Simulated data

June 2024

This section presents two experiments that demonstrate the impact of hyperparameter selection in capturing different patterns in the data. Each experiment simulates a single task with a planned duration of 90 days, exhibiting unique seasonal characteristics and trends. Daily data was generated over a 30-year period, indicating whether the task was performed on each day. Further details on the data generation process can be found in the Appendix.

For these experiments, the validation set consists of start dates for each month over the next year. The following hyperparameter combinations were considered to smooth the annual seasonality: the *smoothing* parameter with values of 0 or 7, and the *cosine* parameter with *smoothing* equal to 0. Each combination was tested with the *alpha* parameter varying between 0.9, 0.95, or 1. The *mc* parameter was not used to simplify the analysis of the graphs, as the data do not present temporal dependence. The project execution time was estimated, and the pinball loss weighted by the right tail was calculated.

Figure 1 presents the results of Experiment 1. In Figure 1-(A), a trend is observed in the data, with the percentage of days worked being higher in more recent years, represented by the darker lines. In Figure 1-(B1), all hyperparameter combinations capture the annual seasonality, with probabilities being higher for lower values of alpha, as they give more weight to more recent years. In Figure 1-(B2), performance improves as alpha decreases, an expected result due to the trend in the data and the choice to give more weight to more recent years. Moreover, *smoothing* and *cosine* performed similarly. This suggests that both the parametric sinusoidal wave approach and the moving average approach are suitable when the data follow the expected behavior.

Figure 2 presents the results of Experiment 2. In Figure 2-(A), the data show more complex seasonalities and no trend. In Figure 2-(B1), smoothing fits well with the data behavior, while the parametric method fails to fit adequately. In Figure 2-(B2), the moving average shows superior performance, indicating that the parametric method employed is insufficient to model more complex patterns.

