## S1 Results of verification experiments on individual events

For the analysis, we have selected 11 events during the summer periods of 2016 and 2017. These events are selected for covering a range of event characteristics with different rainfall intensity, spatial coverage, and duration. Table S1 shows the analyzed events as provided in Ayzel et al. (2019).

**Table S1.** Characteristics of the selected events

| Event # | Start | End | Duration, hours | Maximum extent, $km^2$ | Extent >1 mm $h^{-1}$, % |
|---|---|---|---|---|---|
| Event 1 | 2016-05-23 2:00 | 2016-05-23 8:00 | 6 | 159318 | 42 |
| Event 2 | 2016-05-23 13:00 | 2016-05-24 2:30 | 13.5 | 135272 | 56 |
| Event 3 | 2016-05-29 12:05 | 2016-05-29 23:55 | 12 | 160095 | 72 |
| Event 4 | 2016-06-12 7:00 | 2016-06-12 19:00 | 12 | 150416 | 53 |
| Event 5 | 2016-07-13 17:30 | 2016-07-14 1:00 | 7.5 | 145501 | 62 |
| Event 6 | 2016-08-04 18:00 | 2016-08-05 7:00 | 13 | 168407 | 74 |
| Event 7 | 2017-06-29 3:00 | 2017-06-29 5:05 | 2 | 140021 | 70 |
| Event 8 | 2017-06-29 17:00 | 2017-06-29 21:00 | 4 | 182561 | 60 |
| Event 9 | 2017-06-29 22:00 | 2017-06-30 21:00 | 23 | 160822 | 75 |
| Event 10 | 2017-07-21 19:00 | 2017-07-21 23:00 | 4 | 63698 | 77 |
| Event 11 | 2017-07-24 8:00 | 2017-07-25 23:55 | 16 | 253666 | 63 |

We use three metrics for model verification: mean absolute error (MAE), critical success index (CSI), and fractions skill score (FSS). We have applied threshold rain rates of 0.125, 1, and 5 mm $h^{-1}$ for calculating CSI and FSS. Additionally, for calculating FSS we use neighborhood (window) sizes of 1, 5, 10, and 20 km. Table S2 shows the numbers of corresponding figures which represent verification results averaged for the particular event and metric.

**Table S2.** Reference to the figure numbers for analysed events and verification metrics

| Event # | MAE and CSI | FSS |
|---|---|---|
| Event 1 | S1 | S12 |
| Event 2 | S2 | S13 |
| Event 3 | S3 | S14 |
| Event 4 | S4 | S15 |
| Event 5 | S5 | S16 |
| Event 6 | S6 | S17 |
| Event 7 | S7 | S18 |
| Event 8 | S8 | S19 |
| Event 9 | S9 | S20 |
| Event 10 | S10 | S21 |
| Event 11 | S11 | S22 |

