Supplementary materials for "Dynamic Modeling of Power Outages Caused by Thunderstorms"

Berk Alpay[1], David Wanik[2], Peter Watson[3], Diego Cerrai[3], Guannan Liang[1], and Emmanouil Anagnostou[3]

[1]University of Connecticut, Department of Computer Science and Engineering
[2]University of Connecticut, Department of Operations and Information Management
[3]University of Connecticut, Department of Civil and Environmental Engineering

1 Description of the LSTM model

Recurrent neural networks (RNN) are designed to optimize dynamically interconnected, non-linear parameters [1]. We use this property to capture temporal relationships in chaotic atmospheric contexts. The viability of such a model in atmospheric science was demonstrated in the area of sequential weather forecasting [2]. The RNN is used with long short-term memory (LSTM) structure to adjust for lags between severe weather and outage occurrences, as LSTMs were shown to be capable of learning long-term dependencies between features and response [3]. A mathematical description of the operation of an LSTM unit within a neural network now follows.

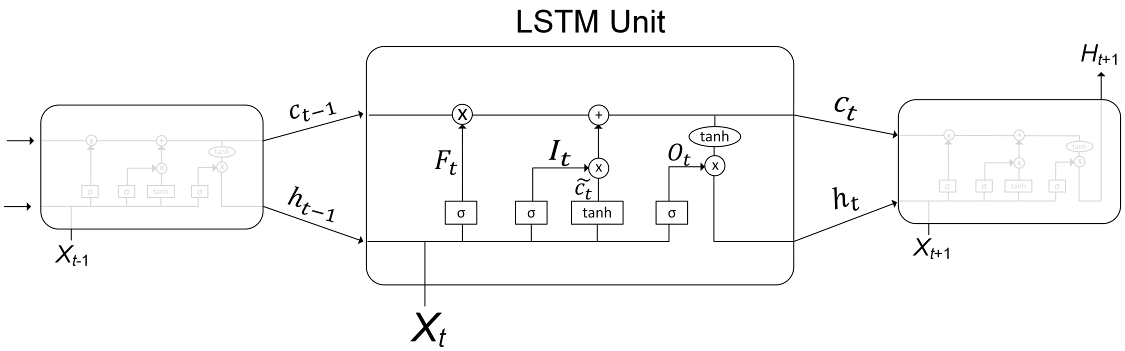

Figure 1: Unrolled architecture of an LSTM node in a neural network. The tth unit has been enlarged.

Let $x = [x_1, ..., x_t, ..., x_T]$ be a timeseries data sample, where x_t for $t \in [1, ..., T]$ represents the t-th observation in the series. In our application, $x_1, ..., x_T$ corresponds to each hour $[h - l, ..., h]$, where h is the current hour and l the lookback period.

From Figure 1, we see that each LSTM unit is composed of the following components: memory cell c_t; input gate i_t; forget gate f_t; and output gate o_t. The memory cell c_t is controlled by the other two gates i_t and f_t: each hour's weather input is written to memory cell c_t based on the activation of signal i_t, and the previous cell memory c_{t-1} will be "forgotten" or "cleared" based on activation of signal f_t. Whereas the upper channel of the illustrated LSTM unit propagates cell memory without much interference, the lower channel provides a short-term memory output through o_t based on the current cell state and a version of long-term memory that is filtered using a hyperbolic tangent transformation. The LSTM unit calculates hidden state h_t as follows:

$$
\begin{aligned}
i_t &= \sigma(x_t U^i + h_{t-1} W^i) \\
f_t &= \sigma(x_t U^f + h_{t-1} W^f) \\
o_t &= \sigma(x_t U^o + h_{t-1} W^o) \\
\tilde{c}_t &= \tanh(x_t U^g + h_{t-1} W^g) \\
c_t &= \sigma(f_t * c_{t-1} + i_t * \tilde{c}_t) \\
h_t &= \tanh(c_t) * o_t,
\end{aligned}
$$

where i_t, f_t, and o_t are the input gate activation vector, forget gate activation vector, and output gate activation vector, respectively. $U^i, W^i, U^f, W^f, U^o, W^o, U^g, W^g$ are weight matrices to be learned during training, $\sigma(x) = \frac{1}{1+e^{-x}}$, $\tanh(x) = \frac{e^{2x}-1}{e^{2x}+1}$, and $*$ is element-wise multiplication.

