

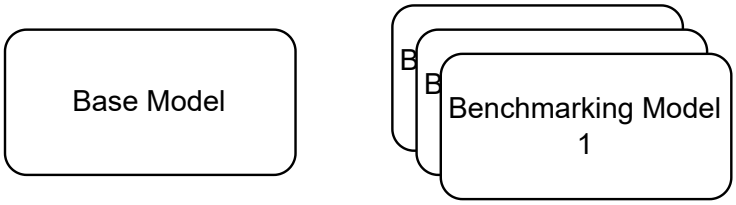
4 Experiments

The experiments, outlined visually in Figure 1, aim to demonstrate that the EnFK can improve the accuracy of a pedestrian system in comparison to a baseline scenario with no data assimilation. In order to better understand the impact of data assimilation on an agent-based model, rather assess the realism of the model itself, we use the "identical twin" approach **?**. In this approach, a 'Base Model' is used. A Base Model is an instance of `StationSim_GCS` that is used to generate 'pseudo-true' data that are taken as the real-world observations in the experiments (in lieu of data from a real crowd).

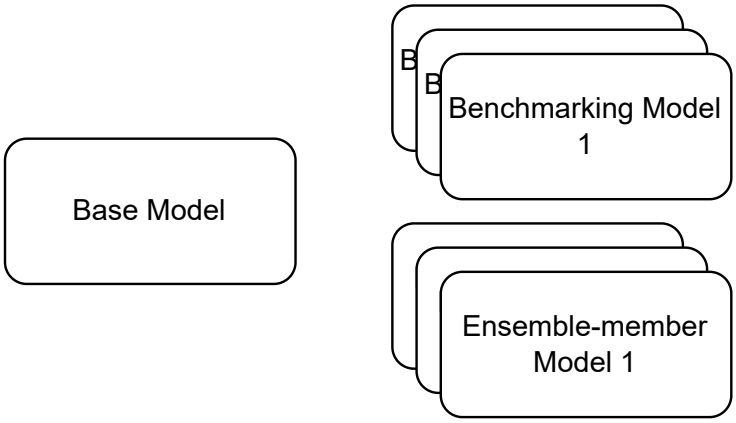
The initial experiment, ('benchmarking', Figure 1a) seeks to establish a benchmark against which to compare subsequent implementations of the EnKF. This is achieved by running an ensemble of models, each initialised as duplicates of a base model which is used to generate pseudo-truth values for the system state.

The second experiment (Figure 1b) seeks to XXXX. It does this by exploring the variation in the accuracy of individual ensemble members. This is achieved

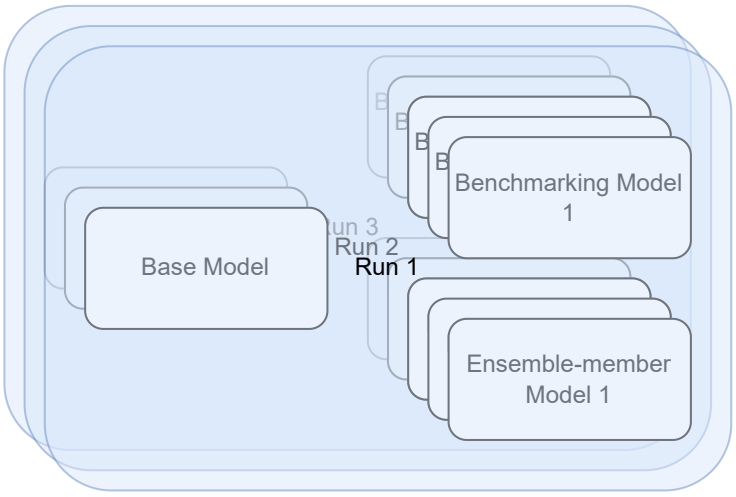
Keiran what is the higher-level purpose of this experiment?



(a) Experiment 1: Benchmarking



(b) Experiment 2: Exploring Ensemble-Member Models



(c) Experiment 3: Implementing the Ensemble Kalman Filter

Figure 1: Graphical outline of the three experiments. Note that the 'base model' is use to generate pseudo-true observation data

by running a single EnKF which maintains a benchmarking ensemble of models, providing a baseline against which to compare results, along with and ensemble of models that are periodically updated by the EnKF assimilation process. In such a situation, we are able to compare the average error per agent in each of the ensemble member models at each assimilation time-step.

The final experiment (Figure 1c) takes this exploration a step further by seeking to capture the variation in error at an ensemble level. This involves running a collection of EnKFs for the same set of model and filter parameters, and in each case gathering data regarding the variation in the error in the ensemble mean state over time, comparing this with the variation in the corresponding collection of benchmark errors.

Again, why?

4.1 Measuring Error

- Talk about measures used when running experiments with multiple EnKFs to ensure that outliers don't skew results:
  - Median instead of mean error.
  - Only considering time-steps when a sufficient number of models are active.

When calculating error, it is also necessary to consider whether an agent is 'active' or not as once an agent has left the simulation they should not be included in an error calculation. However, an agent might be active in the base (pseudo-truth) model and inactive in some or all of the EnKF ensemble members, or vice versa. Here we assume that an agent is active only while it is active in the EnKF ensemble because in a real situation we would not necessarily have access to the true positions of the individuals in the crowd, so could only assess an agent's status from the information available in the ensemble of models. Hence an agent is considered active if its most common (i.e. modal) status across the ensemble is active.

4.2 Experiment 1: Benchmarking

The initial experiment to be performed is to develop a model baseline, establishing the effectiveness of `StationSim_GCS` in modelling a system in the absence of any information whilst running. Error is quantified by calculating the distance between the position of an agent estimated by the benchmarking ensemble of models and the position of the corresponding agent in the base model, d_i:

$$d_i = \begin{cases} |\hat{\mathbf{x}}_i - \mathbf{x}_i| & \text{if } i\text{th agent is active;} \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $\hat{\mathbf{x}}_i$ is the x-y position of the ith agent estimated by the benchmarking ensemble of models and $\mathbf{x}_i$ is the x-y position of the ith agent in the base model. Recall that the 'base model' is an instance of `StationSim_GCS` that is used to generate data that are assumed to be drawn from a real crowd. Inactive agents are excluded from the error calculation (see Section 4.1).

4.3 Experiment 2: Exploring Ensemble Members

Better word than 'exploring' above? Something more specific?

4.4 Experiment 3: Assesing the EnKF

5 Results

5.1 Benchmarking

6 Conclusion