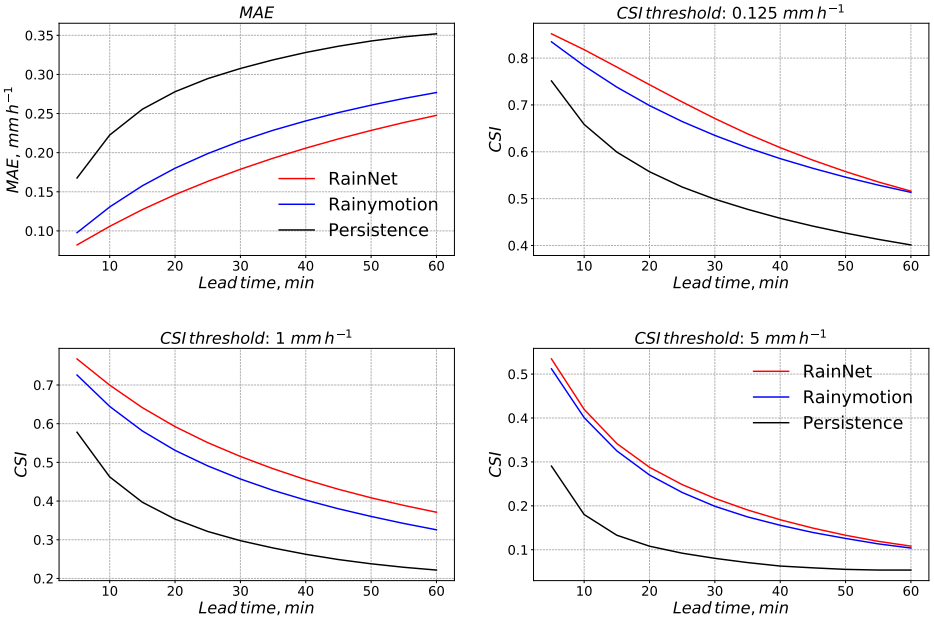


**Figure S1.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 1 (2016-05-23 2:00 – 2016-05-23 8:00)**

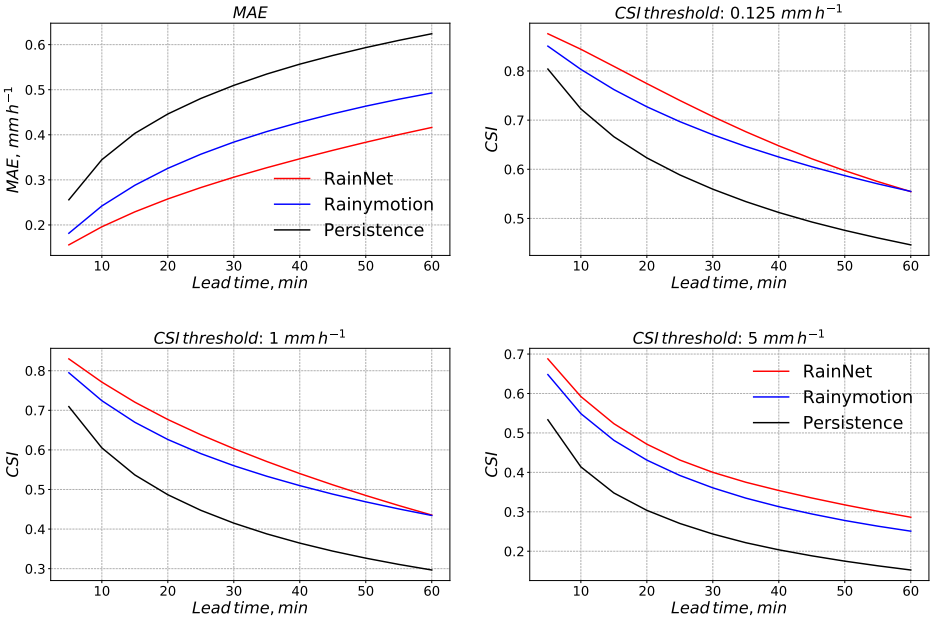


**Figure S2.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 2 (2016-05-23 13:00 – 2016-05-24 2:30)**

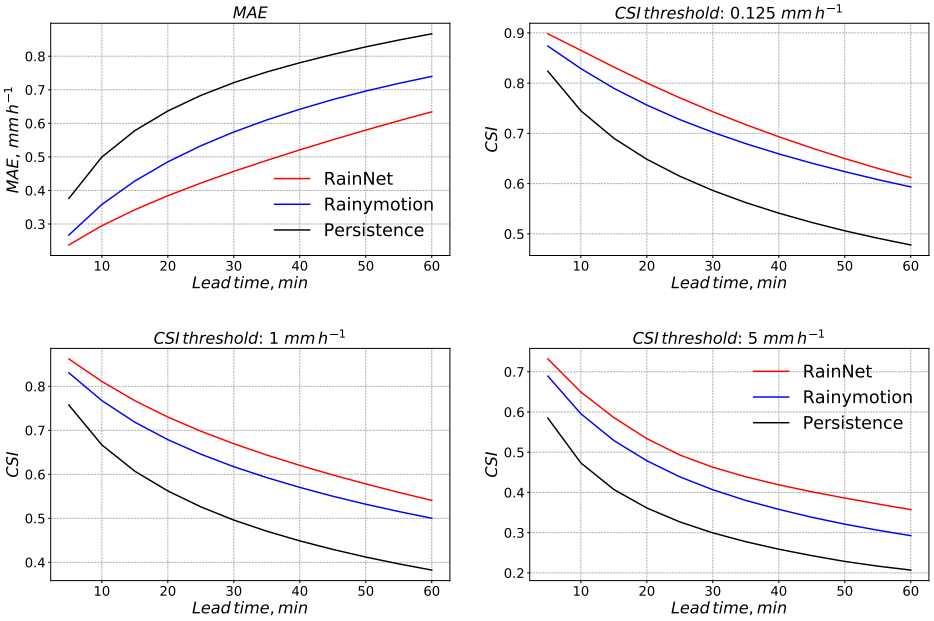


**Figure S3.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 3 (2016-05-29 12:05 – 2016-05-29 23:55)**

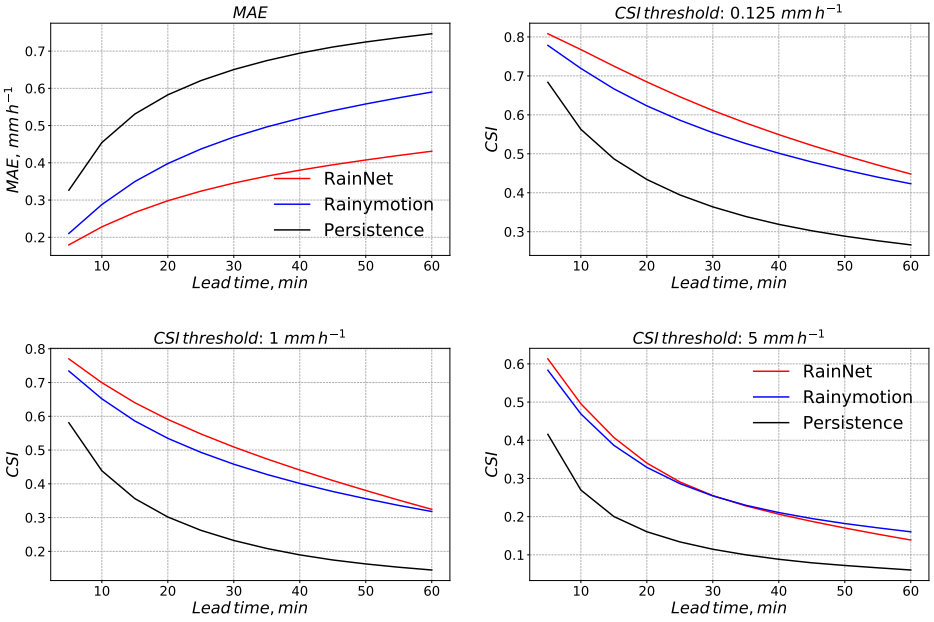


**Figure S4.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 4 (2016-06-12 7:00 – 2016-06-12 19:00)**

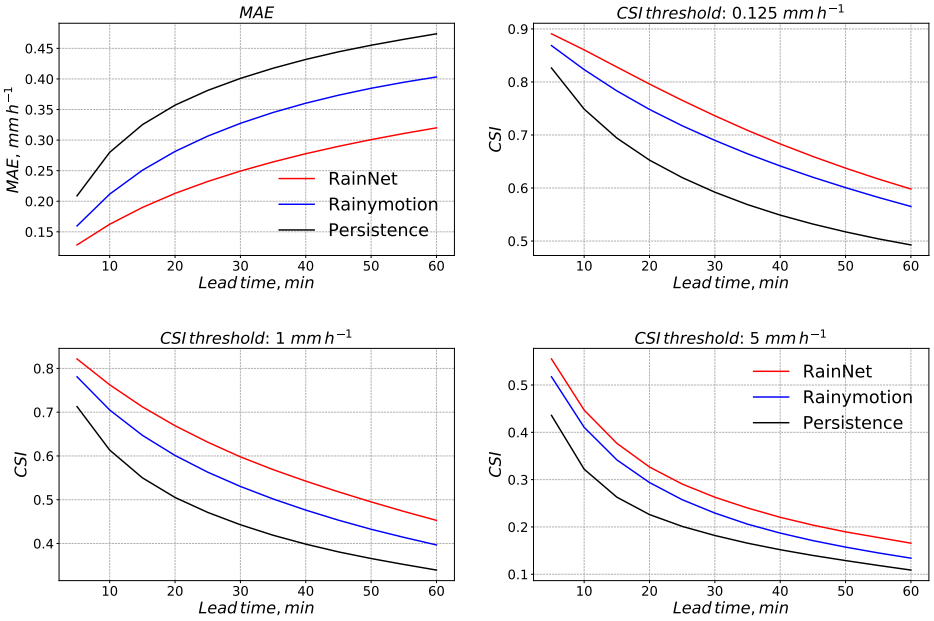


**Figure S5.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 5 (2016-07-13 17:30 – 2016-07-14 1:00)**

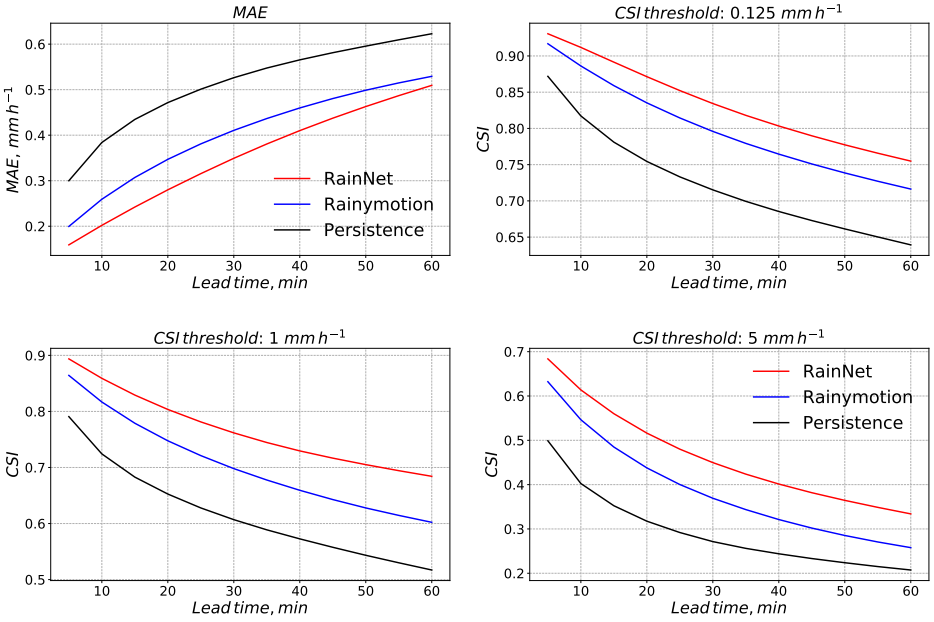


**Figure S6.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 6 (2016-08-04 18:00 – 2016-08-05 7:00)**

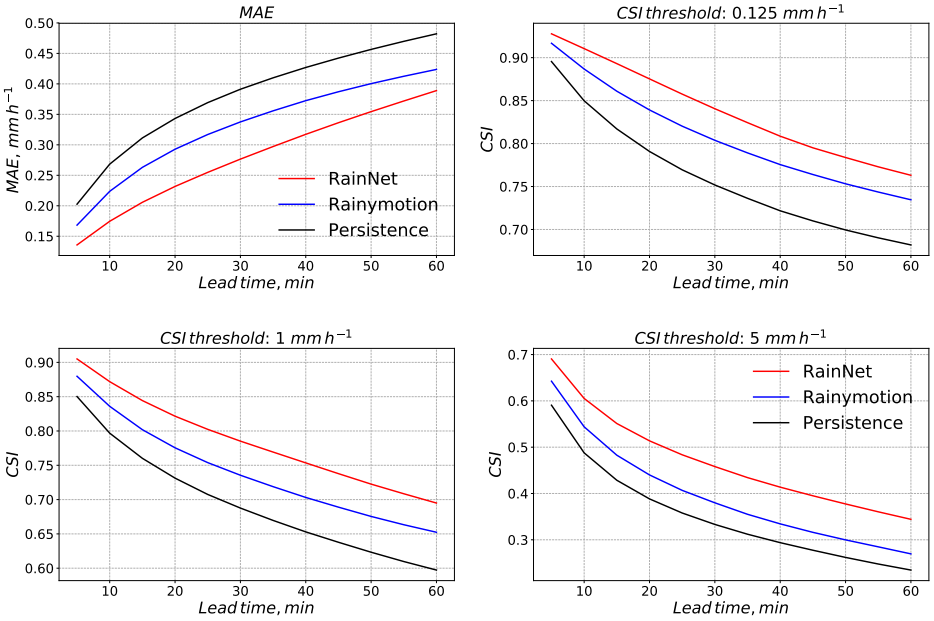