2 Supplementary Tables and Figures

Layer Type	**Input Shape**	**Activation Fnc.**	**Output Shape**
LSTM	(n, 10, 100)	tanh	(n, 128)
Fully connected	(n, 128)	ReLU	(n, 128)
Fully connected	(n, 128)	ReLU	(n, 64)
Fully connected	(n, 64)	ReLU	(n, 16)
Output	(n, 16)	ReLU	(n, 1)

Table 1: Architecture of the LSTM model, where n is the batch size. 100 input features consisting of PCA-reduced weather data and the cyclically encoded hour are input to an LSTM layer of 128 nodes. Fully connected layers then translate the LSTM layer's output into a forecast for the next hour's outages.

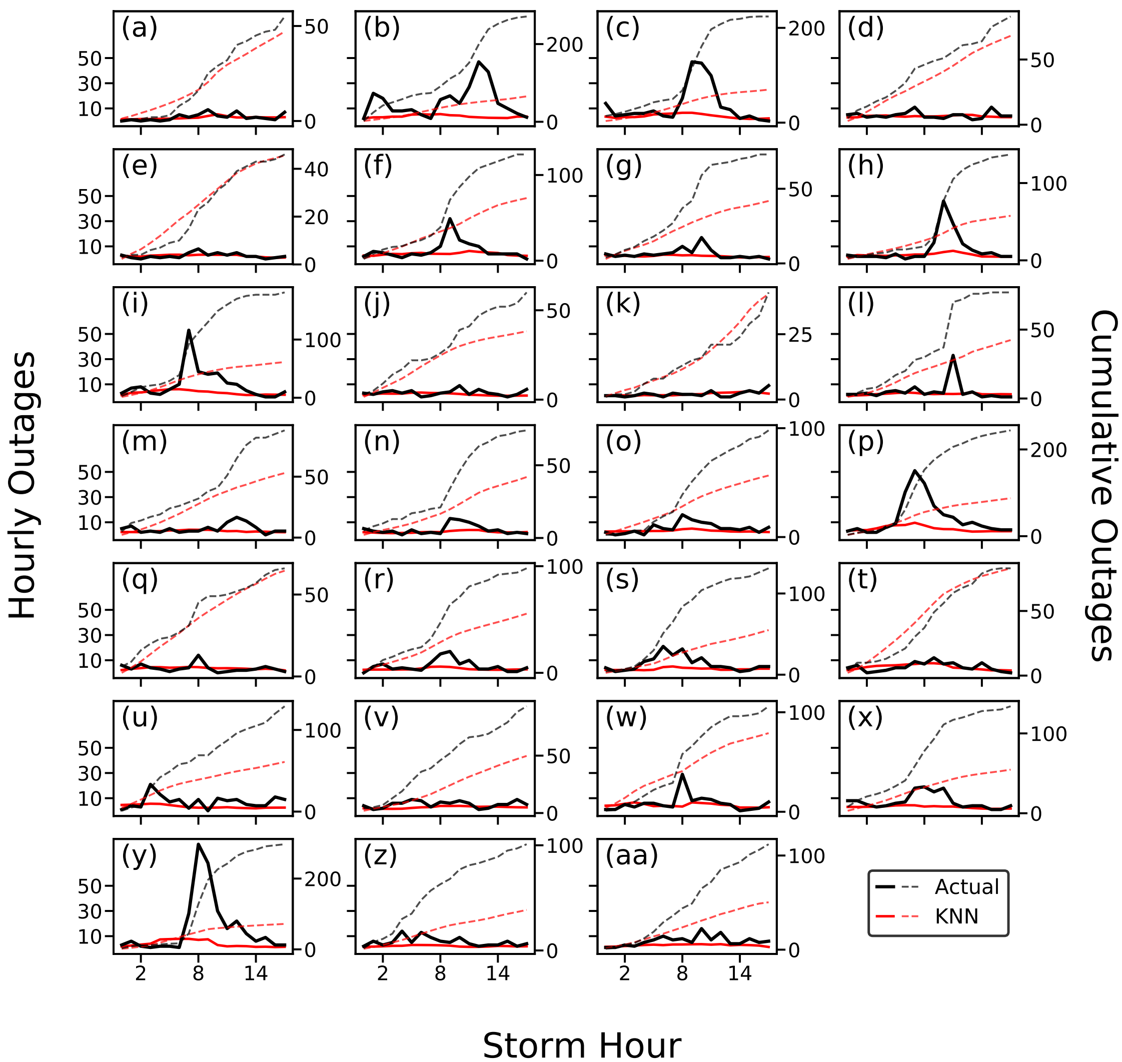

Figure 2: Hourly (solid) and cumulative (dashed) outages for 27 CT thunderstorms, each lasting 18 hours, with corresponding LOSO cross-validated predictions by KNN.

