Convolutional LSTM for Planetary Boundary Layer Height (PBLH) Prediction

Dorsa Ziaei,[1] **Jennifer Sleeman,** [1] **Milton Halem,** [1]
Vanessa Caicedo, [2,3] **Ruben Delgado,** [2,3] **Belay Demoz** [2,3]

[1] University of Maryland, Baltimore County, Dept. of Computer Science & Electrical Engineering, Baltimore, MD 21250 USA
[2] University of Maryland, Baltimore County, Department of Physics, Baltimore, MD 21250 USA
[3] Joint Center for Earth Systems Technology, Baltimore, MD, 21250, USA
dorsaz1@umbc.edu, jsleem1@umbc.edu, halem@umbc.edu, vacaiced@umbc.edu, delgado@umbc.edu, bdemoz@umbc.edu

Abstract

We describe new work that uses deep learning to learn temporal changes in Planetary Boundary Layer Height (PBLH). This work is performed in conjunction with a deep edge detection method that identifies edges in imagery based on ceilometer backscatter signal from LIDAR observations. We implement a convolutional Long Short Term Memory (LSTM) to predict small temporal changes in PBLH estimates. In the presence of rain, clouds, and other unfavorable conditions, PBLH heights are challenging to estimate. The convolutional LSTM acts as an internal state representation of the external partially observable environment, supplementing the deep edge detection method, providing a prediction of PBLH in the absence of a reliable estimation. Convolutional LSTMs trained on image-based frames that define the movements of artifacts in the images, such as Moving MNIST digits, have been used to predict the movement of these artifacts for a set of frames in a sequence. We show how a similar network could be extended to learn more complex movement across frames and learn new information introduced at each frame. Utilizing the convolutional LSTM model with our proposed augmentation methodology applied to ten-minute frames, we predicted the change of the movement of edges identified as the PBL over time with favorable accuracy. We show the result of the prediction of PBL-based edges and evaluate the performance using three different metrics.

Introduction

The Planetary Boundary Layer is the area just above the earth's surface and is the bottom turbulent layer of the troposphere (Stull 1988). The height of the PBL, or PBLH, is identified as the top of the turbulent layer, and is used for air quality forecasting and for air pollution studies. The PBL contains most of the sources for pollution (Stull 1988). PBLH can be calculated using Weather Research and Forecasting models, radiosondes, and also using ground-based Ceilometer observing systems LIDAR technology (Danchovski et al. 2019). There are a number of complexities that hinder accurate estimation of PBLH, such as clouds and transitions from day to night. There has been an effort to improve PBLH estimations by using LIDAR backscatter profiles (Talianu et al. 2006; Compton et al. 2013; Sawyer and Li 2013; Caicedo et al. 2017; Delgado et al. 2018). In previous work by Sleeman et al. (Sleeman et al. 2020), a machine learning derived PBLH (ML-PBLH) was described based on a novel deep boundary layer edge detection method.

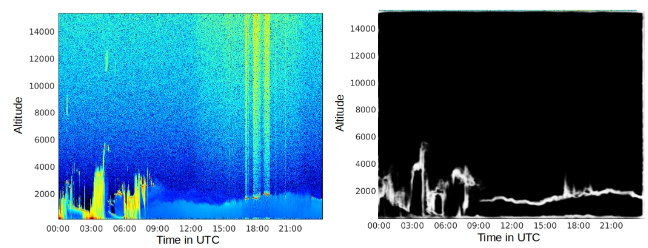

Figure 1: Lufft-CHM15K - UMBC - (left) 24 Hour LIDAR Backscatter Profiles and (right) Backscatter Image Boundary Detection (ML-PBLH)- 12/1/2016.

In Figure 1, we show an example of the backscatter profile and the edges detected for December 1, 2020, using backscatter from a Lufft-CHM15K ceilometer located at UMBC in Baltimore, MD. In Figure 2, we show the PBL heights estimated by our ML-PBLH method denoted by the magenta points. As can be seen in Figure 2, from 0:00 to 9:00 UTC the edge detection method detects erroneous points due to the presence of unfavorable conditions.