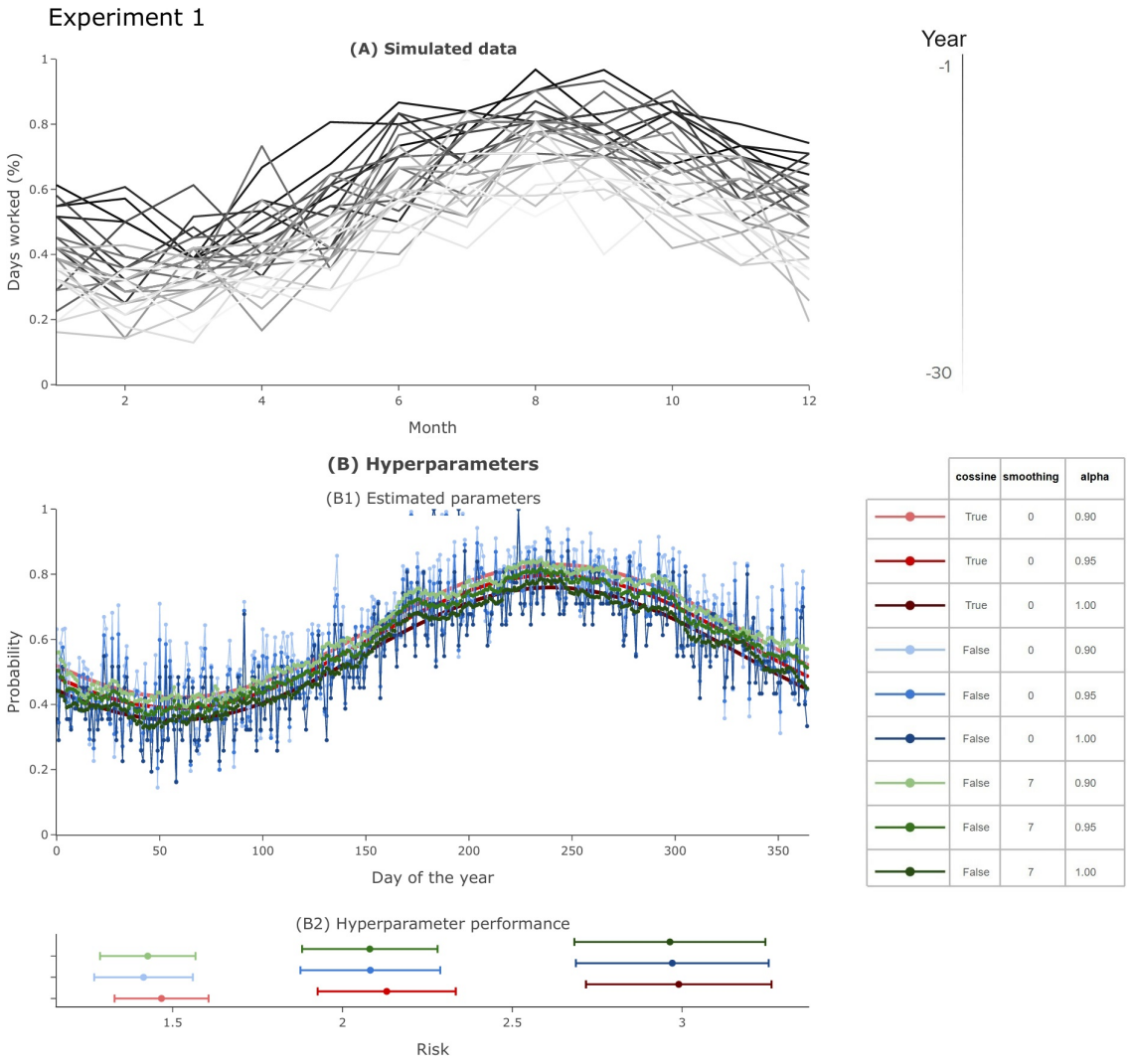

Figure 1: Results of Experiment 1. (A) Percentage of days worked over the 12 months, with each line representing a year in the historical data. (B) Hyperparameters: (B1) Estimated Parameters: Probability throughout the year with different hyperparameter configurations. (B2) Hyperparameter Performance: Risk associated with hyperparameter combinations, measured by right-tail pinball loss.

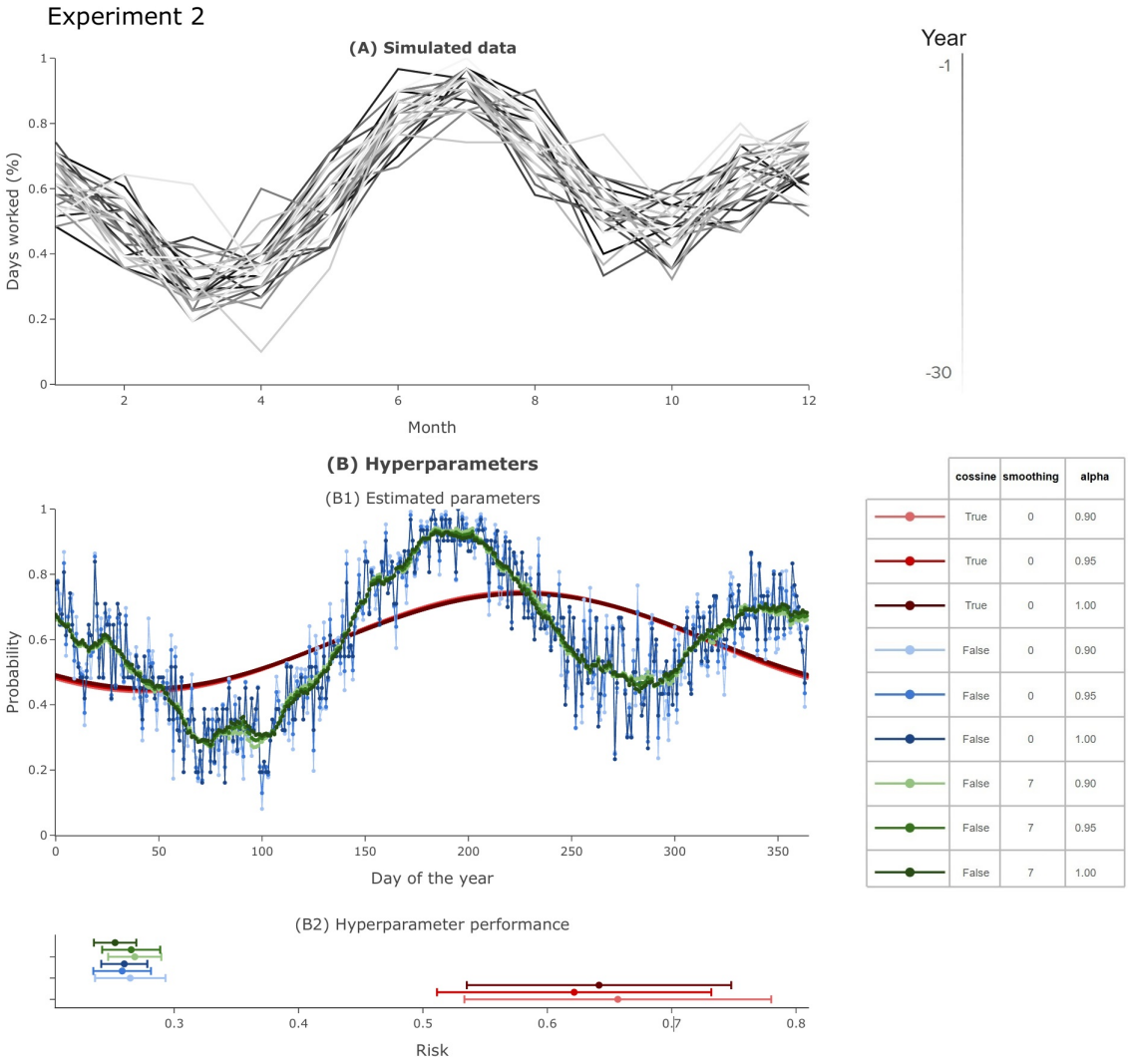

Figure 2: Results of Experiment 2. (A) Percentage of days worked over the 12 months, with each line representing a year in the historical data. (B) Hyperparameters: (B1) Estimated Parameters: Probability throughout the year with different hyperparameter configurations. (B2) Hyperparameter Performance: Risk associated with hyperparameter combinations, measured by right-tail pinball loss.