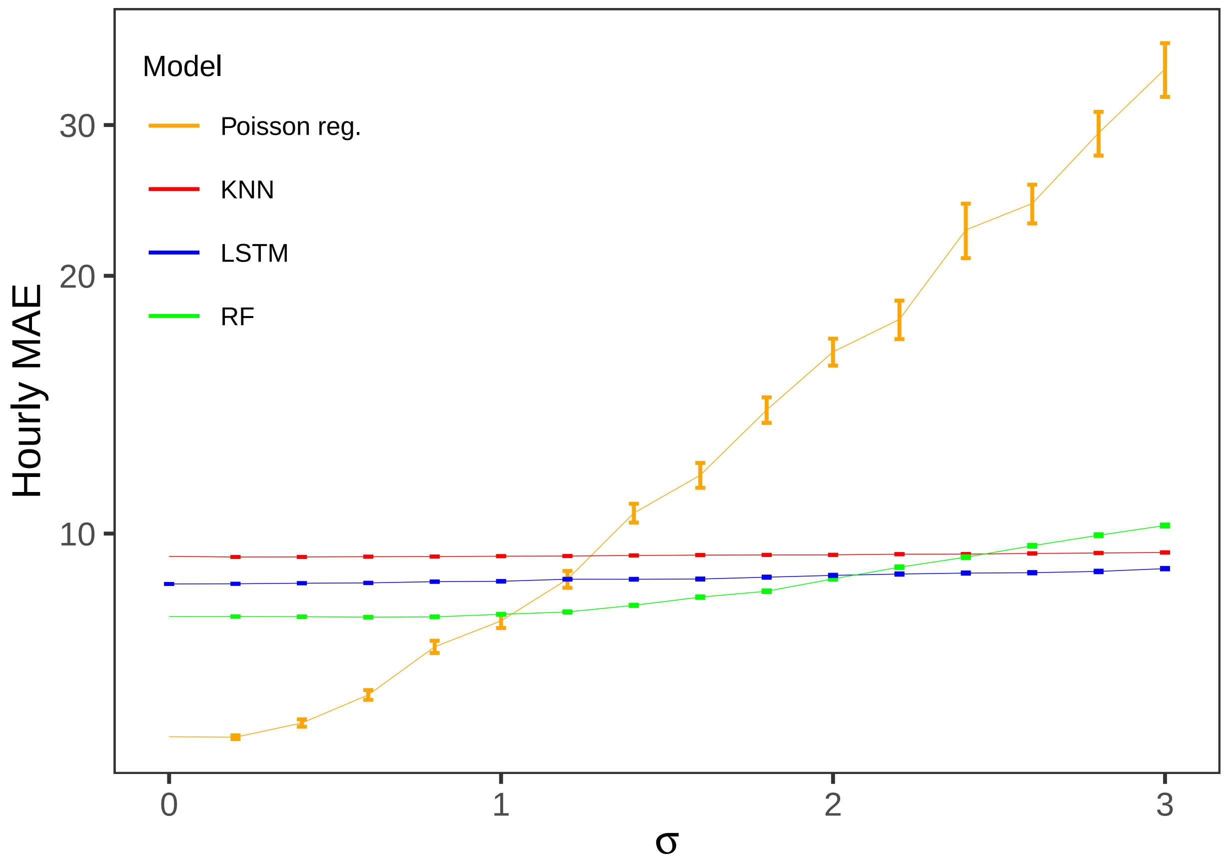

Figure 3: Simulation results for the hourly MAE in log-scale of each dynamic OPM as a function of the level of noise, distributed $\mathcal{N}(0, \sigma)$, in the PCA-scored forecast of a held out thunderstorm ("p") on July 17, 2018. Error bars represent the standard error of each estimate over 300 random samples of noise at each σ.