We address this problem by extending that work and by utilizing a convolutional Long Short Term Memory (LSTM) network to predict small temporal changes in PBLH estimates. Convolutional LSTMs have previously been applied to datasets, such as Moving MNIST, to identify how MNIST digits are moving from frame to frame. These datasets used sets of frames with the same MNIST digits moving around the space across the frames.

We formulate the PBLH estimation prediction as a spatio-temporal image sequence forecasting problem. In sequence forecasting, previously observed data points are used to predict a fixed length of the future data points. We create a dataset of edges based on ceilometer backscatter profiles from December 1st 2016 to December 16th 2016.

The PBL data introduces two new complexities for con-

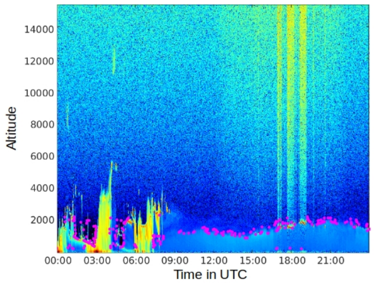

Figure 2: Lufft-CHM15K - UMBC - 24 Hour LIDAR Backscatter Profiles and PBLH Points Generated from our Backscatter Image Boundary Detection (ML-PBLH) - 12/1/2016.

volutional LSTMs: 1) the frames have more information present than datasets used in previous research, and 2) at each frame new information is introduced. Using the existing convolutional LSTM methods from previous research, when applied to the PBL data, the network was unable to learn to predict the small temporal changes. Our proposed augmentation methodology overcomes these challenges and enables the network to learn changes between frames.

Background

Developing an effective prediction model for the PBLH estimates is challenging due to its atmospheric nature and spatio-temporal characteristics. Previous studies on time series atmospheric dataset prediction have been based on conventional and mathematical approaches (Sun et al. 2014; Cheung and Yeung 2012; Reyniers 2008). The application of machine learning is a new perspective in this domain (Shi et al. 2015; Agrawal et al. 2019). A machine learning based model can be trained to predict sequences of data points in near real-time upon receiving new data, that may address the problem of continuous spatio-temporal data analysis better than traditional numerical methods. Recent advances in deep learning for sequential image prediction, such as recurrent neural network (RNN) and long short-term memory (LSTM) models (Cho et al. 2014; Donahue et al. 2015; Sutskever, Vinyals, and Le 2014; Karpathy and Fei-Fei 2017; Srivastava, Mansimov, and Salakhutdinov 2015; Xu et al. 2015) are helpful to tackle the challenge of developing an effective prediction model for the spatio-temporal datasets.

Related Work

Shi et al.(Shi et al. 2015) proposed the convolutional LSTM model for precipitation nowcasting problem. In their work, authors showed the performance of the model first on prediction of frames of Moving MNIST dataset. The moving MNIST dataset has been widely used for evaluating video prediction and image-sequence models (Srivastava, Mansimov, and Salakhudinov 2015). Then they applied their model on a radar echo dataset with 8148 training sequences and showed that they captured the motion of the clouds in images with with the end-to-end convolutional LSTM model. The radar echo dataset, includes radar maps with minimal changes between frames, in terms of shape of the clouds and spatial information.

Agrawal et al. (Agrawal et al. 2019) focused on precipitation forecasting as an image-to-image translation problem. In their paper they utilized a U-net convolutional neural network on a dataset from multi-radar multi-sensor (MRMS) system, developed by NOAA National Severe Storms Laboratory (Zhang et al. 2016).

Yao et al. (Yao and Li 2017) adopted an architecture of convolutional neural network to predict the short-term precipitation on a CIKM AnalytiCup 2017 challenge dataset including radar maps within 1.5 hours contestants (Shenzhen-Meteorological and AlibabaGroup 2017).

This study differs from previous efforts, in that we apply this method to predict small changes in PBLH over time using edge-detected imagery. We describe our methodology to address the added complexities of our data set. To the best of our knowledge this is the first time convolutional LSTMs have been used to try to predict changes in the PBLH.