**Figure S7.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 7 (2017-06-29 3:00 – 2017-06-29 5:05)**

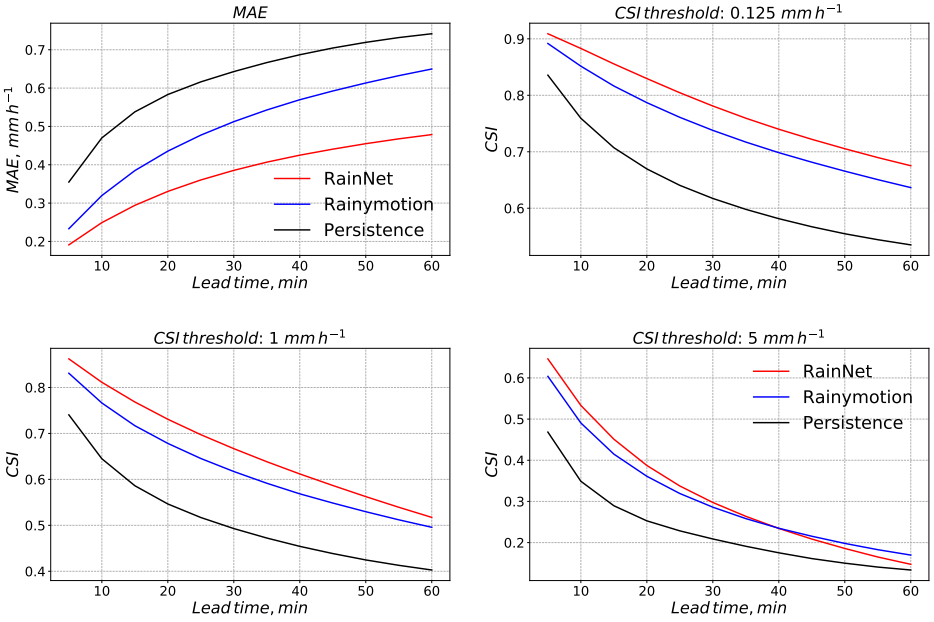


**Figure S8.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 8 (2017-06-29 17:00 – 2017-06-29 21:00)**

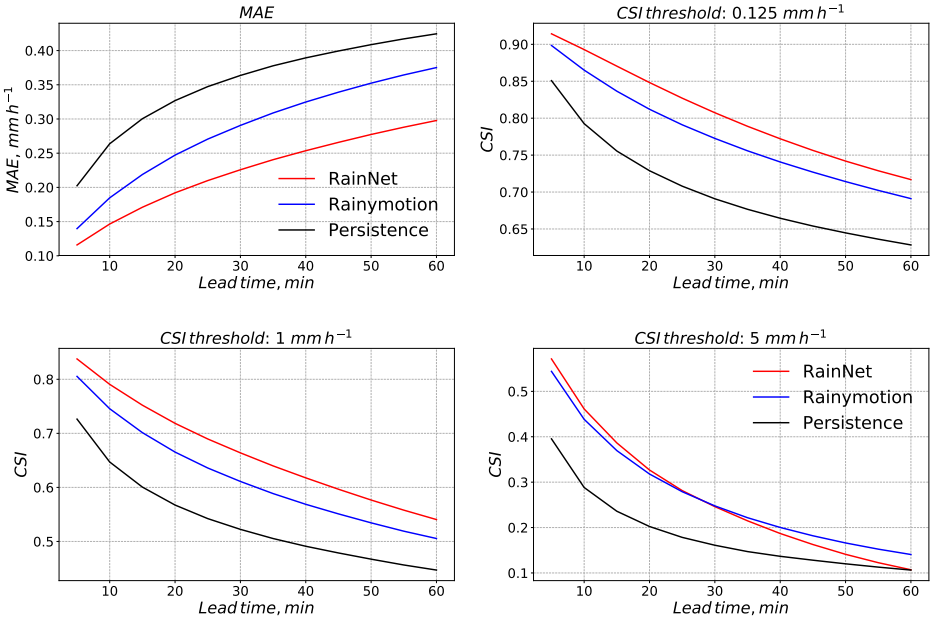


**Figure S9.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 9 (2017-06-29 22:00 – 2017-06-30 21:00)**

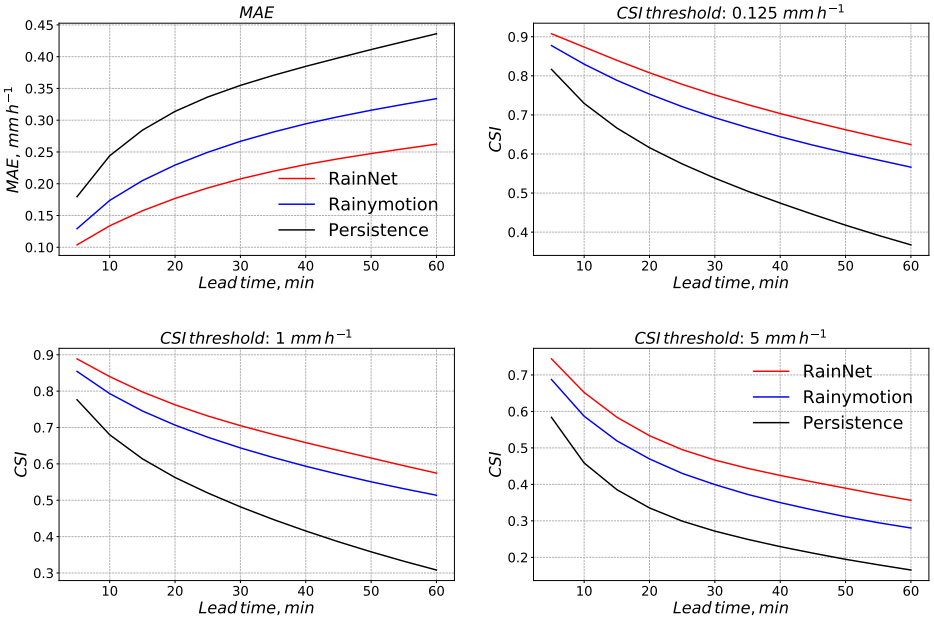


**Figure S10.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 10 (2017-07-21 19:00 – 2017-07-21 23:00)**

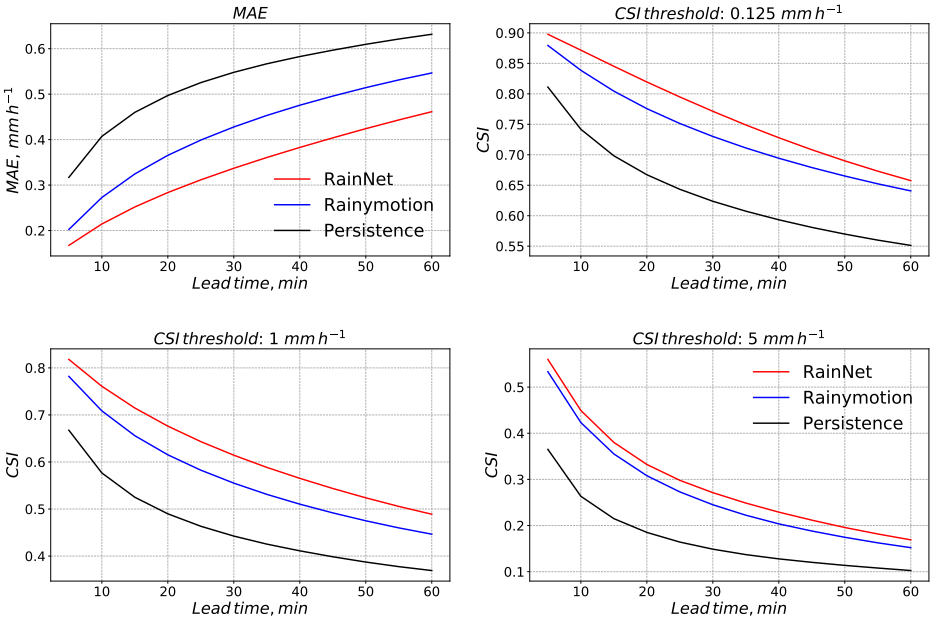


**Figure S11.** Mean Absolute Error (MAE) and Critical Success Index (CSI) for three different intensity thresholds (0.125 mm $h^{-1}$, 1 mm $h^{-1}$, 5 mm $h^{-1}$). The metrics are shown as a function of lead time. All values represent the average of the corresponding metric over **Event 11 (2017-07-24 8:00 – 2017-07-25 23:55)**

