

## Appendix: Concept drift Shallow-LLP:

Hadoop:

The Figures below show the AUC performance of Hadoop's Shallow-LLPs across time for time frame sizes from 6 up to 24 months. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

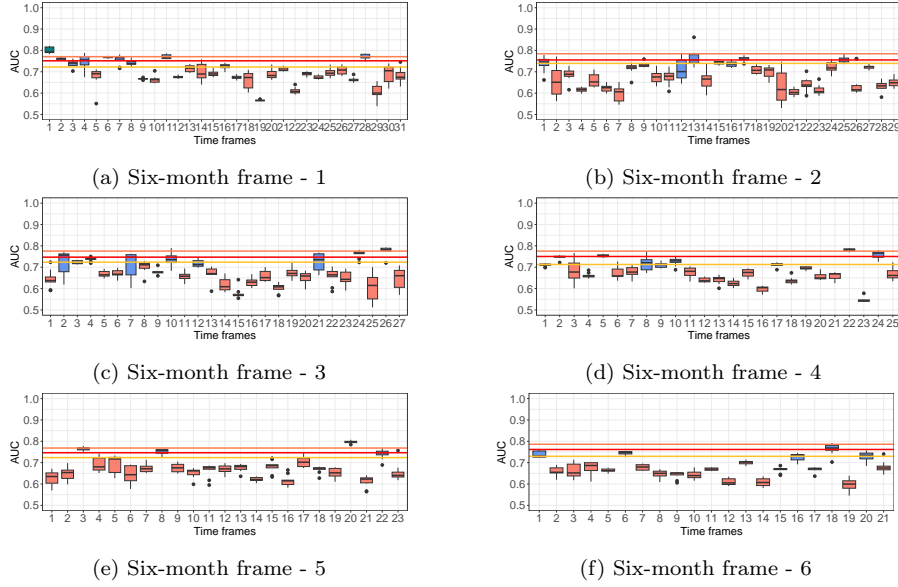


Fig. 1: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

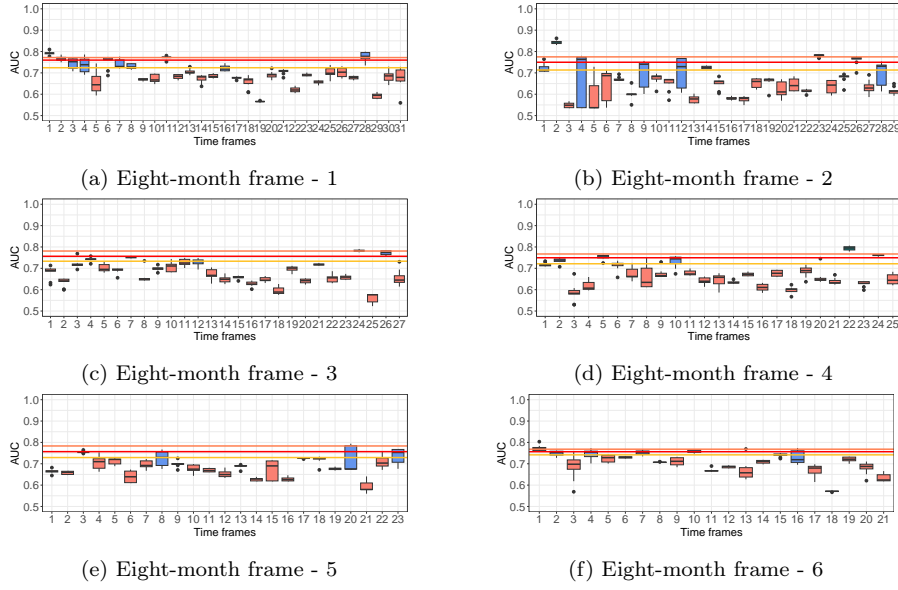