Model Architecture

We utilized a convolutional LSTM architecture, proposed by Shi et al. (Shi et al. 2015). The convolutional LSTM model consists of two networks of stacked LSTM layers: an encoding network and a forecasting network. The use of convolutional layers helps to represent the features of the image sequences. The encoding network compresses the input image sequence into a hidden state tensor and the forecasting LSTM will decompress the hidden state to output the final prediction. The architecture of the model is shown in Figure 3. The power of this convolutional LSTM model is using convolution LSTM layers and designing input, hidden and output vectors as 3D tensors. Convolutional layers are known as the best representation tools, which in combination with LSTM layers perfectly captures the spatio-temporal property of the images. Encoding and forecasting with 3D tensors, where the last two dimensions show rows and columns helps to preserve all of the spatial information. Another key feature of the design is keeping the dimension of all of the states the same by using zero padding. The prediction state has the same dimension as the input state so all of the states can be concatenated in the forecasting network and fed into a 1x1 convolutional layer to generate the final prediction.

The dataset pre-processing pipeline and the model implementation have been implemented in Python using Keras and libraries such as OpenCV, Pillow and Matplotlib for visualization of the results.

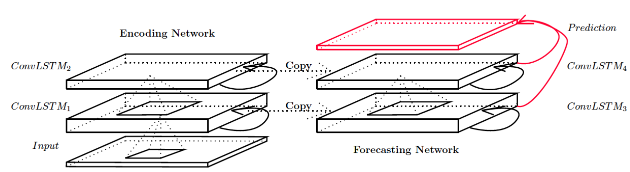

Figure 3: ConvLSTM Architecture (Shi et al.)

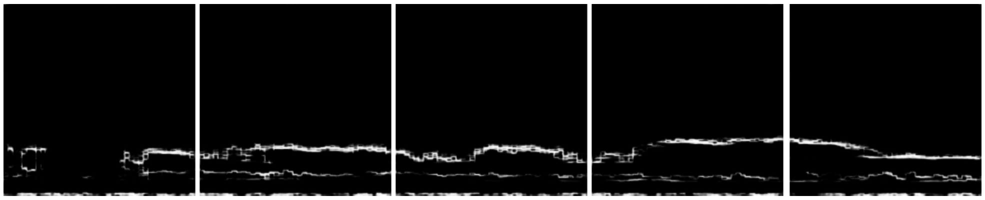

Figure 4: A sequence of 5 PBLH layer images with 10 minute time interval from the raw dataset

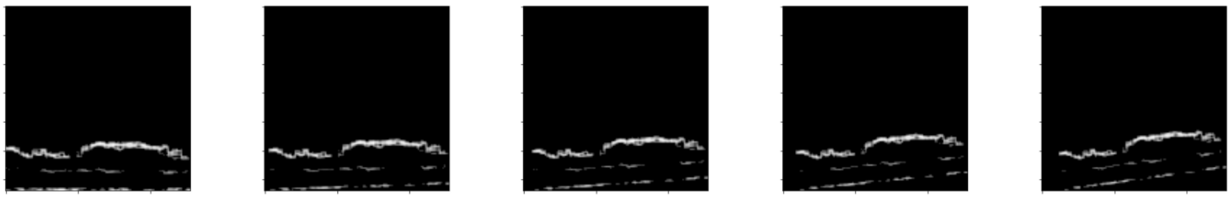

Figure 5: A sequence of 5 PBLH layer images from the synthesized dataset

Methodology

We propose a methodology to solve challenges of processing LIDAR-based backscatter profiles when unfavorable conditions are present. The PBLH edge detection dataset (Sleeman et al. 2020) is used for generating sequences of images (frames) of changing estimated PBLH edges with 10 minute time interval and by applying morphological augmentation methods to predict a given next set of frames in the sequence.

In multiple trials of training the model using the PBLH edge detection dataset, we observed that with frequent changes of shape of the line over time and missing data points due to weather condition, the model was challenged to learn these frame-by-frame changes.

To help smooth the changes between frames of the sequences, we synthesized the images in the dataset using augmentation, which led to homogeneous sequences so that the model could capture the changes in features and position of the line. We generated spatio-temporal sequences of estimated PBLH layer images, where each sequence shows the change of shape and location of the estimated PBLH edges with frames of images.

In this way, we mapped the complex estimated PBLH edge dataset to a smoother spatio-temporal dataset which enabled the convolutional LSTM model to capture the changes between frames in a sequence. With the inclusion of our methodology the network is able to predict the estimated PBLH edges.