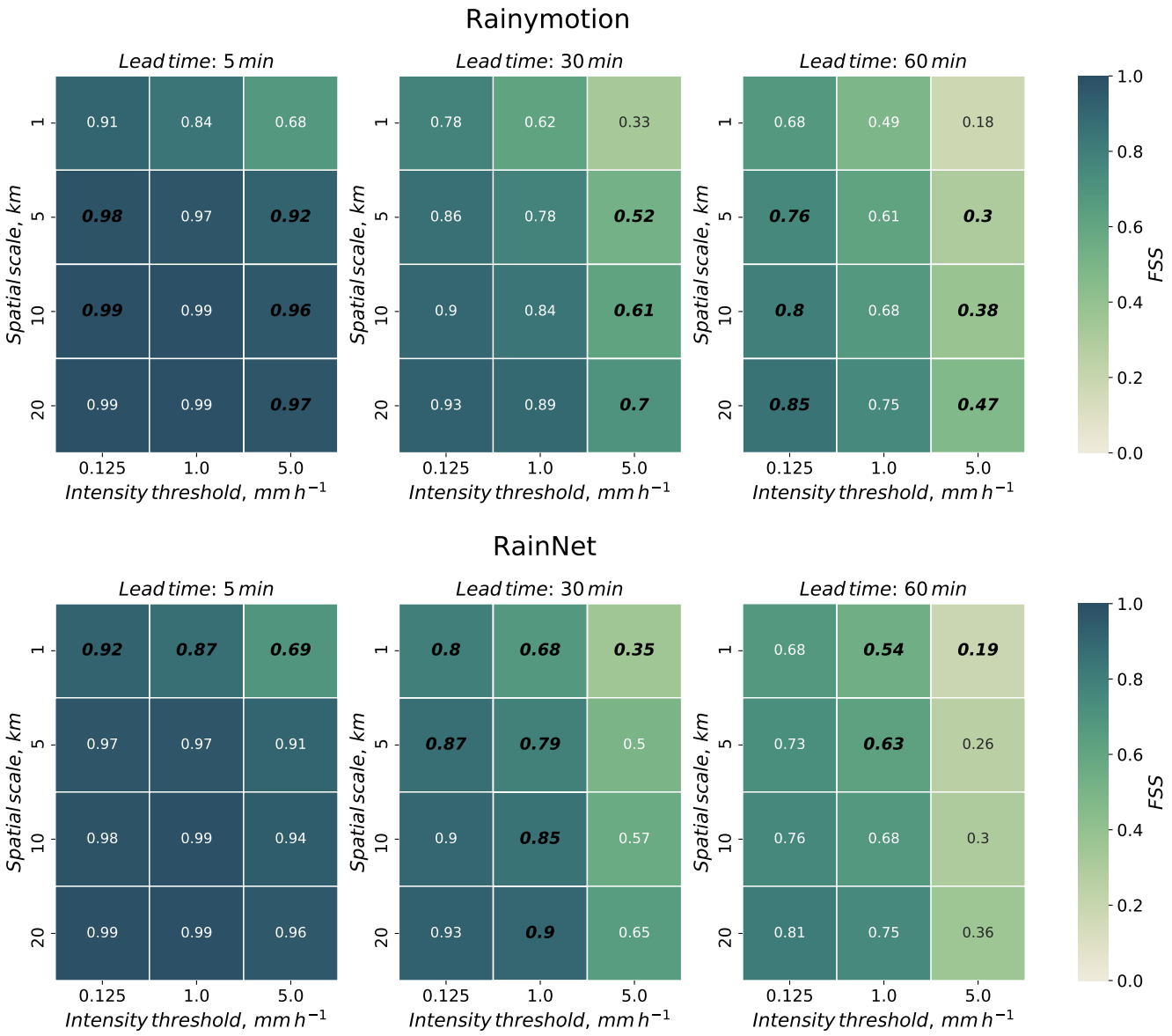


**Figure S12.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 1 (2016-05-23 2:00 – 2016-05-23 8:00)**

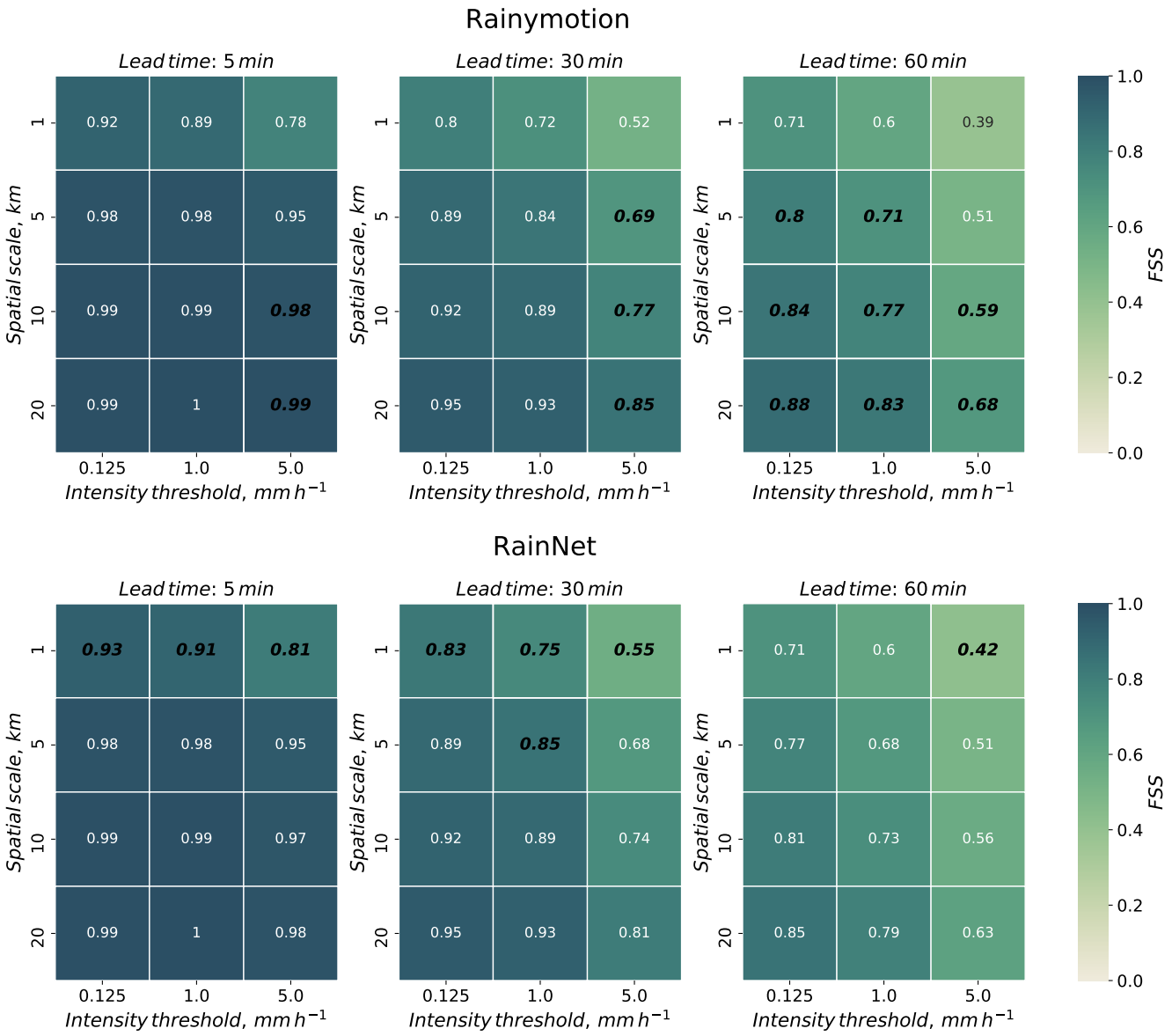


**Figure S13.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 2 (2016-05-23 13:00 – 2016-05-24 2:30)**

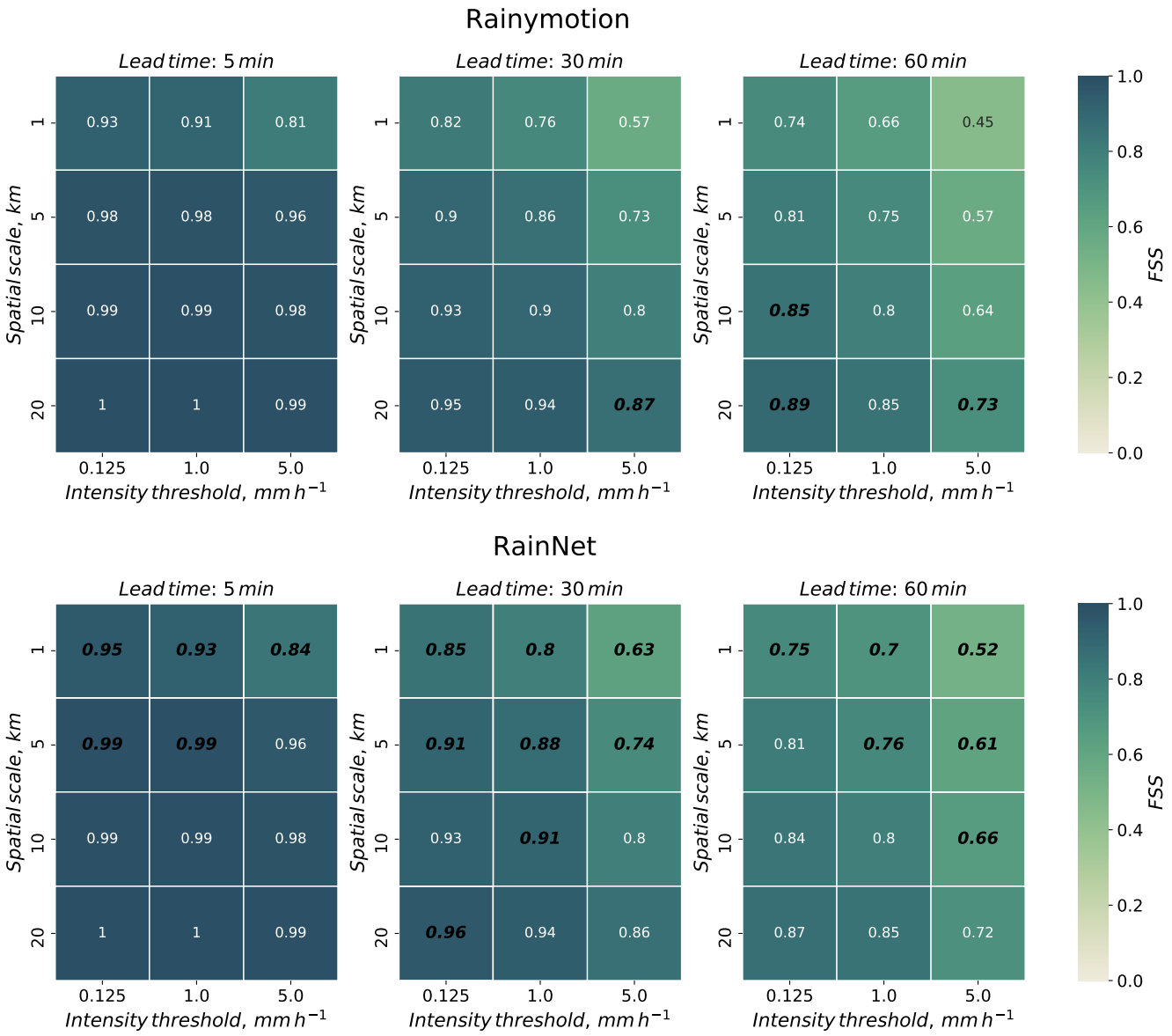


**Figure S14.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 3 (2016-05-29 12:05 – 2016-05-29 23:55)**

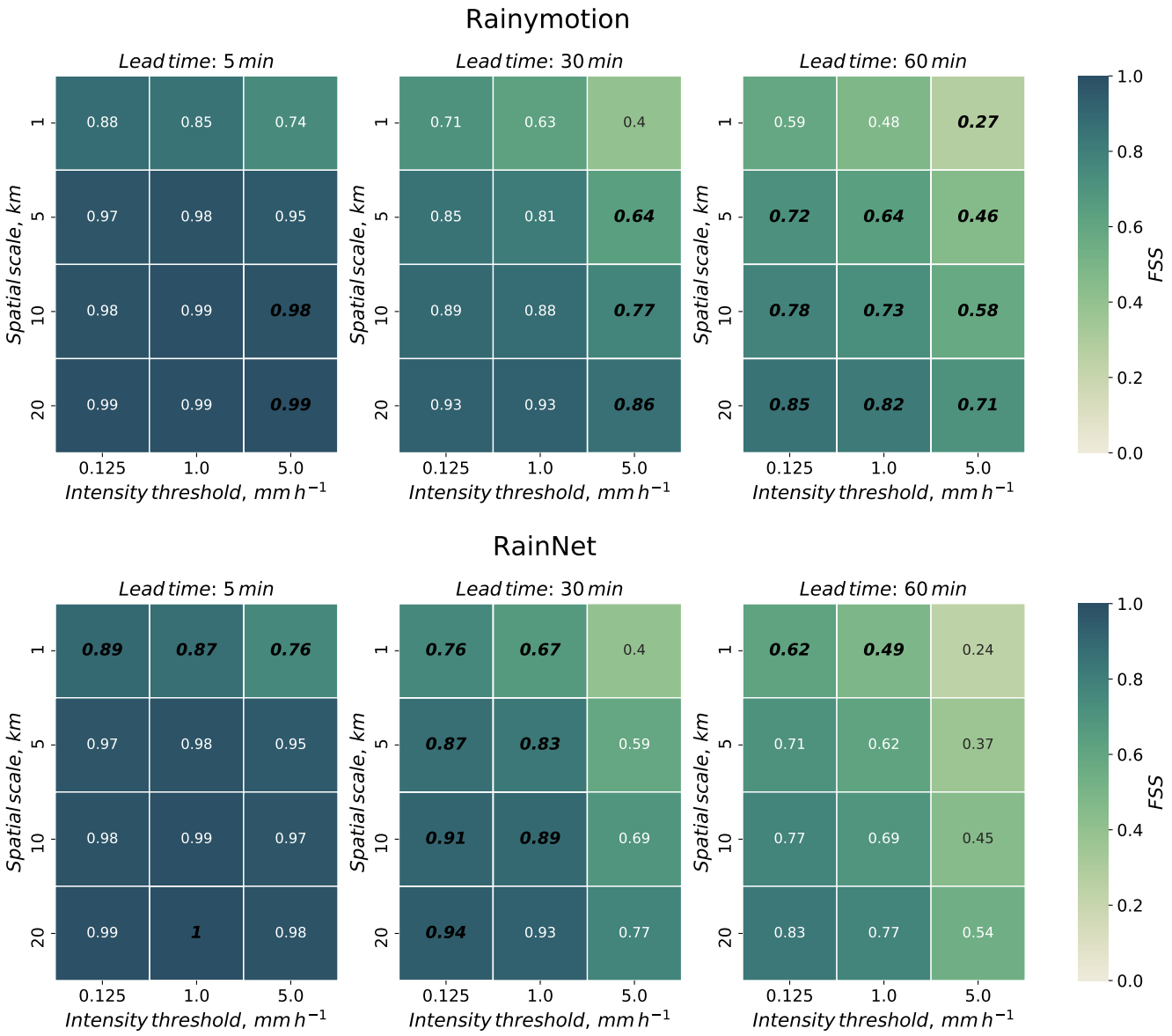