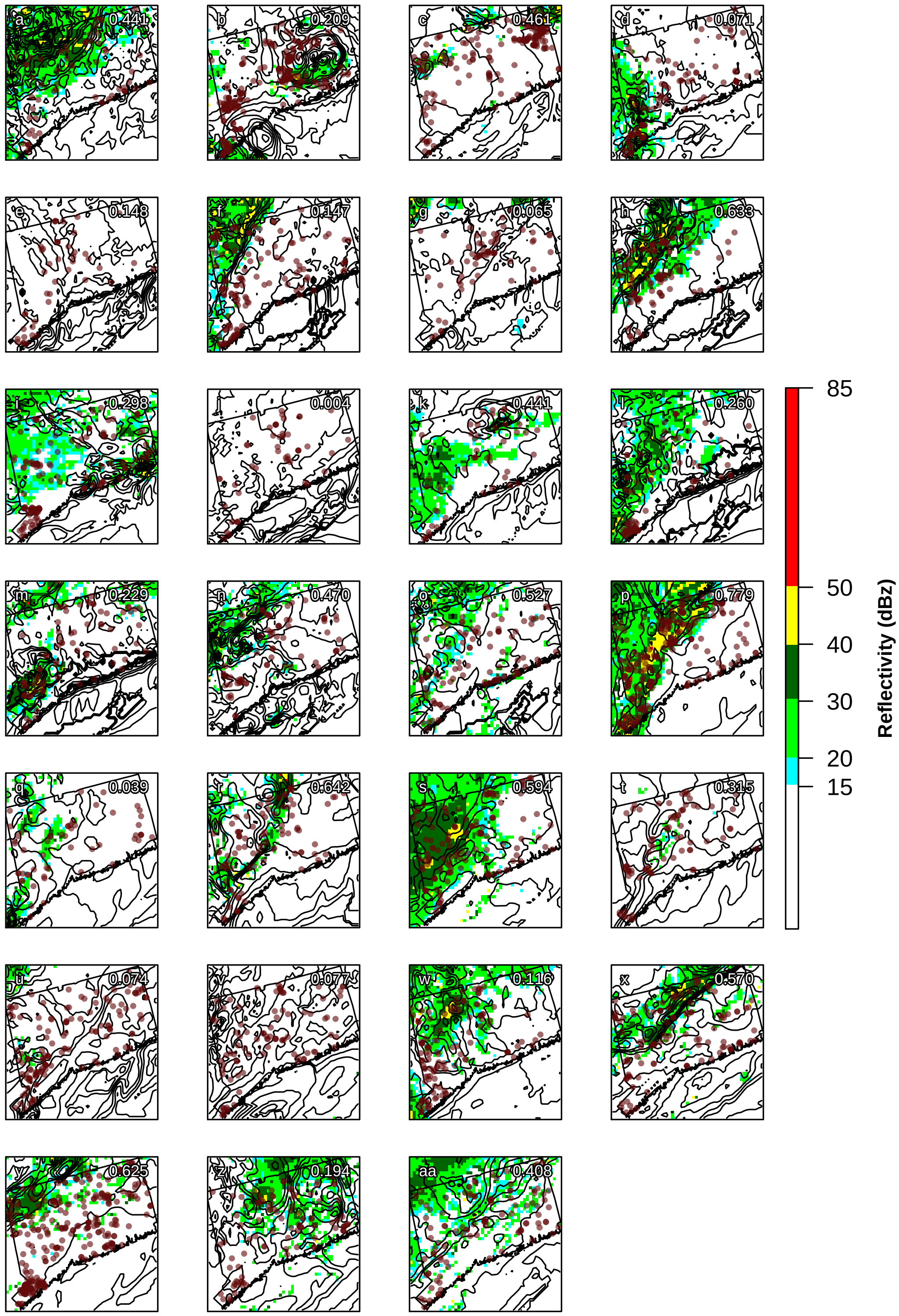

Figure 4: For each storm, the HRRR-simulated zero-hour reflectivity and contours of wind speed at the hour of peak outages, overlaid with outage occurrences accumulated over the 18-hour period. The hourly r^2 — averaged across the dynamic OPMs — is shown in the top right of each panel.