Dataset

To study the behavior of the convolutional LSTM model, we conducted an experiment to train the convolutional LSTM model with a dataset of PBL edge detection images for forecasting next frames in the sequence. The images in this dataset are captured with 10 minutes time interval. A sequence of PBLH edge detection images in the dataset is shown in Figure 4. These images were generated using the method described in work by Sleeman et al.(Sleeman et al. 2020).

In comparison to Moving MNIST dataset, the images in PBLH edge detection dataset are frequently changing in terms of line shape and spatial information. The frames in Moving MNIST dataset contains two repeating patterns (two digits), which slightly moving in a frame. The estimated PBL present in imagery is changing shape by pattern, thickness and continuity of the line and changing location of the line in each frame, which is the biggest challenge for training an image sequence prediction model.

We structured the PBLH edge detection image dataset as sequences with five frames. In order to address the challenge of high frequency of changes in the images from frame to frame, we reduced the variance by applying augmentation on the images. We augmented each image in the dataset with morphological transformations such as rotation and shift. The variance between images (frames) in the raw dataset has been calculated as 361.575 and after augmentation the variance between images decreased to 275.700, which indicates the change between images (frames) has been decreased.

The raw images in the dataset are 885 x 656 pixels, we resized the images to different resolutions (i.e. 32x32, 64x64 and 128x128 pixels). We describe results for the 128x128 pixel images because with higher resolution images, pixelation-based issues are less prominent (no need to apply interpolation). This implies there is some sensitivity to the number of pixels, however more experimentation would be required to understand this sensitivity further.

We generated a training dataset with approximately 10k sequences and used approximately 5000 sequences for training the model and for the held out test dataset used for prediction. We trained the convolutional LSTM model with sequences of 128 x 128 pixel images. A sequence in the synthesized dataset is shown in Figure 5, which shows slight change of shape and spatial information between frames. The third frame from the sequence in Figure 4 has been selected and augmented and visualized in Figure 5, to show

how the dataset has been simplified and synthesized.

Experimental Results

As an experimental study, we trained the model for 15 epochs with "logcosh" loss function and ADAM optimizer and used the trained model as a prediction tool on the test dataset. Figure 6 shows the result of prediction on two single test sequences. In the test phase, three frames from the sequence were considered as the input to the model and prediction was performed on the next two frames. By comparing the predicted frames with the ground truth, we observe that the trained model captured the transformation of the frames as well as slight changes in the shape of the estimated PBLH edge. The model captured the spatial change in frames and predicted the next two frames in the sequence. The prediction model was successful in predicting the next frame (fourth frame). However the fifth frame's prediction could be improved. In general, as prediction frames increase, accuracy decreases. Our current focus is on improving the network for better frame prediction for multiple frames.

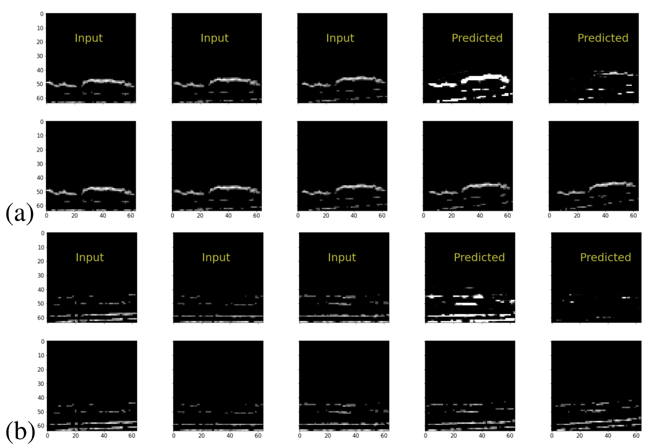

Figure 6: Predicted frames and ground truth of two test PBLH edge detection sequences (a and b) with 128x128 pixel images

For quantitative analysis, we applied the prediction model on a held-out test dataset with 3000 sequences and evaluated metrics such as accuracy, structural similarity index Metric (SSIM) (Larkin 2015) which is a metric for measuring the similarity between two images, probability of detection (POD) (Wehling et al. 2011), which is a metric to quantify the probability to detect a specific flaw, and false alarm rate (FAR) (Barnes et al. 2009) which is the number of false positives that are expected to occur in a given image. Table 1 shows the results of evaluating the model using above metric.