**Figure S15.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 4 (2016-06-12 7:00 – 2016-06-12 19:00)**

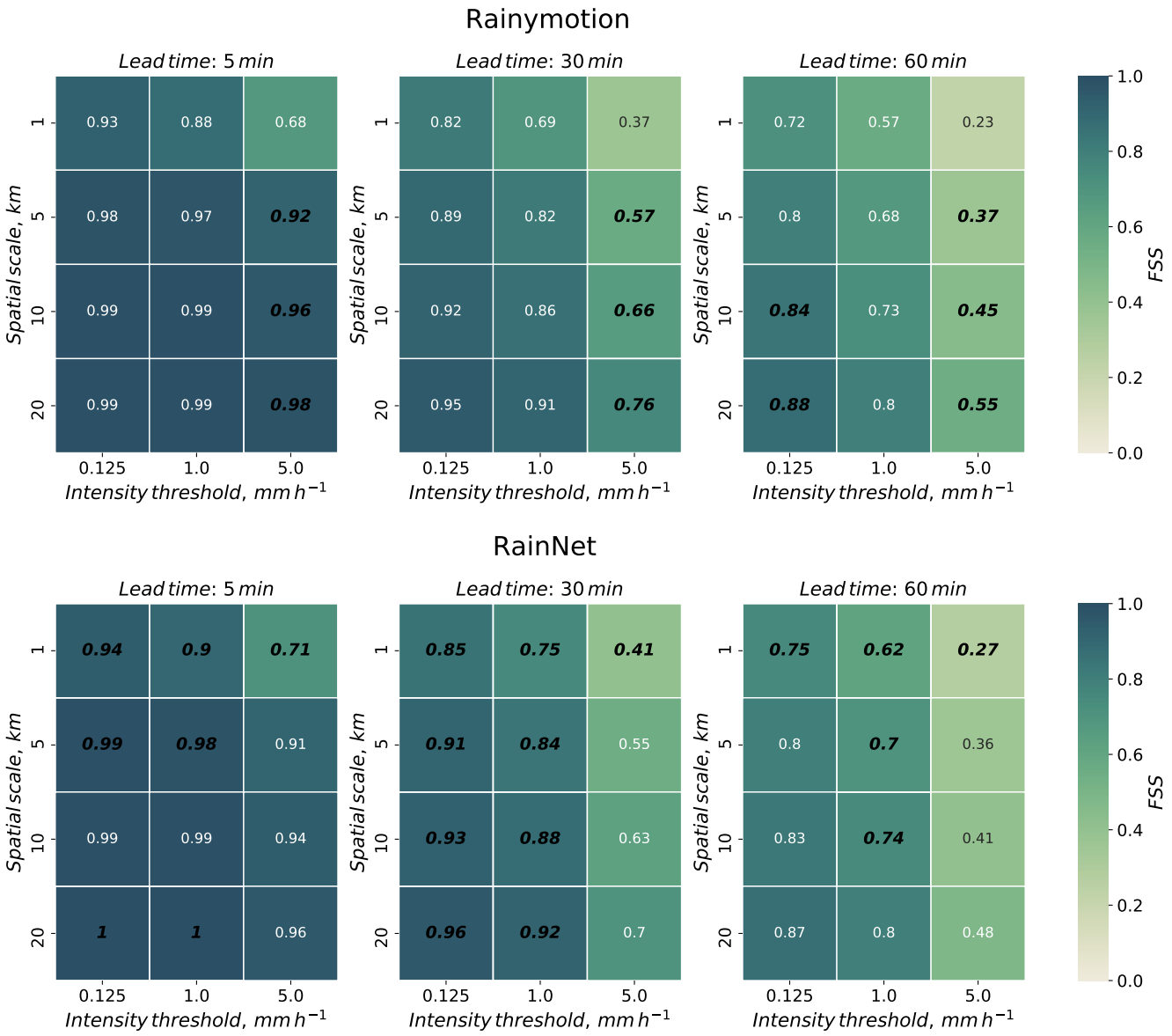


**Figure S16.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 5 (2016-07-13 17:30 – 2016-07-14 1:00)**

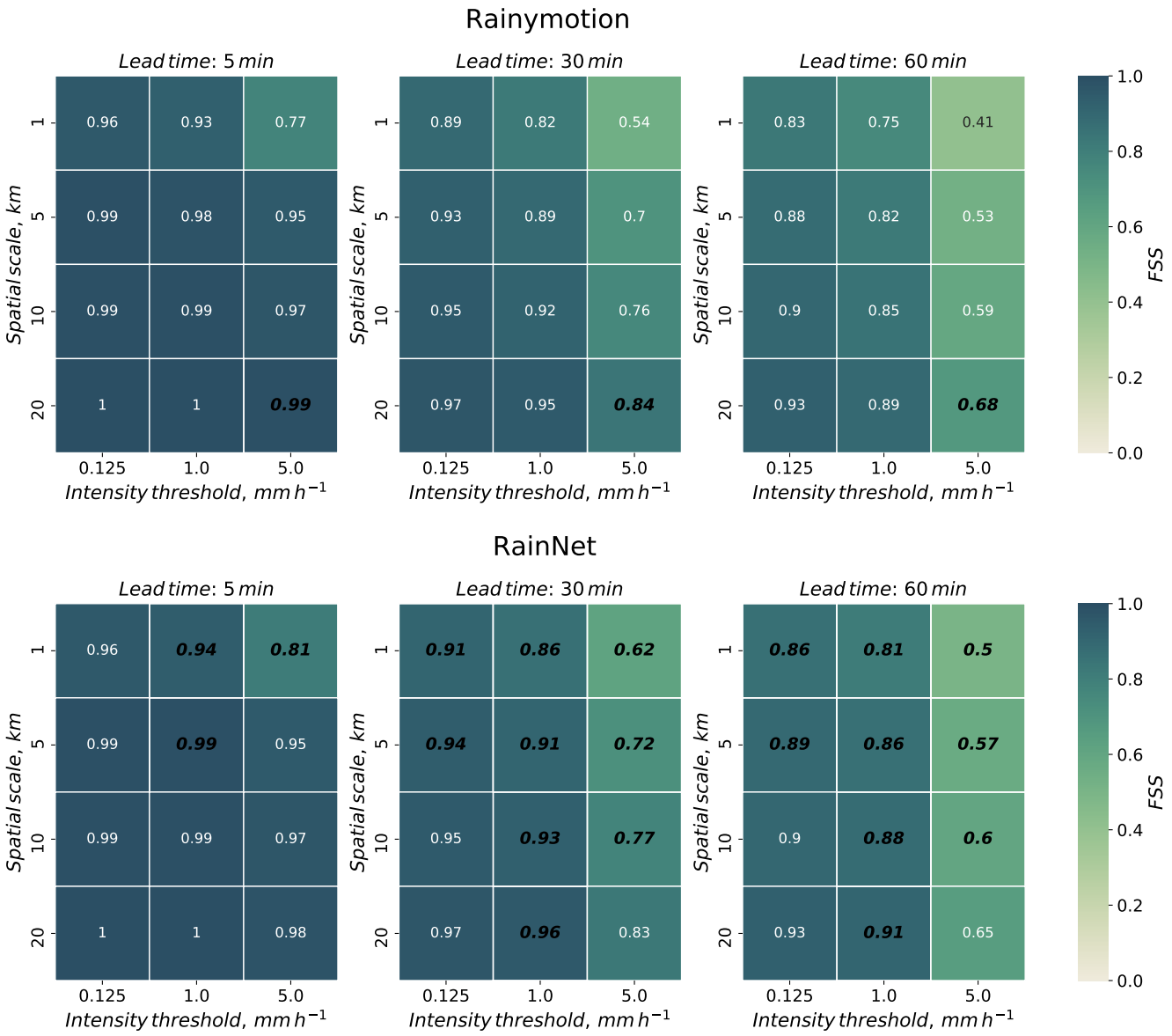


**Figure S17.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 6 (2016-08-04 18:00 – 2016-08-05 7:00)**

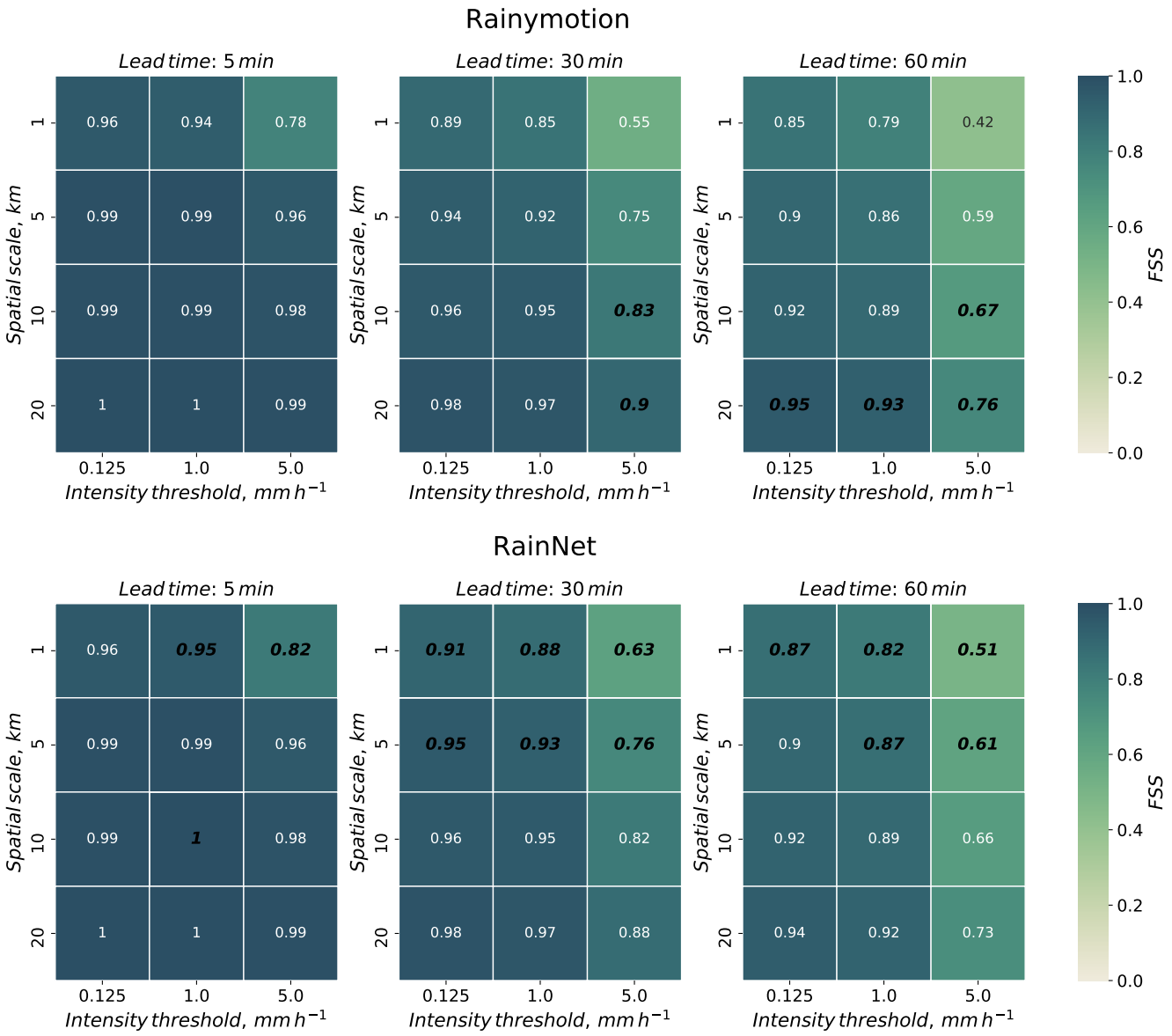


