**Table S3.** Wilcoxon signed-rank test comparing predicted error between the two modeling approaches. Absolute error was calculated and compared for each prediction of observed peak of infectious mosquitoes, maximum mosquito infection rate, and the total number of human cases over the entire season, whereas root mean squared error (RMSE) was used to calculate future forecasts of the total number of mosquitoes observed over the season. 1 indicates the reported daily observed temperature-forced model forecasts had statistically significantly less error than the baseline model and -1 indicates the baseline model forecasts had statistically significant less error.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Forecast Week | Human Cases | Peak Timing | Peak Magnitude | Season Forecast |
| 21 | -1\*\*\* | 1\*\* | -1\*\*\* | -1\*\*\* |
| 22 | -1\*\*\* | 1\*\*\* | -1\*\*\* | -1\*\*\* |
| 23 | -1\*\*\* | 1\* | -1\*\*\* | -1\*\*\* |
| 24 | -1\*\*\* | 0 | -1\*\*\* | -1\*\*\* |
| 25 | -1\*\*\* | 0 | -1\*\*\* | -1\*\*\* |
| 26 | -1\*\*\* | 1\*\*\* | -1\*\*\* | -1\*\*\* |
| 27 | -1\*\*\* | 1\* | -1\*\*\* | -1\*\*\* |
| 28 | -1\*\*\* | -1\* | -1\*\*\* | -1\*\*\* |
| 29 | -1\*\*\* | -1\* | -1\*\*\* | -1\*\*\* |
| 30 | -1\* | 0 | -1\*\*\* | -1\*\*\* |
| 31 | 0 | 1\*\*\* | -1\*\*\* | -1\*\*\* |
| 32 | 1\*\*\* | 1\*\*\* | 0 | -1\*\*\* |
| 33 | 0 | 1\*\*\* | 0 | -1\*\*\* |
| 34 | 1\*\*\* | 1\*\*\* | 1\* | -1\*\*\* |
| 35 | 1\*\*\* | 1\*\*\* | 1\*\*\* | -1\*\*\* |
| 36 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 0 |
| 37 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\* |
| 38 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 39 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 40 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 41 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 42 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 43 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 44 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 45 | 1\*\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 46 | 1\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 47 | 0 | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 48 | 0 | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 49 | 0 | 1\*\*\* | 1\*\*\* | 1\*\*\* |
| 50 | 0 | 1\*\*\* | 1\*\*\* | 0 |
| 51 | -1\*\* | 1\*\*\* | 1\*\*\* | 1\*\*\* |

Asterisks designate differences significant at p<0.05 (\*), p<0.01 (\*\*) and p<0.001 (\*\*\*).