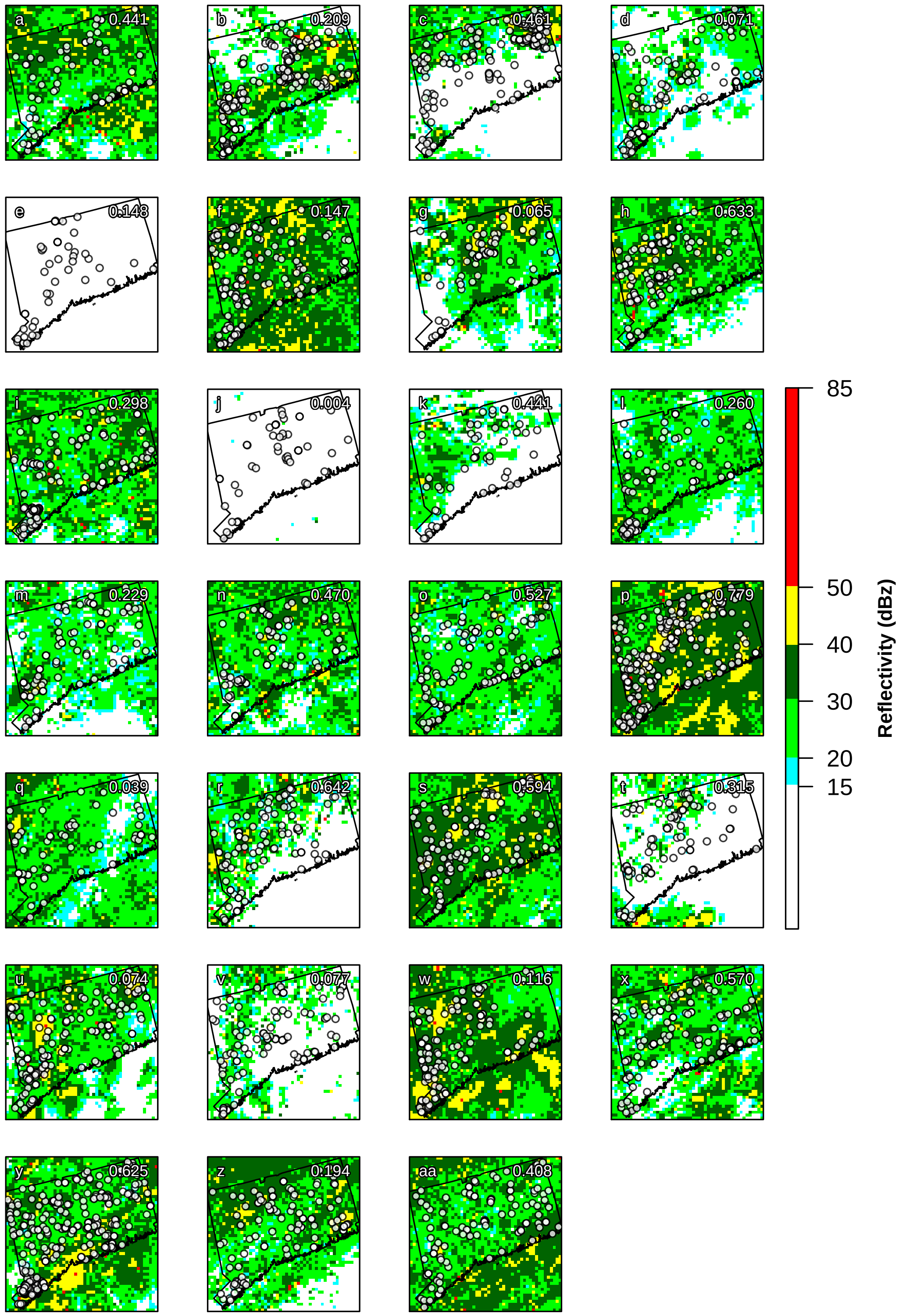

Figure 5: For each storm, the maximum HRRR-simulated zero-hour reflectivity overlaid with outage occurrences accumulated over the 18-hour period. The hourly r^2 — averaged across the dynamic OPMs — is shown in the top right of each panel.

References

[1] Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016; Volume 1.

[2] Zaytar, M.A.; El Amrani, C. Sequence to sequence weather forecasting with long short term memory recurrent neural networks. *Int. J. Comput. Appl.* **2016**, *143*, 7–11.

[3] Hochreiter, S.; Schmidhuber, J. LSTM can solve hard long time lag problems. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1997; pp. 473–479.