**Figure S18.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 7 (2017-06-29 3:00 – 2017-06-29 5:05)**

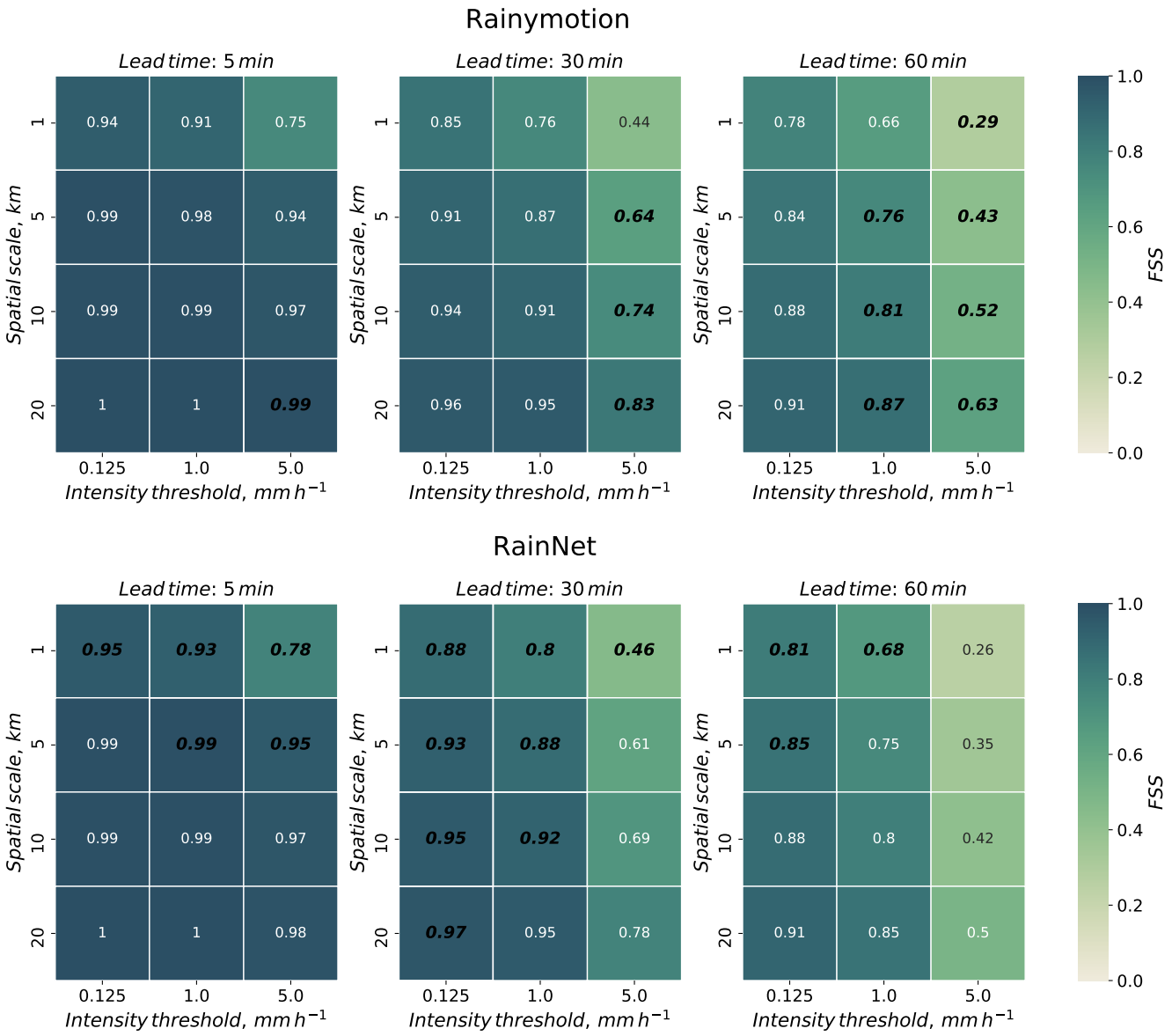


**Figure S19.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 8 (2017-06-29 17:00 – 2017-06-29 21:00)**

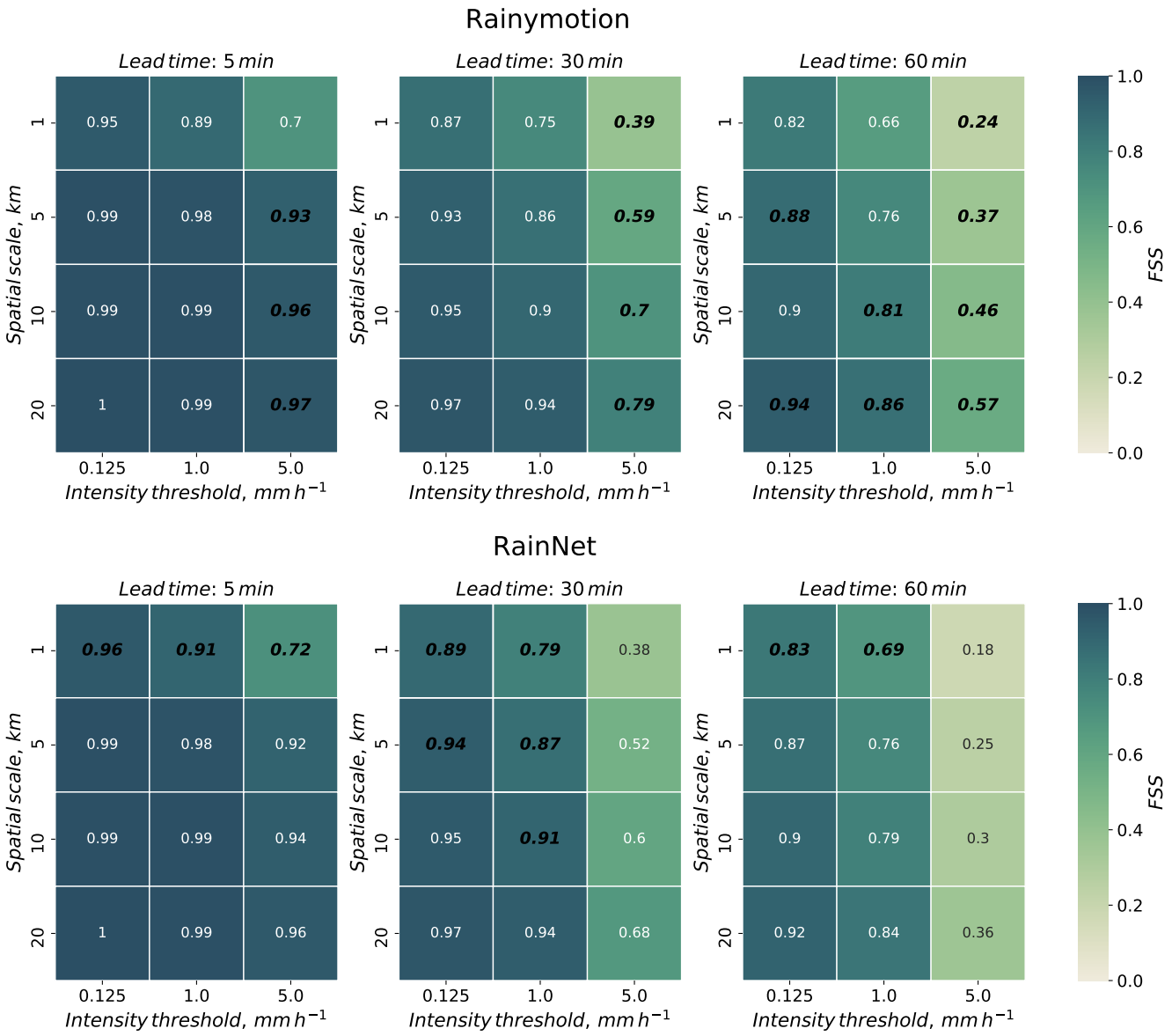


**Figure S20.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 9 (2017-06-29 22:00 – 2017-06-30 21:00)**

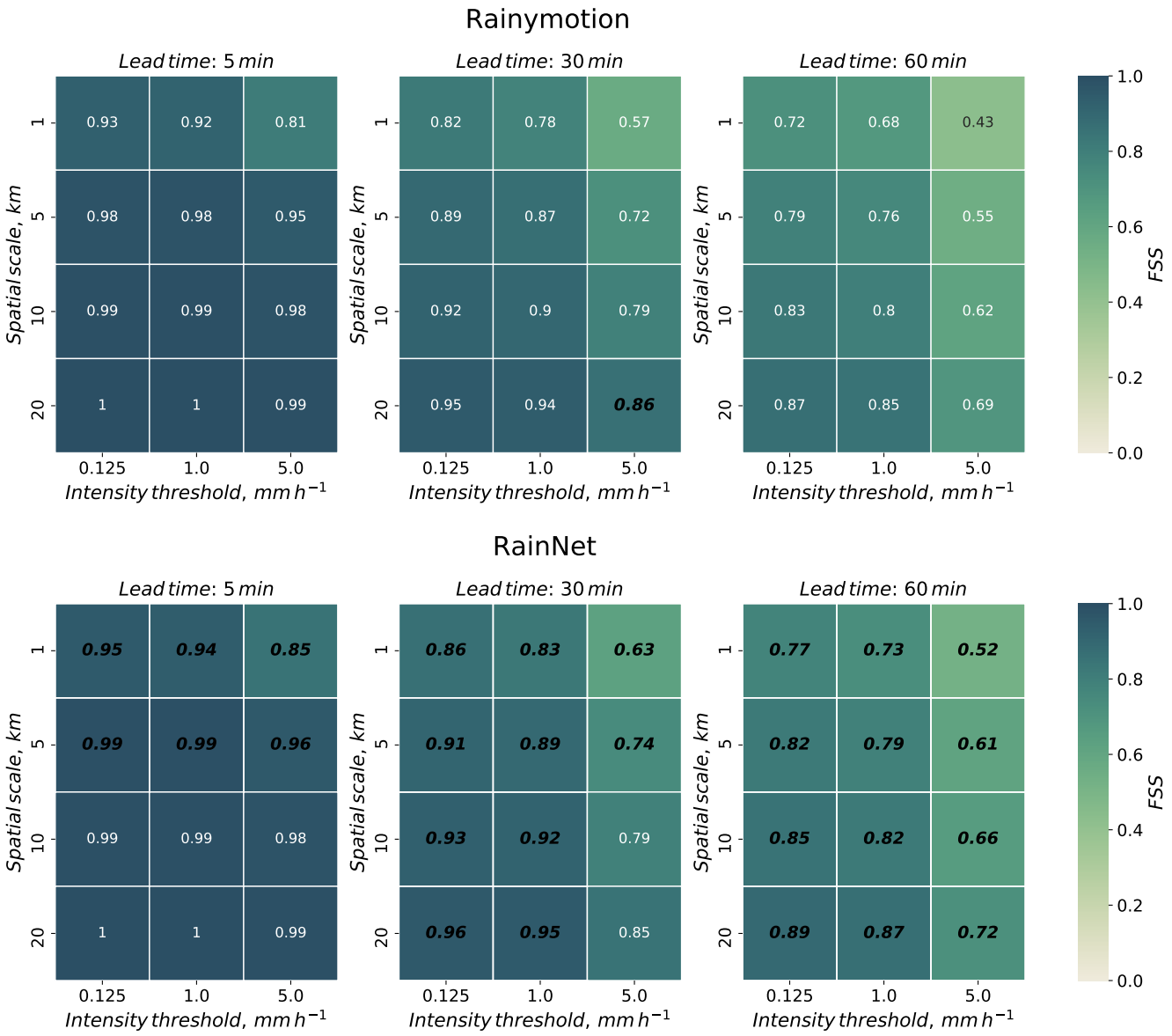