Fig. 2: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

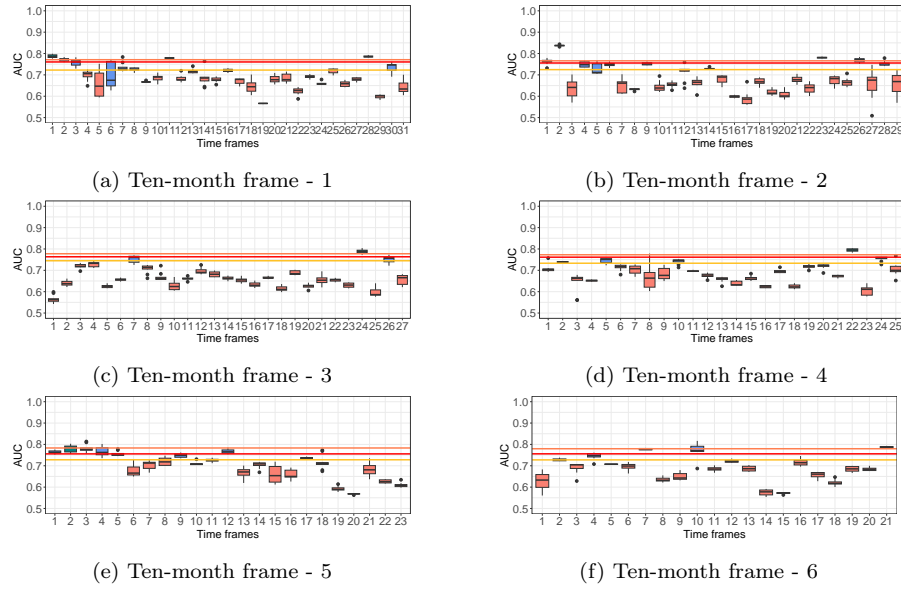


Fig. 3: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

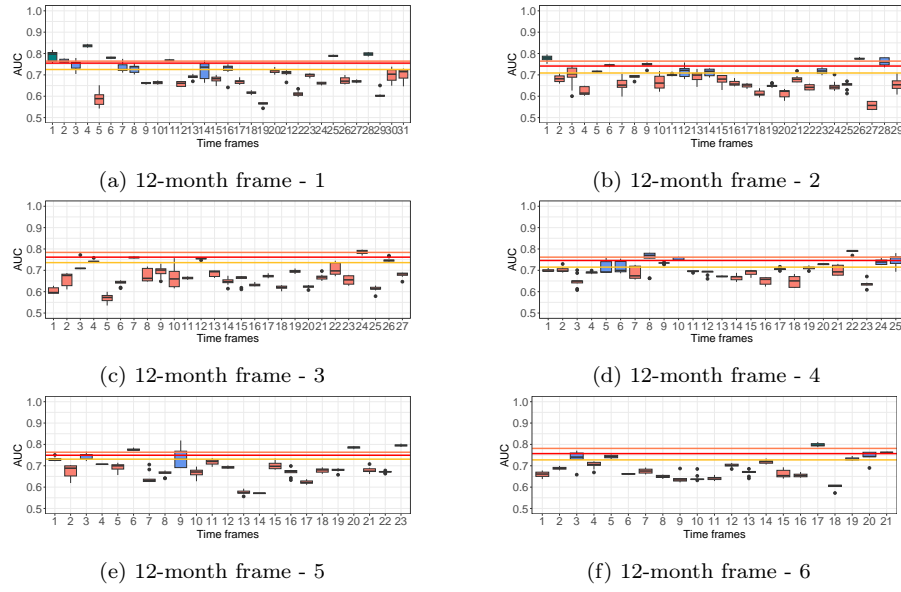


Fig. 4: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

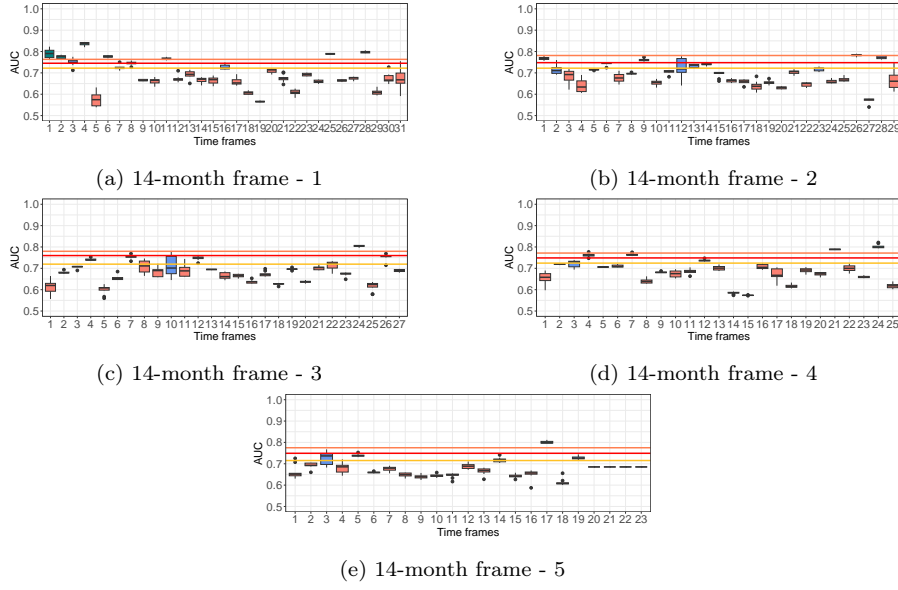


Fig. 5: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

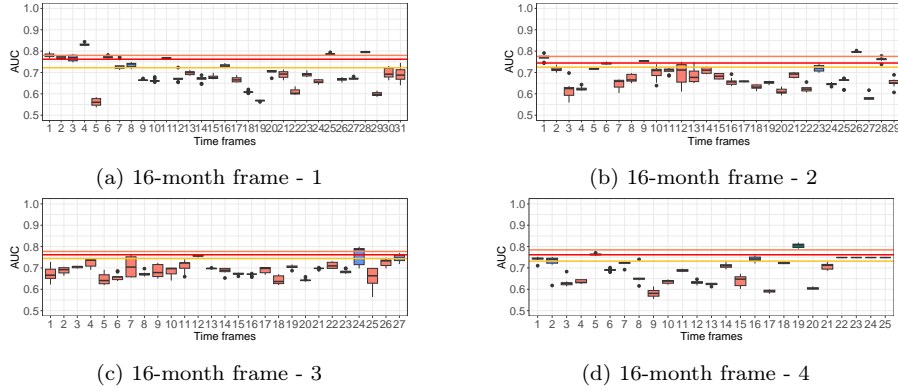