The SSIM metric of the table 1 shows that predicted images have 83.88 percent of similarity with the ground truth images, which shows the quality of predicted images. POD shows the accuracy of the test which is 98.10 percent for our test prediction and FAR, the number of occurred false positives is 3.89. Overall, the metrics used to measure performance of the predicted images (frames) are favorable.

Table 1: Evaluation of results on a heldout test dataset

Image size/Metrics	**Accuracy**	**SSIM**	**POD**	**FAR**
128x128 Images	97.67	83.88	98.10	3.89

Accuracy alone is not an indicative metric to evaluate the performance of a machine learning model, and additional metrics should be considered as well (Gaur 2020). We can perceptually conclude the above point by comparing the relatively high accuracy result with visualized predictions, considering the imperfect prediction for the last frame (fifth frame) in the sequence. For future work, we will consider using a larger size dataset with adjustments in length of sequences (increasing number of frames in a sequence) and tuning model parameters, to train a more generalized model for the task of PBLH prediction.

Conclusions and Future Work

In the presence of unfavorable conditions, PBLH heights are challenging to estimate. We described a convolutional LSTM that can supplement existing edge detection methods in a partially observable environment. The LSTM provides a prediction of the estimated PBLH in the absence of a reliable estimation. In this work, we described a way to apply convolutional LSTM to edge-detected PBLH backscatter output and show how our augmentation methodology can extend existing methods for predicting small changes in the estimated PBL across frames. We show how we overcame training deficiencies when the images have a significant amount of information and when new information is present in each frame. We described how we developed an image sequence dataset. The PBLH edge detection images have a lot of information change due to the turbulent nature of the PBL. Predicting the next set of frames in such a datasets is still very challenging. Our future work includes extending the model architecture and the augmentation and image transformation, as well as input sequences length and size adjustment, to be able to predict small temporal changes in PBLH estimates with more content information change.

Acknowledgments

This work has been funded by the following grants: NASA grant NNH16ZDA001-AIST16-0091 and NSF CARTA grant 17747724

References

Agrawal, S.; Barrington, L.; Bromberg, C.; Burge, J.; Gazen, C.; and Hickey, J. 2019. Machine learning for precipitation nowcasting from radar images. *arXiv preprint arXiv:1912.12132* .

Barnes, L. R.; Schultz, D. M.; Gruntfest, E. C.; Hayden, M. H.; and Benight, C. C. 2009. Corrigendum: False alarm rate or false alarm ratio? *Weather and Forecasting* 24(5): 1452–1454.

Caicedo, V.; Rappenglück, B.; Lefer, B.; Morris, G.; Toledo, D.; and Delgado, R. 2017. Comparison of aerosol lidar retrieval methods for boundary layer height detection using ceilometer aerosol backscatter data. *Atmospheric Measurement Techniques* 10(4).

Cheung, P.; and Yeung, H. 2012. Application of optical-flow technique to significant convection nowcast for terminal areas in Hong Kong. In *The 3rd WMO International Symposium on Nowcasting and Very Short-Range Forecasting (WSN12)*, 6–10.

Cho, K.; Merrienboer, B. V.; Çaglar Gülçehre; Bahdanau, D.; Bougares, F.; Schwenk, H.; and Bengio, Y. 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *ArXiv* abs/1406.1078.

Compton, J. C.; Delgado, R.; Berkoff, T. A.; and Hoff, R. M. 2013. Determination of planetary boundary layer height on short spatial and temporal scales: A demonstration of the covariance wavelet transform in ground-based wind profiler and lidar measurements. *Journal of Atmospheric and Oceanic Technology* 30(7): 1566–1575.

Danchovski, V.; Dimitrova, R.; Vladimirov, E.; Egova, E.; and Ivanov, D. 2019. Comparison of urban mixing layer height from ceilometer, radiosonde and WRF model. In *AIP Conference Proceedings*, volume 2075, 120005. AIP Publishing.

Delgado, R.; Caicedo, V.; Demoz, B.; Szykman, J.; Sakai, R.; Hicks, M.; Posey, J.; Atkinson, D.; and Kironji, I. 2018. Ad-Hoc Ceilometer Evaluation Study (ACES): Lidar/Ceilometer Mixing Layer Heights and Network. In *AGU Fall Meeting Abstracts*.