**Figure S21.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 10 (2017-07-21 19:00 – 2017-07-21 23:00)**

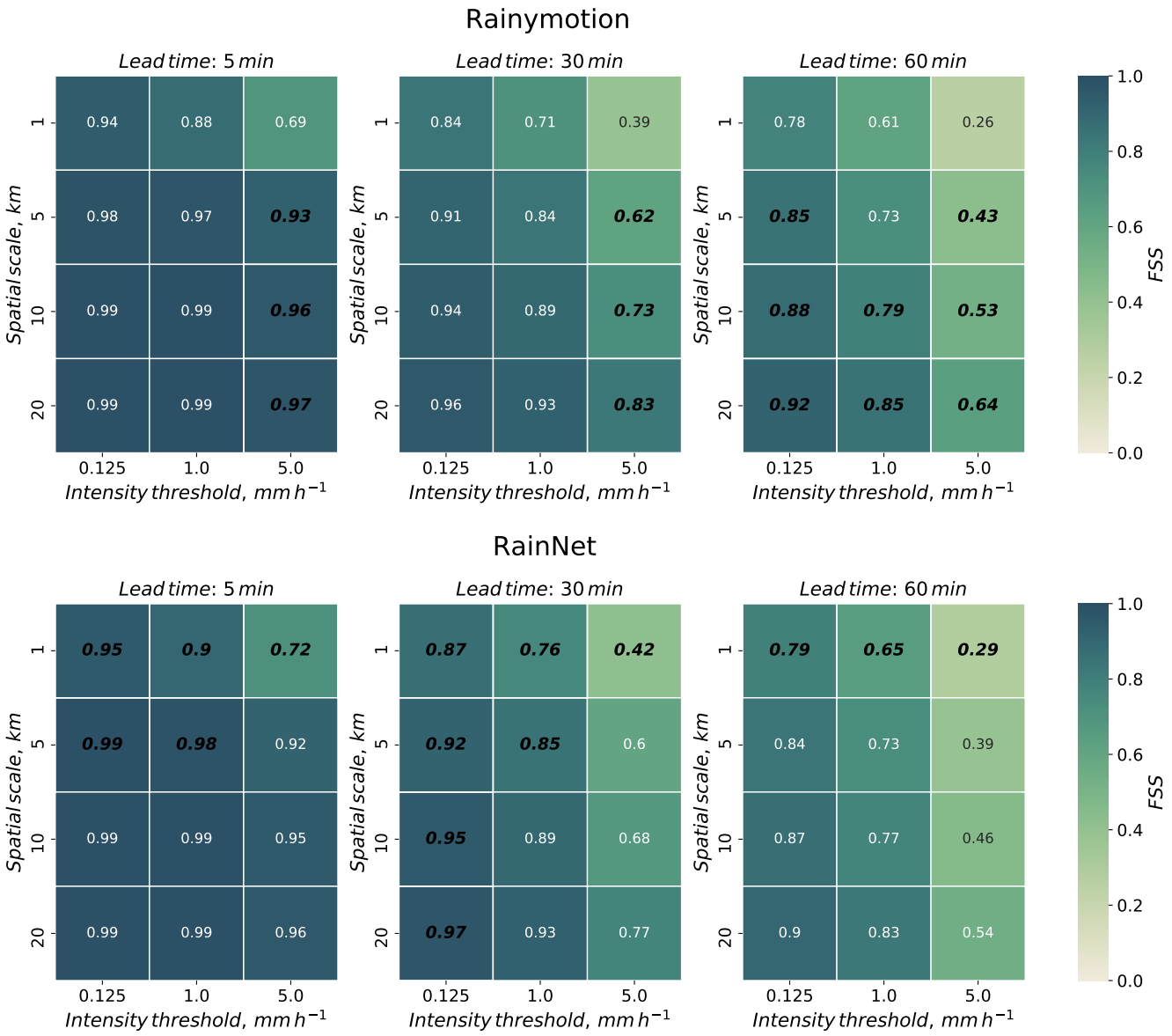


**Figure S22.** Fractions Skill Score (FSS) for Rainymotion (*top panel*) and RainNet (*bottom panel*), for 5, 30, and 60 minutes lead time, and spatial window sizes of 1, 5, 10 and 20 km, and for intensity thresholds of 0.125, 1, and 5 mm $h^{-1}$. In addition to the color code of the FSS, we added the numerical FSS values. The FSS value of the model which is significantly superior for a specific combination of window size, intensity threshold, and lead time is typed in bold black digits, for the inferior model in regular. All values represent the average of the FSS over **Event 11 (2017-07-24 8:00 – 2017-07-25 23:55)**

## References

Ayzel, G., Heistermann, M., and Winterrath, T.: Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1), Geoscientific Model Development, 12, 1387–1402, https://doi.org/10.5194/gmd-12-1387-2019, https://www.geosci-model-dev.net/12/1387/2019/, 2019.