Fig. 6: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

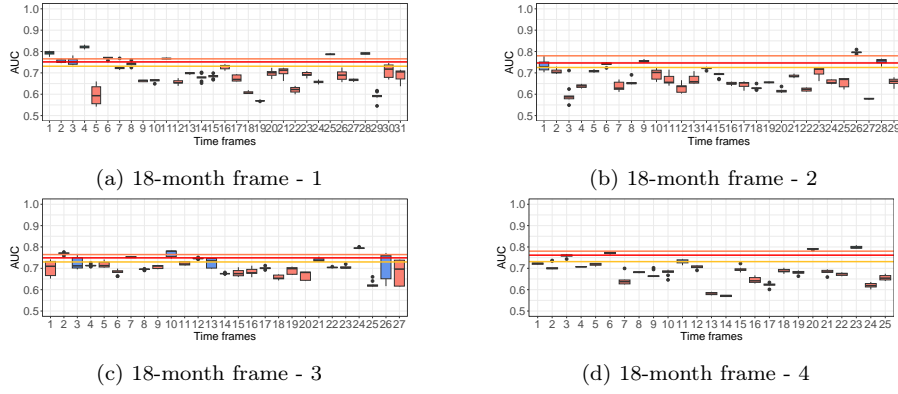


Fig. 7: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

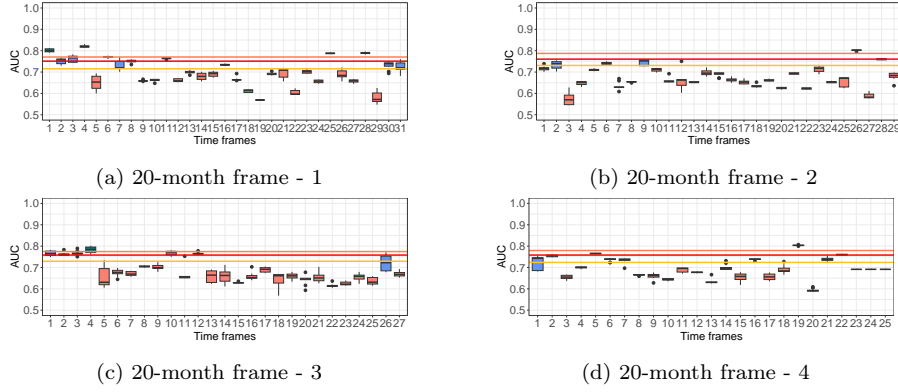


Fig. 8: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

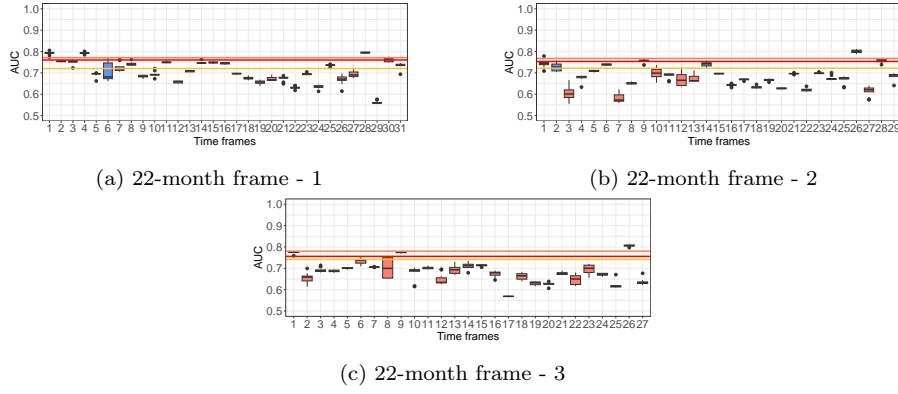


Fig. 9: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.