Donahue, J.; Hendricks, L. A.; Rohrbach, M.; Venugopalan, S.; Guadarrama, S.; Saenko, K.; and Darrell, T. 2015. Long-term recurrent convolutional networks for visual recognition and description. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2625–2634.

Gaur, Y. 2020. Precipitation Nowcasting using Deep Learning Techniques.

Karpathy, A.; and Fei-Fei, L. 2017. Deep Visual-Semantic Alignments for Generating Image Descriptions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39: 664–676.

Larkin, K. G. 2015. Structural Similarity Index SSIMplified: Is there really a simpler concept at the heart of image quality measurement? *arXiv preprint arXiv:1503.06680* .

Reyniers, M. 2008. *Quantitative precipitation forecasts based on radar observations: Principles, algorithms and operational systems*. Institut Royal Météorologique de Belgique Brussel, Belgium.

Sawyer, V.; and Li, Z. 2013. Detection, variations and intercomparison of the planetary boundary layer depth from radiosonde, lidar and infrared spectrometer. *Atmospheric environment* 79: 518–528.

ShenzhenMeteorological; and AlibabaGroup. 2017. CIKM AnalytiCup. URL http://www.cikmconference.org/CIKM2017/CIKM_AnalytiCup_task1.html.

Shi, X.; Chen, Z.; Wang, H.; Yeung, D.-Y.; Wong, W.-K.; and Woo, W.-c. 2015. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems* 28: 802–810.

Sleeman, J.; Halem, M.; Caicedo, V.; Demoz, B.; Delgado, R. M.; et al. 2020. A Deep Machine Learning Approach for LIDAR Based Boundary Layer Height Detection. In *IEEE International Geoscience and Remote Sensing Symposium*.

Srivastava, N.; Mansimov, E.; and Salakhudinov, R. 2015. Unsupervised learning of video representations using lstms. In *International conference on machine learning*, 843–852.

Srivastava, N.; Mansimov, E.; and Salakhutdinov, R. 2015. Unsupervised Learning of Video Representations using LSTMs. In *ICML*.

Stull, R. B. 1988. Mean boundary layer characteristics. In *An Introduction to Boundary Layer Meteorology*, 1–27. Springer.

Sun, J.; Xue, M.; Wilson, J. W.; Zawadzki, I.; Ballard, S. P.; Onvlee-Hooimeyer, J.; Joe, P.; Barker, D. M.; Li, P.-W.; Golding, B.; et al. 2014. Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bulletin of the American Meteorological Society* 95(3): 409–426.

Sutskever, I.; Vinyals, O.; and Le, Q. V. 2014. Sequence to Sequence Learning with Neural Networks. In *NIPS*.

Talianu, C.; Nicolae, D.; Ciuciu, J.; Ciobanu, M.; and Babin, V. 2006. Planetary boundary layer height detection from LIDAR measurements. *Journal of Optoelectronics and Advanced Materials* 8(1): 243.

Wehling, P.; LaBudde, R. A.; Brunelle, S. L.; and Nelson, M. T. 2011. Probability of detection (POD) as a statistical model for the validation of qualitative methods. *Journal of AOAC International* 94(1): 335–347.

Xu, K.; Ba, J.; Kiros, R.; Cho, K.; Courville, A. C.; Salakhutdinov, R.; Zemel, R.; and Bengio, Y. 2015. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. In *ICML*.

Yao, Y.; and Li, Z. 2017. CIKM AnalytiCup 2017: Short-Term Precipitation Forecasting Based on Radar Reflectivity Images. In *Proceedings of the Conference on Information and Knowledge Management, Short-Term Quantitative Precipitation Forecasting Challenge, Singapore*, 6–10.

Zhang, J.; Howard, K.; Langston, C.; Kaney, B.; Qi, Y.; Tang, L.; Grams, H. M.; Wang, Y.; Cocks, S. B.; Martinaitis, S. M.; Arthur, A.; Cooper, K.; Brogden, J. W.; and Kitzmiller, D. 2016. Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities. *Bulletin of the American Meteorological Society* 97: 621–638.