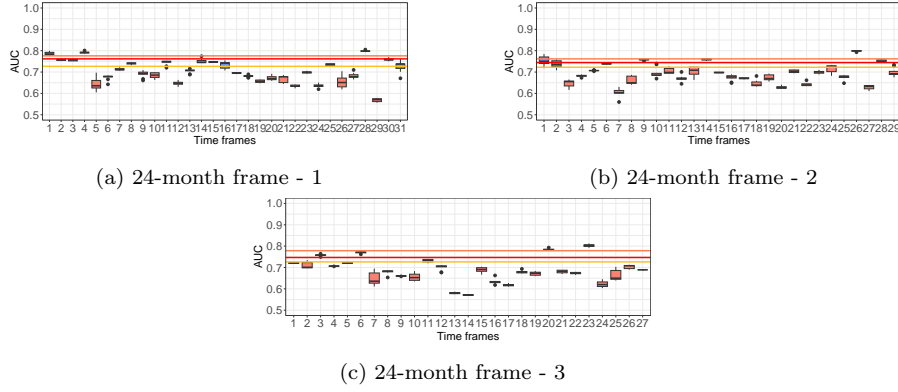


Fig. 10: AUC performance of Hadoop's models across time. The lines represents the 1st quantile, median and 3rd Quantile for the baseline AUC performance. Red boxplots show time frames with a performance statistically below the baseline, blue boxplots show no statistical difference and green boxplots show time frames with a performance statistically better than the baseline.